BTO Research Report No. 682

Earth Observation Data Integration Pilot Project 5 – Developing community and crowd-sourced validation of 'Living Maps'

Authors

Stuart E. Newson, David J. Turvey, Samuel Neal, Simon Gillings

Report of work carried out by the British Trust for Ornithology under contract to Defra (SD0493)

March 2016

The British Trust for Ornithology, The Nunnery, Thetford, Norfolk IP24 2PU
Registered Charity No. 216652

Scope of report
We outline findings from the project: "Earth Observation Data Integration Pilot Project 5 – Developing community and crowd sourced validation of 'Living Maps'", under contract from Defra (SD0493).

Research Team
Project lead: Dr Simon Gillings – British Trust for Ornithology
Researcher: Dr Stuart Newson – British Trust for Ornithology
Technology advisor: Dave Turvey – British Trust for Ornithology

With contributions and input from: Dawn Balmer, Andy Musgrove, James Pearce-Higgins, David Noble, Gavin Siriwardena, Mike Toms, Karen Wright – British Trust for Ornithology

Subcontractors
Sam Neal – Norfolk Biodiversity Information Service (NBIS)

Project Steering Group
Project manager: Lynn Heeley – Joint Nature Conservation Committee (JNCC)
Richard Alexander – Natural England (NE)
Colin Chapman – Welsh Government
Paul Gilbertson – JNCC
Paul Robinson - JNCC
Douglas Scott – Welsh Government

When referring to this report please use the following citation:

Newson, S.E., Turvey, D.J. & Neal, S. & Gillings, S. (2016). *Earth Observation Data Integration Pilot Project 5 – Developing community and crowd sourced validation of 'Living Maps'*. BTO Research Report 682, British Trust for Ornithology, Thetford.

Acknowledgements
Thanks to Tim Angell, Neal Armour-Chelu, Dawn Balmer, Gerry Barnes, Lindsey Bilston, Simon Bower, Anne Casey, Jenny Chamberlin, Ben Darvill, Geoff Doggett, Jake Fiennes, Katy Froud, Tim Gatti, Ross Haddow, Adam Hinchliffe, John Hiskett, Paul Holley, Martin Horlock, Tom Hunt, Tony Leech, Helen Leith, Emily Nobbs, David North, Marya Parker, Stuart Paston, Simon Pickles, Lorna Shaw, Edward Stocker, Tim Strudwick , Sarah Taigel, Heidi Thompson, Tim Venes, Doreen Wells, Steve Whitbread, David White, Russell Wilson for completing questionnaires and interviews on volunteer communities (Chapter 4; includes affiliations).

Thank to Katherine Boughey (Bat Conservation Trust), Katie Cruickshanks (Butterfly Conservation) and Hayley New (Plantlife) for providing membership and volunteering information for the spatial analysis of volunteer communities (Chapter 4).

Thanks to Alison Johnston and Mark Miller for provided statistical input in planning the power analysis (Chapter 5).

Thanks to Matt Aitkenhead, Crona Hodges, Jamie Williams and Karen Wright for completing the vendor questionnaire (Chapter 6).

Contents

i Authorship and affiliations

Stuart Newson[1], Dave Turvey[1], Sam Neal[2] & Simon Gillings[1]
With contributions and input from Dawn Balmer[1], Andy Musgrove[1], James Pearce-Higgins[1], David Noble[1], Gavin Siriwardena[1], Mike Toms[1], Karen Wright[1].

[1] British Trust for Ornithology, The Nunnery, Thetford, Norfolk, IP24 2PU, UK
[2] Norfolk Biodiversity Information Service, c/o Community and Environmental Services, Norfolk County Council, 6th Floor, County Hall, Martineau Lane, Norwich, NR1 2SG, UK

ii Abbreviations used in text

BAP	Biodiversity Action Plan, e.g. BAP habitats
BC	Butterfly Conservation
BCT	Bat Conservation Trust
BBS	BTO/JNCC/RSPB Breeding Bird Survey
BSBI	Botanical Society of the British Isles
BTO	British Trust for Ornithology
CEH	Centre for Ecology and Hydrology
EO	Earth Observation, e.g. EO products
EODIP	Earth Observation Data Integration Pilot
JNCC	Joint Nature Conservation Committee
LERC	Local Environmental Records Centre
NBIS	Norfolk Biodiversity Information Service
NBMP	National Bat Monitoring Programme
NPMS	National Plant Monitoring Scheme
OS	Ordnance Survey
RSPB	Royal Society for the Protection of Birds
TCV	The Conservation Volunteers
VGI	Volunteered Geographic Information

Executive Summary

1) Background

a) Earth Observation data offers great potential for a range of terrestrial surveillance and management issues. Living Maps – land cover maps with a focus on priority semi-natural habitats – are being developed using state of the art data and remote sensing analyses.

b) The purpose of this report is to scope out how volunteers could be engaged in the validation of the Norfolk Living Map, and how transferable proposed techniques will be to other regions of the UK.

2) The size of the task

a) Many Norfolk habitats are very rare and sampling is likely to require a stratified approach

b) Many Norfolk Living Map habitat classes are difficult to identify; some can be validated at the desk only if additional data layers are available; others can be validated in the field only at key times when botanical features are evident. A field trial is recommended to confirm ease of identification by volunteers, and to develop and test identification and training material.

c) Significant artefacts and ambiguous habitat classes require clarification before the data are fit for validation by volunteers.

3) Generic methodological issues

a) The validation task could be divided into a desk-based component and a field-based component. For the former, volunteers would validate selected individual parcels; for the latter, volunteers should be directed to validate parcels within grid squares to capitalise on travel costs and allow squares to be joined to provide continuous validation where needed.

b) The desk task should include validation of a sample of parcels for superabundant habitat classes (Gardens; Urban; possibly also Hedgerows and field margins); removing these habitats from the field task will significantly reduce the number of parcels needing to be checked and make 500-m squares a viable resolution for field-based validation.

4) Capacity of the volunteer community

a) Interviews and questionnaires were completed across a broad spectrum of groups spanning charities, councils, leisure groups, recorder networks and conservation agencies.

b) Interviews suggest that up to 3,500 volunteers may exist in Norfolk, many preferring to self-select their local area for validation (i.e. unstructured surveying).

c) Analysis of existing schemes suggest structured scheme capacity of 0.5 volunteers per 10,000 residents, rising to 1–2 volunteers for unstructured schemes. Based on the current Norfolk population (878,000) we could expect 44 volunteers for a structured scheme or 88–176 for an unstructured scheme. These figures are significantly lower than those estimated from interviews, possibly owing to differences between national and local promotion and appealing to potentially different communities.

5) Survey design and statistical power analysis

a) Both field- and desk-based validation show potential but will require different optimisation.

b) We recommend a desk-based assessment of 200–400 parcels of each habitat type, with a focus on superabundant easily identified habitats and any rare habitats that can be identified remotely with the use of additional data layers.

c) Structured field-based sampling will be required to ensure coverage of rare habitats, which will also achieve coverage of many other common and widespread habitats. In terms of grid resolution, using 500-m grid squares provides the best balance of sufficient parcels to warrant the travel without too many to make a survey impractical (provided Gardens and Urban have been dealt with at the desk).

d) As a rule of thumb, 50 squares per habitat are needed to derive a robust error estimate; more if few parcels co-occur in a square, and more if spatial autocorrelation of error is judged to be a serious issue.

e) Sufficiently precise countywide estimates of error could be produced with a sample of c630 squares selected for presence of rare habitats; this would achieve coverage of common habitats but their error estimates may be biased. Stratification by habitat is achievable at the county scale with a sample of c1700 500-m squares which is at the upper end of volunteer capacity.

f) The power analysis provides a useful analytical framework for optimising the sampling strategy once clarity has been gained on the ease of habitat identification, ideally based on a field pilot.

g) Local communities should be encouraged to undertake unstructured validation of a network of 500-m (i.e. self-selected) squares to produce local maps, with the aim of providing qualitative information on commission errors.

h) It will be important to build in procedures for collecting information that will facilitate quality control of volunteer data. This should include using multiple volunteers to validate a sample of the same habitat parcels for both the field and desk-based components, and using control sites in desk-based validation where the habitat has been validated by experts.
i) Transferability of habitat-based stratification is dependent on the number of habitat classes in future maps and how often they co-occur; high and low values respectively will inflate required sample sizes and challenge volunteer capacity.

6) Technology review

a) A questionnaire was formulated to quantify the operational, functional and distributional aspects of technological solutions offered by six vendors.
b) The gap between the project requirements and open-source products is significant and would require considerable systems development to achieve a solution. The gap between the product requirements and proprietary solutions is less, but would still require considerable development.
c) There are two potential routes to providing a solution: a) approach existing vendors that provide systems capability and work with them to extend their solution to meet the needs of the project. This may include tailoring of both software and infrastructure, or b) Take existing open source software and a vendor with the capability to extend this, as well as the infrastructure to implement the solution for the project, and commission a development and maintenance contract for the project. The resultant implementation may then be moved back into the public domain.
d) The costs for producing a solution are estimated to be upwards of £150,000 with an annual cost of at least £2,000 per year to run the system.
e) Only two existing vendors identify their capacity to scale the project from the Norfolk Living Map to a larger solution.

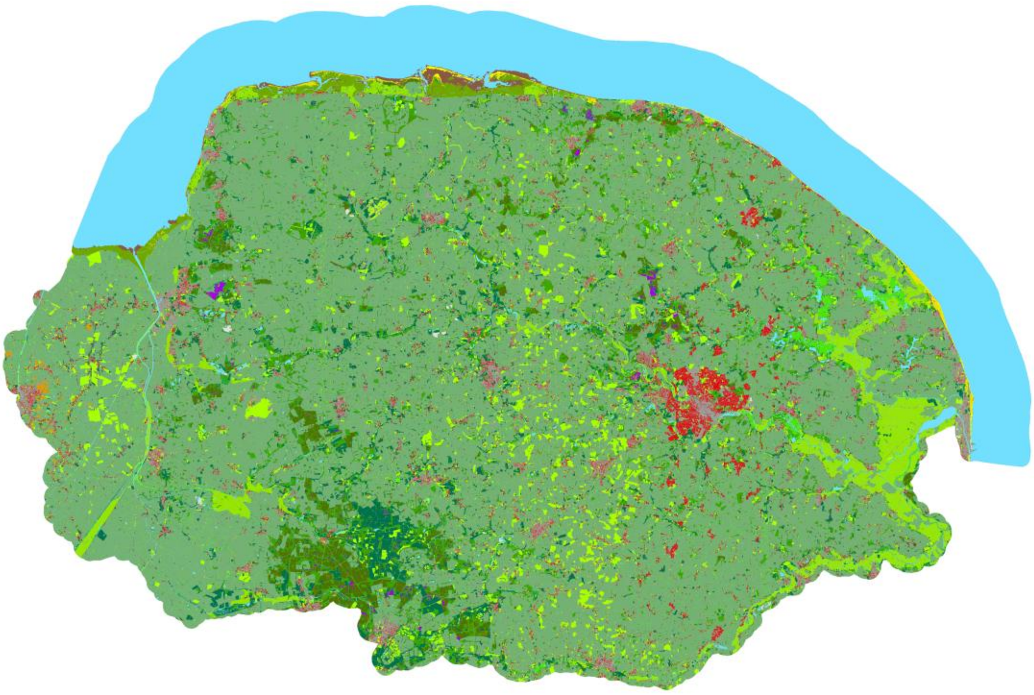
7) Communication strategy

a) A clear message on why the project is needed, why volunteers are required and how the resulting data will be used is critical.
b) Local promotion is important and should use a variety of media whilst not neglecting national volunteer groups with local presence. Promotion should focus on making people aware of the project and how to sign-up and request a square.
c) Volunteers should be provided with training and other material to maximise uptake and the quality of data collected. Volunteers should be supported with information and materials to make the process of taking part as simple and as engaging as possible, including information on interesting features to look for when visiting a location.
d) Support is critical for a project involving new methods and new volunteers. It will be important to provide engaging feedback during the project to motivate and encourage further volunteer uptake.
e) Upon completion, all data would be stored in the Collaboration Node and all users should be notified of any issues highlighted by the validation task. The latter would be easier if users had to register before getting their (free) copy of the map data.

8) Overall conclusions: feasibility of validating the Norfolk Living Map with volunteers

a) There is willingness in the local community to assist with the validation task but the level and type of activity may differ from what is required for a statistically robust validation exercise.
b) The volunteer capacity for small scale but intensive desk-based validation component exists and the broad methods are well defined. Finer detail concerning which habitats to focus on requires more information which will be best provided by pilot work.
c) The volunteer capacity for the larger field component, ideally utilising a habitat-based stratified sample of 500-m squares is more problematic and will require significant effort to recruit and train. Diverting observers from covering large self-selected areas will be important.
d) Technology is at the heart of the desk-based solution and is easily defined but costly. With an expectation of only c10 desk-based volunteers per county, development costs will look very high unless a long view can be taken on the re-use of the technology in subsequent iterations of the map and in other regions. GIS-based solutions will be cheaper to develop but more costly to implement.
e) It is our view that a smartphone application is the most effective way to complete the field-based validation but a desktop version with printable maps may aid engagement.
f) Several actions are required before validation with volunteers can begin, including resolution of artefacts, clarity of purpose and future plans, pilot fieldwork and full specifications for technology tendering.

1 Introduction



Caption: At a coarse scale the Norfolk Living Map is an impressive product, but its validity at parcel scale needs to be assessed.

1.1 Project background

Information about habitat quality and extent is required for a wide range of uses, including national statistics, estimating natural capital, for numerous conservation and research purposes and to inform local and national planning decisions. Such information can most cost-effectively be derived across large areas from remotely sensed data. Reflecting this, use of data and technology are now central to Defra's operations, and the Earth Observation Data Integration Pilot (EODIP) is the first initiative from Defra's new Earth Observation Centre of Excellence. One of EODIP's goals has been to develop new land cover maps, with a special focus on priority semi-natural habitats. Such "Living Maps" have been developed for Norfolk (Medcalf *et al*. 2014) and Wales, and tested in parts of northern England and Scotland, but there is a desire to apply this approach now more widely. Before this can be done, the Living Map requires robust validation to provide measures of accuracy and ground-truthed information to feed into future versions of the Living Map. Citizen science is seen as an effective and efficient way to achieve this at a large scale.

Citizen science has been central to biodiversity monitoring for around a century. In other fields there are many successful examples of citizen science that have led to new scientific discoveries, including unravelling protein structures (Khatib *et al*. 2011) and discovering new galaxies in the Galaxy Zoom application (Clery 2011). In the context of land cover, engaging with volunteers to validate global land cover datasets shows huge potential (e.g. Aitkenhead & Aalders 2011). In the past few years there has been a very rapid growth of interest in the exploitation of crowd-sourced data, or what is more commonly referred to as volunteered geographic information (VGI) in the remote-sensing field (Goodchild & Li 2012, Schuurman 2009). Volunteers providing the data may vary greatly from enthusiastic but naive and untrained individuals to highly skilled professionals of considerable experience. Whilst the value of such VGI is, however, often limited by concerns associated with its quality (e.g. Goodchild & Li 2012, Fonte *et al*. 2015), it is possible to build into any design, methods to formally assess data quality (e.g. Goodchild & Li 2012, Foody *et al*. 2013).

1.2 Aims, objectives and structure of the report

The purpose of this project as specified in ITT 22713 is to develop "a new approach to building volunteer support to validate Living Maps" including the provision of costed technological solutions to implement a validation task and to identify how easily such a procedure could be transferred to other UK regions. This remit does not include consideration of more traditional methods of validation, such as the employment of professional experts to complete the validation task. There was also no requirement to provide cost estimates for resourcing needed to promote and run the validation task which are likely to add significantly to the outlined technology solution costs.

Transferability is an essential thread and is considered throughout the project, where factors such as the breadth and distinctiveness of habitats, access to technology and volunteer capacity may vary. Whilst the project will consider the collection and analysis of data as part of the validation process, the scope of this project is on validation, and not on enhancing the Norfolk Living Map. In this respect the boundaries of habitat parcels are to be assumed correct.

To achieve its objectives, the study has required a number of interdependent tasks. The production of the Norfolk Living Map involved complex analyses of large data sets and it will be necessary to know where errors and uncertainties are most likely to have arisen to help focus recording effort in the locations and habitats of greatest uncertainty (Chapter 2). This will help to identify possible methods of validation considered in this scoping study, which could range from sending observers into the field to visit certain points, habitat patches or grid squares, or crowd-sourcing with individual volunteers remotely checking Living Map parcels against other data sources, such as earth observation imagery or photographs (Chapter 3). For efficient use of effort we also must assess whether data collected by surveyors who participate in existing structured schemes can be used to validate EO products.

Identifying the most appropriate approach from a statistical standpoint needs to be balanced against the likely uptake of different methods by volunteers, requiring assessment of the willingness and competence of different surveyor groups to visit new locations, differentiate among habitats and collect and submit useful data (Chapter 4). It will also be necessary to estimate the number of samples needed and both the number and distribution of volunteers to provide a statistically robust measure of the accuracy of the Living Map (Chapter 5).

Answers to such questions are contingent on how technology will be used. For example, can apps on smartphones help surveyors locate focal habitat patches and submit data? Are there ways in which people with reduced mobility may participate via the internet? This project will review existing technological tools and compare them against what bespoke systems (such as new smartphone apps) could deliver (Chapter 6). Once appropriate field methods and data collection protocols have been scoped, it will be necessary to outline a communication strategy (Chapter 7) to recruit surveyors in sufficient number and in the right places, encourage them to participate and submit data, and provide adequate feedback to ensure continued participation. One aspiration for Living Maps is that they "live" through use by stakeholders and communities, so a legacy of use is potentially important for the communication strategy.

The results of this work will assist Defra in deciding how to validate the Norfolk Living Map and in considering the potential to develop Living Maps in other regions of England. Ultimately, efficiently produced maps of the extent and condition of habitats across the country could inform a wide range of decisions from local planning to natural capital accounting.

1.3 Previous examples of land cover validation using volunteers

A number of studies have looked at the potential for crowd-sourcing to contribute to land cover map validation activities. We summarise below some of the main projects that have addressed similar or related questions, and research that has considered how best to use such data.

The Geo-Wiki project (Fritz *et al.* 2009) asks online participants to use aerial imagery via Google Earth as well as any local knowledge that the volunteer may have to make classification choices on which land cover type they are observing given a predefined classification scheme.

The Land Use/Cover Area Frame Survey (LUCAS) (http://www.lucas-europa.info/) is an example of a more authoritative, non-crowd based attempt at capturing land cover data. LUCAS, commissioned by Eurostat, deploys professional surveyors to many locations across the European Union to determine land cover/land use, record transects, and take photographs of the landscape. By virtue of LUCAS's means of data collection using professional surveyors, creating a comprehensive dataset using this method would not be possible. However, these data can be used as a means of validation of other EO products.

For the purposes of defining land cover, due to the complexity of the earth's surface, all measurements contain error to an unknown extent. It is therefore very difficult to precisely describe and categorise features of land cover. This error is true for both remote sensing classification and classification via human interpretation of aerial imagery. Foody (2002) discusses this with respect to remote sensing, emphasising that ground truth measurements are still a form of classification and thus contain some degree of error. Having error in both the land cover data and volunteered data, requires that steps be taken to ensure the methods for collecting data allow for the opportunity of the highest quality products. This means understanding people's perceptions of land cover in order to provide training or additional material to assist in the classification process.

The majority of experiments measuring quality of crowdsourced volunteered land cover classifications come from experiments run through the Geo-Wiki project (Perger *et al.* 2012, See *et al.* 2013, Foody *et al*. 2013, Comber *et al*. 2013, Comber *et al*. 2014). Perhaps most notably See *et al*. (2013) reports on an experiment in which expert and non-expert participants in a Geo-Wiki campaign were asked to classify land cover given aerial imagery for the purposes of measuring participants' accuracy rates. Three experts visually classified land cover from aerial imagery Control points generated by three experts which were used to measure how accurate the crowdsourced participants' classifications were. Averaged accuracy rates ranged from 66%-76% for the full set of participants, with experts reaching a maximum of 84%, and non-experts reaching a maximum of 65%. Comber *et al*. (2013) also uses crowdsourced classification data gathered from Geo-Wiki but focuses on the level of agreement between expert and non-expert classification of land cover type, rather than reporting accuracy rates measured against control points. They conclude by illustrating map outputs that show obvious visual differences between expert and non-expert classification choices. Comber *et al*. (2014) further states that expertise in classification has a general influence but is varied across land cover classes.

Similarly to Geo-Wiki, the OpenStreetMap (http://www.openstreetmap.org) dataset is comprised of crowdsourced geographic information that research has identified as potential data to assist, support, and validate other land use mapping projects. Arsanjani *et al.* (2013) analyzed OpenStreetMap contributions to quantify the accuracy of participants' land use (opposed to land cover) classifications in an urban setting compared to other non-crowdsourced land-use datasets.

The study concluded that OpenStreetMap, and in general other forms of crowdsourced geographic data, can be valid data sources for mapping land use.

1.4 Definitions and components of accuracy and error

There are several ways in which accuracy and error can be calculated and expressed. These are not specific to the remote sensing field and there are parallels in many fields ranging from species distribution modelling to medical research. To introduce the concepts we use a simple example involving the classification of parcels as Orchards or not (Table 1.1). This classification (or confusion) matrix shows the number of instances (or events) according to a two-way breakdown of how parcels were classified from EO data and what they were in reality (assumed to be error free, but see later). True positives are the instances where both the EO classification and field visit concur that parcels are Orchards. True negatives are the cases where truly non-Orchard parcels (e.g. woods) are classified as non-Orchard by algorithms. An algorithm that performed well would return many true positives and many true negatives. A poorly performing algorithm would also (or instead) return false positives (cases of woodlands misclassified as Orchards) and false negatives (cases of Orchards misclassified as Woodlands).

Table 1.1. A simple two-way classification matrix containing artificial data.

		Field visit confirms patch is an Orchard	
		YES	NO (it's a wood!)
EO algorithm classified parcel as Orchard	YES	True Positives (n = 90)	False Positives (n = 210)
	NO (classified as a Woodland)	False Negatives (n = 10)	True negatives (n = 690)

The numbers of parcels in each quadrant can be used to calculate different accuracy and error metrics as follows. The True Positive Rate (also known as completeness or Producer's Accuracy) assesses what proportion of the known cases are identified correctly:

$$TP\ rate = \frac{TP}{TP + FN} \qquad (1)$$

However, this ignores cases where woodlands were incorrectly assigned in the Norfolk Living Map to Orchards (False Positives). This can give us another measures of accuracy, often referred to as Precision (Correctness, Reliability or User's Accuracy), the proportion of positive classifications that are true:

$$Precision = \frac{TP}{TP + FP} \qquad (2)$$

From these two measures of accuracy two measures of error can be derived; Omission Error expresses the proportion of known cases that were missed and Commission Error expresses the proportion of identified cases that were mistaken:

$$Omission\ Error = 1 - TP \qquad (3)$$

$$Commission\ Error = 1 - Precision \qquad (4)$$

Overall accuracy can be calculated by dividing the total number of correct cases (TP + TN) by the total number of cases in the matrix (TP + TN + FP + FN), or using the following equation (cf. Rutzinger et al. 2009).

$$Accuracy = \frac{TP}{TP + FP + FN} \qquad (5)$$

Using the artificial data in Table 1.1 and focussing only on parcels classified as Orchards, we may conclude the True Positive rate is 90% (Omission Error = 10%), because 90 of the 100 Orchards parcels were identified. However, in this example, a large number of Orchards were missed by the classification (e.g. wrongly assigned to Woodland), giving Precision of 30% (Commission error of 70%). Combining these two measures in equation 5, an overall value for accuracy is produced, which for this example would be 29% (90 / 310). The example in Table 1.1 considers an overly simple two-class analysis (Orchard vs not Orchard) but the same principles apply when there are more than two classes.

For scarce habitats, it is easy to determine the 'column' error in Table 1.1, because we can focus structured sampling around rare habitats. Sampling theory can be used to identify statistical power to detect levels of error for such designs (Chapter 5). It is much harder to estimate the 'row' error, because we need to find instances of scarce habitats hidden within more common classifications e.g. a handful of orchards within a large number of woodland parcels. With limited information on likely false negative error rates it will be extremely difficult to assess statistical power for designing a robust process for assessing false negative errors.

For the power analyses in Chapter 5 we focus on estimating Omission Error and following discussion among the Steering Group we assess sampling regimes that return estimates of Omission Error within a 10% wide range, i.e. ±5% points on error estimates.

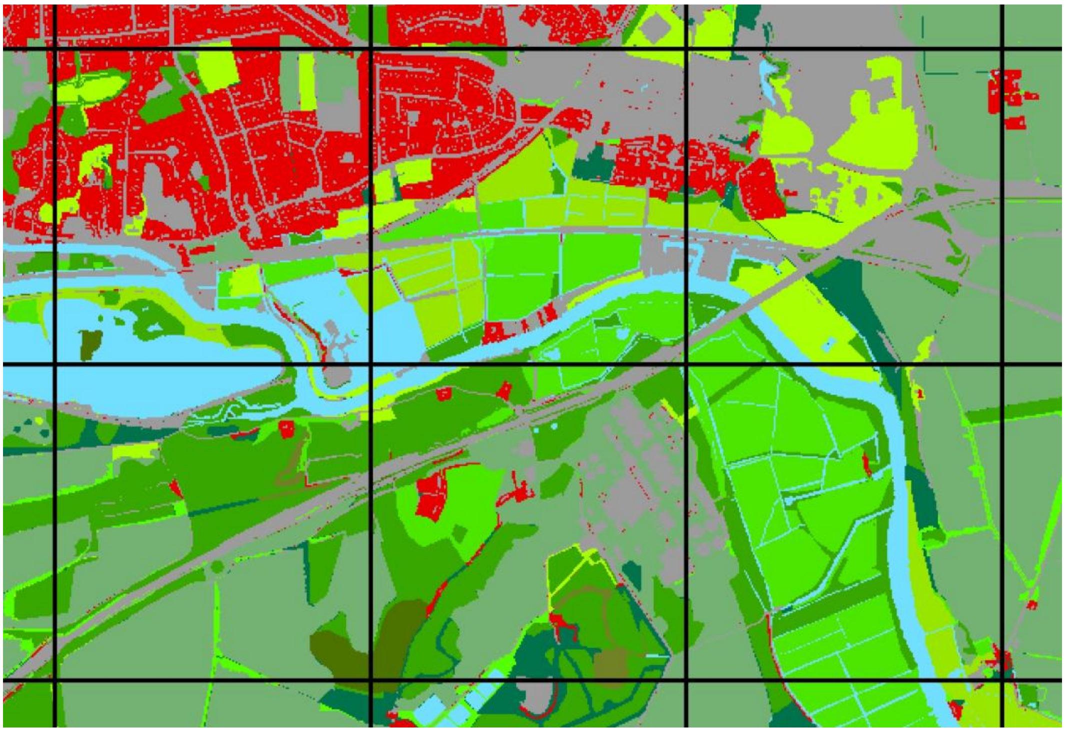
1.5 Transferability

VGI has been used around the globe including in some of the remotest locations so in principle there is no reason why a volunteer-based validation task in Norfolk could not be transferred to other regions of the UK. A key caveat concerns the precise sampling strategy devised and its volunteer resource requirements, which may be less transferable.

1.6 Conclusions and Recommendations

- There is a growing literature on the use of citizen scientists (volunteers) to provide data (volunteered geographic information).
- A full assessment of accuracy needs to consider both false positives and false negatives.
- Power analyses for structured sampling may need to focus on Omission Error.
- Commission error is likely to be difficult to assess.

2 Size of the task: preliminary analyses of the Norfolk Living Map



Caption: Living Map extract for southwest Norwich showing 1-km grid; red areas contain thousands of individual parcels of Gardens.

To be able to scope out the requirements for a robust volunteer-based validation procedure, it is essential to understand the Norfolk Living Map data. This includes understanding the number and size of parcels and how they are distributed. It requires an understanding of any issues there might be with the data, how or whether it is possible to take account of these in the validation process, and what implications these might have for the volunteer or volunteer engagement.

2.1 Satellite imagery and map production

Initial habitat classification rules were first developed by Environment Systems Ltd for case study regions as part of the Defra/JNCC-funded project Making Earth Observation Work for UK Biodiversity Conservation. These rules were then adapted and applied to the entire county of Norfolk. See Medcalf *et al*. (2014) for further information on the methodology and rules used for developing the Norfolk Living Map. Table 2.1. lists the habitats contained in the Norfolk Living Map, along with how these relate to BAP, Annex 1 or other habitat descriptions, and the broad rule base on which each habitat is based. It should be noted that many of the habitat classes in the Norfolk Living Map do not directly relate to priority habitats (of any other definition) or they include subclasses that relate to condition or fine-scale structure. Also, it is an acknowledged shortcoming of the Norfolk Living Map that the EO data did not permit the identification of some priority habitats (e.g. reedbeds). These two facts - inclusion of unanticipated habitats or habitat subclasses and exclusion of anticipated habitats - make it very difficult to assess the transferability of any proposed validation methods to new areas. In subsequent sections of the report we have assumed that all 35 habitat classes will require validation. Existing validation information is insufficiently robust to allow habitats to be removed from the task on the basis of high certainty (see Section 2.5), although habitats of low policy priority could be excluded and additional data sources could be used in an initial GIS exercise to mask out well defined habitats (see Section 2.4.1). Terminology has also proven challenging, with several of the habitat class names being ambiguous. As advised by the Steering Group, analyses in this report focus on what has referred to by Environmental Systems as the “unmerged map”. A merged map was also produced where data were submitted to a small polygon removal process, whereby parcels of some habitats that were less than or equal to $50m^2$ were reassigned to the class belonging to their nearest neighbour.

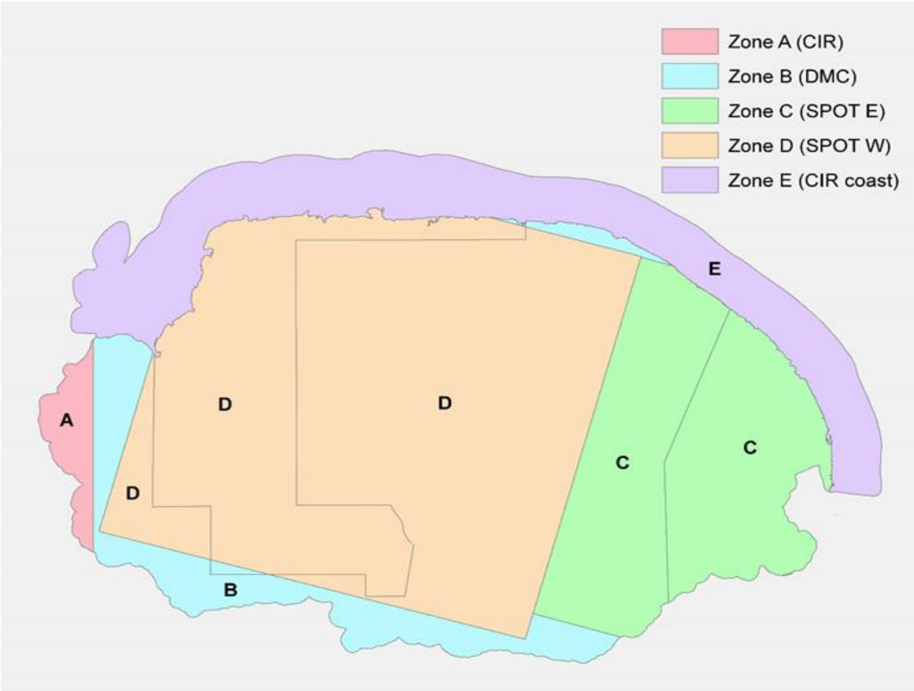
Table 2.1. The 35 Habitat classes in the Norfolk Living Map (note that a 36th class, Semi-improved (poor) was combined with Semi-improved grassland at the recommendation of the Steering Group prior to analysis. This table was produced by Environment Systems as a readme document to support the Norfolk Living Map data under contract to NBIS.

Habitat	**BAP / Annex 1 or other habitat description**	**Classification basis**
Arable	Arable.	Spectral
Bare Ground	Urban, disturbed ground.	Spectral
Beach	Sandy beach, vegetated shingle, shallow sand dune.	Spectral, Distance to sea
Bracken	Bracken.	Spectral
Coastal and Floodplain Grazing Marsh (Coastal)	Coastal and floodplain grazing marsh.	Spectral, Topographic, Distance to water
Coastal and Floodplain Grazing Marsh (high productivity)	Coastal and floodplain grazing marsh.	Spectral, Topographic, Distance to water
Coastal and Floodplain Grazing Marsh (low productivity)	Coastal and floodplain grazing marsh.	Spectral, Topographic, Distance to water
Coastal and Floodplain Grazing Marsh (medium productivity)	Coastal and floodplain grazing marsh.	Spectral, Topographic, Distance to water
Coastal Dune Heathland	Coastal Dune Heathland	Spectral, Topographic, Distance to water
Coastal Saltmarsh (established)	Coastal saltmarsh.	Spectral, Topographic, Distance to water
Coastal Saltmarsh (pioneer)	Coastal saltmarsh, coastal sediment (sand or mud).	Spectral, Topographic, Distance to water
Coastal Sand Dunes	Coastal sand dunes, vegetated shingle.	Spectral, Topographic
Coastal Sand Dunes (scrub)	Coastal sand dunes, vegetated shingle.	Spectral, Topographic
Coastal Sediment	Intertidal sediment, sand or mudflat.	Spectral, Topographic
Coniferous Plantation	Coniferous plantation.	Spectral
Dune Grassland	Dune grassland, vegetated shingle.	Spectral, Topographic
Felled Woodland	Felled woodland; primarily felled coniferous plantation.	Spectral
Fen, Marsh and Swamp	Fen, marsh and swamp, mire.	Spectral
Gardens	Gardens, small areas of parkland or improved grassland.	Mastermap
Hedgerow or Field Margin	Hedgerow, field margin.	Mastermap, Spectral, Topographic
Humid dune slacks	Dune slack, sand dune.	Spectral, Topographic
Improved (scrub)	Improved grassland or parkland	Spectral, Manual

	containing trees or scrub.	
Improved Grassland	Improved grassland.	Spectral
Lowland Heathland	Lowland heathland; acid grassland; wet grassland; fen, marsh and swamp; mire.	Spectral
Lowland Heathland (Scattered)	Lowland heathland; acid grassland; wet grassland; wet heath; fen, marsh and swamp; mire.	Spectral
Lowland Mixed Deciduous Woodland	Lowland mixed deciduous woodland, hedgerows.	Spectral
Maritime Cliff and Slopes	Cliffs and cliff vegetation, acid grassland	Spectral, Topographic, Distance to water
Orchard	Traditional and non-traditional orchards.	Spectral, Manual
Scrub	Scrub	Spectral
Semi-improved (scrub)	Scrubby acid grassland, calcareous grassland, neutral grassland.	Spectral
Semi-improved grassland	Acid grassland, calcareous grassland, neutral grassland.	Spectral
Semi-improved grassland (wet)	Acid grassland, calcareous grassland, neutral grassland, rush pasture.	Spectral
Urban	Buildings, roads, caravan parks.	Mastermap, Vectormap, Spectral
Waterbodies	Ponds, rivers, lakes, aquifer fed fluctuating water, drains, saline lagoons, sea.	Mastermap (inland) Spectral (marine)
Woodland Rides	Semi improved grassland, bracken, lowland heath.	Mastermap, Spectral, Topographic

Whilst the Norfolk habitat map was developed using an extensive image stack, the majority of the images only provided partial spatial coverage of Norfolk, with the exception of the Landsat scenes. This necessitated development of five separate classification rulesets, based on different image combinations, comprising satellite images at a range of spatial resolutions. Figure 2.1 illustrates how different EO data sources and timeframes were used in different parts of Norfolk, and how the Norfolk Living Map can be considered as comprising data from five zones, where data are likely to be most comparable within zones. These zones are important because accuracy and errors may differ among zones owing to the differing use of datasets. For example, the use of high resolution imagery in zone E facilitated the identification of strictly coastal habitats but could also have affected how well non-coastal habitats were classified in the coastal zone. Also, although both zones C and D relied upon SPOT imagery, the timing of images differed and also one spanned the Broads, the other clay soils, which may cause differences in classification accuracy. For these reasons it is likely to be necessary to derive zone-specific accuracy metrics which will likely necessitate stratified sampling within zones.

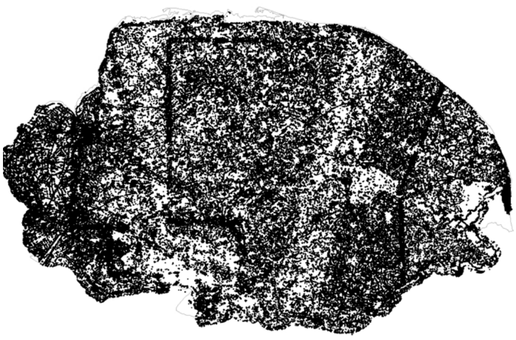
Figure 2.1. Zones defined by the classification join areas taken from Medcalf *et al*. (2014), labelled A–E. These zones are used in power analyses in Chapter 5.



It is clear from the Norfolk Living Map, that there are some artefacts in the data that in part match the boundaries of classification zones. Some of these were identified by NBIS during their appraisal work during MEOW Phase 2 and relate to artificially inserted field margins at the edges of the pilot regions (Figure 2.2a). Other issues include large-scale artefacts in the classification of Improved grassland (Figure 2.2b) and other improved grassland categories (Appendix A). Not all of these errors map directly onto the classification zones (Figure 2.1). Although individual volunteers may not see the map in its full form, so such artefacts may be less apparent, any errors of a systemic nature should be removed prior to fieldwork to ensure the volunteer resource is used effectively and to reduce loss of confidence in the product among stakeholders. Some of these errors are likely to be unique to the Norfolk Living Map – a consequence of the incremental way the Norfolk pilot map was created – but such issues should be monitored in future maps prior to entering a volunteer-based validation task.

Figure 2.2. Examples of artefacts in the Norfolk Living Map.

a) Hedgerow and Field Margins



b) Improved grassland

An issue that is Important to mention, but difficult to take account of in the validation process, is that whilst most of the images used in the remote sensing data stack were cloud-free, extensive cloud cover was present in the June and July Landsat scenes. Environment Systems Ltd hand-digitised and incorporated cloud cover extent into the classification rule-base to ensure that these images were not used where cloud was present. This has most significance for the classification of Arable and Improved Grassland, where time-series and summer imagery formed a particularly important part of the class rules. Providing separate accuracy metrics for the intersection of cloud cover and zones could inflate required sample sizes beyond the capacity of the local recorder community.

Recommendation: we recommend that professional desk-based work is required to clean the data and remove obvious artefacts before the data are used in volunteer-based validation. In terms of transferability, this may be specific to the Norfolk pilot. Confirmation will be needed as to whether low priority habitats can be excluded from the validation task.

2.2 The number, distribution and size of habitat parcels

To understand how big a job it would be to validate the Norfolk Living Map, it is important to know:

- how many habitat parcels there are of each habitat. If a habitat were sufficiently rare, it may be possible for volunteers to validate all the parcels of that habitat, but in practice more likely that there will be a range from scarce to very common habitats where the survey design will need to be sufficiently flexible (e.g. a stratified random sampling design) to allow for the targeting of rare habitats to ensure that a sufficient sample of habitat parcels of that habitat are validated.
- how many habitat parcels there are in sample areas. For example, if a volunteer were assigned a 1-km square, (sampling unit commonly used in biological recording), within which to validate all habitat parcels, how feasible would this be, and what is the range in the number of parcels that a volunteer might be asked to validate within this area? There is likely to be a compromise here between maximising the value of sending a volunteer to a remote location and what they can feasibly achieve within a reasonable time frame. This will be related to the size of the parcels. For example, if the parcel size of a particular habitat was small, and the habitat was clumped in its distribution, the job of validation may be considerably harder to perform in some parts of Norfolk than others.

The total area of Norfolk is 557,330 hectares (5,573.30 km^2) within which there are 4,441,282 habitat parcels. The average parcel size is 0.13 hectares, but size ranges from <0.01 to 753 hectares (7.53 km^2). As seen in Table 2.2. there is a large difference between how rare or common different habitats are, and ranges from Improved grassland with scrub (Improved scrub) with only 78 parcels of the habitat in Norfolk to 895,592 for Gardens and 1,936,692 of Urban.

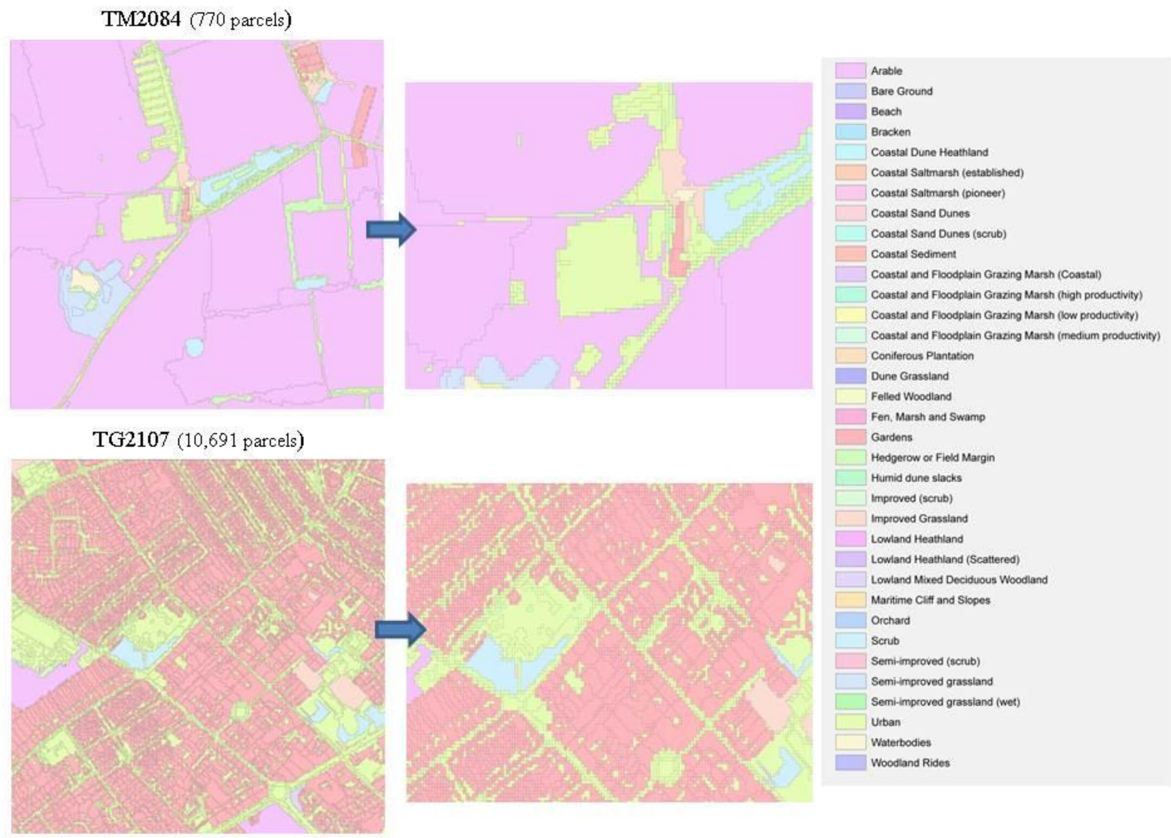
Table 2.2 Summary statistics for Norfolk Living Map data based on unmerged data. Habitats are sorted in increasing order of prevalence (number of parcels).

Habitat	Total area in	Number of	Mean parcel size in
Improved (scrub)	35	78	0.45 (0.04 - 3.08)
Humid dune slacks	14	152	0.09 (0.00 - 1.20)
Coastal Dune Heathland	13	170	0.08 (0.00 - 1.72)
Felled Woodland	282	181	1.56 (0.00 - 11.33)
Orchard	511	183	2.79 (0.12 - 33.19)
Maritime Cliff and Slopes	103	308	0.33 (0.00 - 14.18)

Lowland Heathland (Scattered)	363	352	1.03 (0.00 - 37.37)
Coastal Sand Dunes (scrub)	61	853	0.07 (0.00 - 1.79)
Beach	503	1,598	0.31 (0.00 - 70.38)
Dune Grassland	534	1,821	0.29 (0.00 - 66.93)
Semi-improved (scrub)	1,826	1,917	0.95 (0.00 - 36.15)
Bare Ground	334	2,724	0.12 (0.00 - 20.43)
Coastal Sand Dunes	528	3,200	0.16 (0.00 - 38.17)
Semi-improved grassland (wet)	69	3,467	0.02 (0.00 - 7.00)
Coastal Sediment	2,119	3,812	0.56 (0.00 - 752.74)
Lowland Heathland	831	4,109	0.20 (0.00 - 18.66)
Coastal and Floodplain Grazing Marsh	1,432	4,113	0.35 (0.00 - 123.33)
Coastal Saltmarsh (pioneer)	1,536	4,249	0.36 (0.00 - 80.22)
Fen, Marsh and Swamp	2,311	4,966	0.47 (0.00 - 136.87)
Coastal Saltmarsh (established)	2,662	6,178	0.43 (0.00 - 164.83)
Bracken	850	6,267	0.14 (0.00 - 10.78)
Coniferous Plantation	10,447	11,298	0.92 (0.00 - 93.13)
Coastal and Floodplain Grazing Marsh (low	2,002	19,432	0.1 (0.00 - 28.72)
Coastal and Floodplain Grazing Marsh (high	10,348	20,367	0.51 (0.00 - 34.06)
Woodland Rides	1,262	28,153	0.04 (0.00 - 6.27)
Coastal and Floodplain Grazing Marsh	5,773	38,443	0.15 (0.00 - 115.27)
Arable	355,220	49,527	7.17 (0.00 - 266.45)
Lowland Mixed Deciduous Woodland	31,971	68,698	0.47 (0.00 - 81.79)
Improved Grassland	25,859	83,270	0.31 (0.00 - 152.91)
Scrub	7,456	100,245	0.07 (0.00 - 27.01)
Semi-improved grassland	27,419	204,109	0.13 (0.00 - 164.68)
Waterbodies	10,006	398,440	0.03 (0.00 - 541.39)
Hedgerow or Field Margin	7,117	536,318	0.01 (0.00 - 18.32)
Gardens	15,130	895,592	0.02 (0.00 - 5.15)
Urban	30,403	1,936,692	0.02 (0.00 - 24.62)
All habitats combined	**557,330**	**4,441,282**	

Any survey method that is adopted would need to be able to deal with the different complexities of landscapes present within the county. Perhaps the most important difference can be observed between rural and urban areas. To illustrate this, Figure 2.3 shows two 1-km squares; a typical rural square (TM2084) and an urban square (TG2107). The rural square, which comprises some buildings (Urban), but is largely Arable totals 770 habitat parcels. In contrast, in heavily urban TM2107, there are 10,691 habitat parcels, which mainly comprise Gardens and Urban parcels.

Figure 2.3. Maps showing two contrasting 1-km squares, to illustrate the large difference between rural and urban areas in terms of the complexity of the landscape and the size of the validation task a volunteer could encounter if 1-km squares were used for the basis of sampling.



Recommendations: A purely random sampling design is unlikely to yield sufficient coverage of rare habitats to be able to quantify the error rate for these habitats, and some targeted sampling towards these habitats (a stratified random sampling design) will be necessary. Validating all Garden and Urban habitat parcels within an area is impractical in built up areas for all but the smallest grid resolutions. Consideration should be given to survey designs that limit field validation of these, with greater emphasis on desk-based validation where this is possible. Hedgerows and field margins is another habitat class where their inclusion may inflate the number of parcels in an area beyond acceptable levels for field-based recording.

2.3 Habitat pixellation

Parcels of the Norfolk Living Map have irregular pixelated boundaries (see Figure 2.3 and Appendix B). Even where the parcel boundaries have been derived from underlying features mapped in a smooth manner by OS Mastermap, other data limitations of the remote sensing procedure have resulted in irregular parcel boundaries. This presents three problems: 1) thin diagonal features (e.g. Roads, hedgerows) can be reduced to a string or disconnected line of single-pixels of the same habitat class, thereby inflating the number of parcels relative to the number of features that observers must validate; 2) in heterogenous landscapes, identification of parcels may be difficult in

the field; 3) validation that relies on a strong local ownership of the Map may suffer if the local community judges the Map to be substandard to existing maps (e.g. OS).
Recommendation: we recommend that some consideration is given to whether some professional desk-based work is needed to remove the pixellation, potentially using Mastermap data.

2.4 Distinctiveness of different habitats for desk-based and field-based assessment

It is essential to understand how easily different habitats could be identified by volunteers. We consider this separately for desk-based and field-based recording; the desk-based perspective would rely upon aerial photography, potentially with additional layers of information.

For the purposes of this report, for field identification, we consider participants as belonging to one of three broad groups ('experts', 'fieldworkers' and 'novices'). We define experts as skilled field recorders with high proficiency in identify habitats (probably with detailed botanical knowledge). Fieldworkers may be involved in biological recording, which could include habitat recording, but are not themselves habitat specialists (for example Breeding Bird Survey surveyors). Novices would include individuals or groups who are not normally involved in habitat recording or survey work, but have an interest in their local area (for example: ramblers).

When considering the distinctiveness of different habitats, it is important to judge whether identification is only possible from aerial photography or in the field at a particular time or times of year. In relation to future Living Maps, it is also important to consider priority habitats that are not present in Norfolk. In this report we consider here the ease by which other BAP priority habitats not present in Norfolk can be identified, but accept that a living map for another area of the country, could include Annex 1, Phase 1, land cover type or other habitat classifications.

For desk-based validation, the Crick Framework[1] is particularly helpful in identifying which priority habitats should be identifiable from earth observation data with ancillary data layers. The Crick Framework considers habitats as belonging to one of five tiers, where habitat flagged with tier 5 and probably most of the options classed as tier 4 are unlikely to be possible to classify using earth observation techniques. Additional unpublished information and comments on which to make an assessment was provided by the Steering Group members, and BTO staff independently considered the ease by which habitat identification could be made from a desk- and field-based perspective. The majority view is summarised in Table 2.3 for the Norfolk habitats and additional BAP Priority habitats. However, it is clear that more thought is needed on this, particularly for field-based validation, and whether volunteers are able to correctly distinguish some habitats in the field. This will depend in part on the quality of training and supporting materials. Targeted training material, with examples where difficulties in identification are most likely and a continual learning process should be built into the process. This would allow volunteers to reflect on the information they contribute. For desk-based validation regular feedback and evaluation of their data could be made through the use of control habitat parcels (of known habitat).

[1] http://jncc.defra.gov.uk/page-6281

Table 2.3. Identifiability of habitat classes included in (a) the Norfolk Living Map and (b) additional BAP Priority habitats from a desk- and field-based perspective.

(a) Norfolk Living Map

Norfolk Living Map habitat names	DESK-BASED			FIELD-BASED				Comments
	Aerial imagery alone	Imagery plus other layers (secondary habitat)	Seasonality of Identification	Novice	Fieldworker	Expert	Seasonality of identification	
Arable	Y	Y	Y	Y	Y	Y	Y?	Can be difficult to distinguish from early pasture reseed. Could be some seasonality depending on growth stage.
Bare Ground	N	Y	N	Y	Y	Y	N	Clear definition important
Beach	Y	Y	N	Y	Y	Y	N	
Bracken	N	N	Y	Y	Y	Y	N	May depend on density of bracken to count as "bracken"
Coastal and Floodplain Grazing Marsh (Coastal)	N	Y (Y)	Y	Y (Y)	Y (Y)	Y (Y)	Y	
Coastal and Floodplain Grazing Marsh (high productivity)	N	Y (N)	Y	Y (N)	Y (N)	Y (N)	Y	
Coastal and Floodplain Grazing Marsh (low productivity)	N	Y (N)	Y	Y (N)	Y (N)	Y (N)	Y	
Coastal and Floodplain Grazing Marsh (medium prod.)	N	Y (N)	Y	Y (N)	Y (N)	Y (N)	Y	
Coastal Dune Heathland	N	Y	N	N	Y	Y	N	
Coastal Saltmarsh (established)	N	Y (N)	N	Y(N)	Y (Y)	Y	N	
Coastal Saltmarsh (pioneer)	N	Y (N)	N	Y(N)	Y (Y)	Y	N	Clear definition of established v pioneer, but should be possible for fieldworker
Coastal Sand Dunes	Y	Y	N	Y	Y	Y	N	Is a question over how identifiable a sand dune is from dune grassland (same for below)
Coastal Sand Dunes (scrub)	Y	Y	N	Y (Y)	Y	Y	N	
Coastal Sediment	Y	Y	N	Y	Y	Y	N	
Coniferous Plantation	N?	N?	Y	Y	Y	Y	Y	Identifiable from two images – winter and summer, or maybe even just winter

Dune Grassland	N	N	N	N	Y	Y	N	
Felled Woodland	Y?	Y?	N	Y	Y	Y	N	Depends on how recent images are. May rapidly become covered e.g. bracken
Fen, Marsh and Swamp	N	N	N	N	Y	Y	N	
Gardens	Y	Y	N	Y	Y	Y	N	
Hedgerow or Field Margin	Y	**Y**	N	Y	Y	Y	N	
Humid dune slacks	N	**Y**	N	N	Y	Y	N	Potentially may be seasons – if drying out in summers makes job difficult
Improved (scrub)	N (N)	Y (Y)	Y	N (Y)	Y (Y)	Y (Y)	Y	improved grassland with scrub
Improved Grassland	N (N)	Y (Y)	Y	N (N)	Y (Y)	Y (Y)	Y	
Lowland Heathland	N (N)	Y (N)	N	Y (N)	Y (Y)	Y (Y)	N	
Lowland Heathland (Scattered)	N (N)	Y (N)	N	Y (N)	Y (Y)	Y (Y)	N	Scattered patches of heathland. A clear definition is needed for “scattered”, “scrub”
Lowland Mixed Deciduous Woodland	N	Y	N	Y	Y	Y	N	
Maritime Cliff and Slopes	Y	Y	N	Y	Y	Y	N	
Orchard	N	Y?	N	Y	Y	Y	N	
Scrub	Y	Y	N	Y	Y	Y	N	
Semi-improved (scrub)	N (N)	Y (Y)	Y	N (Y)	Y (Y)	Y (Y)	Y	semi-improved grassland with scrub. This may depend on the precise definition.
Semi-improved grassland	N (N)	Y (Y)	Y	N (N)	Y (Y)	Y (Y)	Y	This may depend on the precise definition.
Semi-improved grassland (wet)	N (N)	Y (N)	Y	N (N)	Y (Y)	Y (Y)	Y	This may depend on the precise definition.
Urban	Y	Y	N	Y	Y	Y	N	includes roads outside urban areas
Waterbodies	Y	Y	N	Y	Y	Y	N?	smaller Waterbodies may be seasonal
Woodland Rides	Y	Y	N	Y	Y	Y	N	

(b) Additional UK BAP Priority habitats

UK BAP broad habitat	**UK BAP priority habitat**	**DESK-BASED**			**FIELD-BASED**				**Comments**
		Aerial imagery alone (secondary habitat)	**Imagery plus other layers (secondary habitat)**	**Seasonality of Identification**	**Novice**	**Fieldworker**	**Expert**	**Seasonality of identification**	
Terrestrial and Freshwater Habitats									

Rivers and Streams	Rivers (updated December 2011*)	Y? (Y)	Y (Y)	N	Y	Y	Y	N	
Standing Open Waters and Canals	Oligotrophic and Dystrophic Lakes	Y (N)	Y (N)	N	N	N?	Y	N	
	Ponds	Y (N?)	Y (N)	N	Y	Y	Y	N	
	Mesotrophic Lakes	Y (N)	Y (N)	N	N	N?	Y	N	
	Eutrophic Standing Waters	Y (N)	Y (N)	N	N	Y	Y	N	
	Aquifer Fed Naturally Fluctuating Water Bodies	Y (N)	Y (N)	N	N	Y	Y	Y	
Arable and Horticultural	Arable Field Margins	Y?	Y?	Y	Y	Y	Y	N	May depend on plant growth or management.
Boundary and Linear Features	Hedgerows	Y	Y	N	Y	Y	Y	N	
Broadleaved, Mixed and Yew Woodland	Traditional Orchards	Y (N)	N	N	Y	Y	Y	N	Field validation needed for traditional orchards
	Wood-Pasture and Parkland (updated December 2011)	Y (N)	Y (N)	N	Y	Y	Y	N	
	Upland Oakwood	Y (N)	Y (N)	N	Y	Y	Y	Y	
	Lowland Beech and Yew Woodland	Y (N)	Y (N)	N	Y	Y	Y	Y	
	Upland Mixed Ashwoods	Y (N)	Y (N)	N	Y?	Y	Y	Y	
	Wet Woodland	Y (N)	Y (N)	N	N	Y	Y	Y	
	Lowland Mixed Deciduous Woodland	Y (N)	Y (N)	N	Y?	Y	Y	Y	
	Upland Birchwoods	Y (N)	Y (N)	N	Y	Y	Y	Y	
Coniferous Woodland	Native Pine Woodlands	Y	Y	N	Y	Y	Y	N	
Acid Grassland	Lowland Dry Acid Grassland	N (N)	N (N)	Y	N	Y	Y	N	Desk-based not reliable
	Lowland Calcareous Grassland	N (N)	N (N)	Y	N	Y	Y	N	Not always possible to distinguish unimproved from improved grassland
	Upland Calcareous Grassland	N (N)	N (N)	Y	N	Y	Y	N	
Neutral Grassland	Lowland Meadows	N (N)	N (N)	Y	Y	Y	Y	N	Desk-based not reliable
	Upland Hay Meadows	N (N)	N (N)	Y	N	Y	Y	Y	Fairly distinct, but may be difficult to separate without other information from flower-rich calcareous grassland

Improved Grassland	Coastal and Floodplain Grazing Marsh	N (N)	N (N)	N	Y	Y	Y	N	
Dwarf Shrub Heath	Lowland Heathland	N (N)	N (N)	N	Y	Y	Y	N	
	Upland Heathland	N (N)	N (N)	N	Y	Y	Y	N	For desk-based, not possible to tell from blanket bog, mountain heaths and willow scrub
Fen, Marsh and Swamp	Upland Flushes, Fens and Swamps	N (N)	N (N)	N	N	Y	Y	N	
	Purple Moor Grass and Rush Pastures	N (N)	N (N)	Y	N	Y	Y	N	
	Lowland Fens	N (N)	N (N)	N	N	Y	Y	N	
	Reedbeds	N (N)	N (N)	N	Y	Y	Y	Y?	
Bogs	Lowland Raised Bog	N (N)	N (N)	N	Y	Y	Y	N	
	Blanket Bog	N (N)	Y? (Y?)	N	Y	Y	Y	N	
Montane Habitats	Mountain Heaths and Willow Scrub	N	N	N	Y	Y	Y	N	
Inland Rock	Inland Rock Outcrop and Scree Habitats	Y (Y)	Y (Y)	N	Y	Y	Y	N	
	Calaminarian Grasslands	Y (N)	Y (N)	Y	N	y?	Y	Y?	
	Open Mosaic Habitats on Previously Developed Land	Y (N?)	Y (N?)	N	N	Y?	Y	N	
	Limestone Pavements	Y	Y	N	Y	Y	Y	Y	
Supralittoral Rock	Maritime Cliff and Slopes	Y	Y	N	Y	Y	Y	Y	
Supralittoral Sediment	Coastal Vegetated Shingle	N (N)	N (N)	N	Y	Y	Y	Y	
	Machair	N (N)	N (N)	N	Y	Y	Y	Y	
	Coastal Sand Dunes	N (N)	N (N)	N	Y	Y	Y	Y	
Relevant Marine Habitats									
Littoral Rock	Intertidal Chalk	N	Y	N	Y	Y	Y	Y	
Littoral Sediment	Coastal Saltmarsh	Y (N)	Y	N	Y	Y	Y	Y	
	Intertidal Mudflats	Y (N)	Y	N	Y	Y	Y	Y	

2.4.1 Desk-based identification of habitats

Views expressed in this section are heavily based on information provided by Richard Alexander (Natural England). In terms of identifying priority habitats from aerial photography, very few priority habitats can be distinguished accurately from aerial photography alone. The use of stereo aerial photography and near infrared data can help improve mapping though. For woodlands, coniferous plantations can be distinguished to a reasonable degree from deciduous woodlands but identifying individual priority woodland types is very difficult. Whilst young orchards can be readily distinguished from aerial photography, traditional orchards cannot be accurately identified without field validation. It is almost impossible to distinguish priority grasslands from aerial photography alone as they are defined in terms of plant communities. Coastal habitats have already been mapped to priority habitat level from aerial photography as part of the Regional Coastal Monitoring Programme, although this requires a reasonable to high level of expertise. Wetland habitats can be particularly difficult to distinguish, although large reed-beds can often readily be identified. In the uplands it is possible to separate grass moorland from heather dominated vegetation, but blanket bog is defined primarily in terms of peat depth rather than vegetation cover. In summary, whilst it is possible to identify probable areas of priority habitat from aerial photography there will generally be a high level of uncertainty around it. In terms of validating the Living Maps, it is more likely that aerial photography can be used to determine that a habitat parcel has been misclassified or that it 'might' be the priority habitat than to confirm that it definitely is that habitat.

For some habitats, by including ancillary data, valuable information about the spatial context of the area being mapped is provided, which can help the validation process. Although many ancillary datasets are available, the value of any specific layer will depend on the spatial resolution, the date and amount of time it will remain relevant, the quality of the data, and where the data originated. The value of ancillary data in relation to specific priority habitats is looked at in more detailed within the Crick Framework, which would guide decisions on what ancillary data were used in desk-based validation. The most frequently used ancillary datasets in support of habitat mapping include the geology, soils, elevation, slope and aspect, hydrological features. More specific ancillary data which may be used for specific circumstances include field boundaries, tidal boundaries, urban zonation and exposure for sub-montane habitats. A key concern in the use of these additional datasets is that many have been used on the production of the Norfolk Living Map. The validation task should be conducted in an independent manner. For example, cross referencing a parcel of class Waterbodies with an OS map would not provide independent validation because the OS data were a primary source in the classification of Waterbodies.

The identification of some non-priority habitats is less problematic. For Norfolk a large number of parcels have been classified as Gardens and Urban, a situation likely to be repeated in many lowland counties. These two habitat classes are easily distinguished by eye using aerial imagery, either directly from the appearance of the parcel or by interpreting the parcel with respect to its immediate neighbours, i.e. in purely spectral terms a garden may look like a priority habitat but its configuration and position relative to houses and the broader urban matrix will aid manual identification.

Conclusions: Few priority habitats can be identified from aerial photography alone, but ancillary datasets can help with the identification of some habitats. The Crick Framework provides a useful guide for deciding which ancillary data would help the validation process but for this to be an independent assessment these should not be the same ones used in the map's production. Desk-based identification of some non-priority habitats, such as Gardens and Urban should be straightforward.

2.4.2. Field-based identification of habitats

In relation to field-based validation, a field worker, with a suitable habitat key, should be able to identify and validate the majority of priority habitats, although accessibility can be a problem in some areas and habitats. We believe that with sufficient training and guidance, it should be possible to accurately identify most priority habitats, at least to distinguish broad habitat types. Classes that include aspects of quality (e.g. differing productivity levels for floodplain grasslands) may be particularly problematic, and may be best to exclude these from the field-based validation process.

Recommendation: We would recommend before any large-scale survey was introduced that a field trial was carried out to determine the ease by which different habitats can be identified by volunteers in the field. Concurrently, pilot participants should test any habitat identification material to ensure it is as streamlined and unambiguous as possible.

2.5 Previous work to ground-truth the accuracy of the Norfolk Living Map pilot

As part of the Phase 2 report (Medcalf *et al.* 2014), NBIS carried out some work to independently ground-truth the Norfolk map at sample of 622 locations (points), using additional data layers or in a small number of cases through making a field visit. Because different imagery and rule sets were used to produce the Norfolk Living Map in the east and west of the county, (Zones C and D in Figure 2.1), accuracy was considered separately by NBIS in these two areas.

Originally we anticipated that this initial work in Medcalf *et al*. (2014) would directly inform the sampling design and validation approach here, but for most habitats, apart from perhaps Arable, there were insufficient ground validation points to provide meaningful results (see Table 2.4). Some rare habitat types were not assessed at all by this method. However, as a guide, looking across the sample, it suggests that the broad level of % error for a habitat is likely to be about 22–11%, although there is the possibility that this may range widely between habitats.

Recommendation: we recommend that the habitat-specific regional measures of error presented in Medcalf *et al*. (2014) are not relied on or used in simulations to inform decisions on the required number of samples required for validation.

Table 2.4. Summary of the accuracy tables (presented here as % error) from Medcalf *et al*. (2014) for West and East Landscapes of the Norfolk Living Map. Sample size is expressed in number of parcels (identified from random points). For Coastal and Floodplain Grazing Marsh, the % error is for broad Coastal and Floodplain Grazing Marsh habitat class and not for the secondary habitat (e.g. low productivity).

Habitat	West Landscape % error rate (sample size)	East Landscape % error rate (sample size)
Arable	14% (241)	8% (59)
Bare Ground	-	-
Beach	-	-
Bracken	67% (6)	-
Coastal and Floodplain Grazing Marsh (Coastal)		

Coastal and Floodplain Grazing Marsh (high productivity)	46% (26)	8% (63)
Coastal and Floodplain Grazing Marsh (low productivity)		
Coastal and Floodplain Grazing Marsh (medium prod.)		
Coastal Dune Heathland	-	-
Coastal Saltmarsh (established)	12% (22)	-
Coastal Saltmarsh (pioneer)	77% (13)	-
Coastal Sand Dunes	42% (12)	-
Coastal Sand Dunes (scrub)	-	-
Coastal Sediment	22% (9)	-
Coniferous Plantation	25% (24)	-
Dune Grassland	50% (2)	-
Felled Woodland	100% (2)	-
Fen, Marsh and Swamp	0% (1)	0% (2)
Gardens	17% (6)	0% (2)
Hedgerow or Field Margin	-	-
Humid dune slacks	-	-
Improved (scrub)	-	-
Improved Grassland	100% (1)	-
Lowland Heathland	67% (6)	-
Lowland Heathland (Scattered)	13% (15)	-
Lowland Mixed Deciduous Woodland	41% (46)	0% (2)
Maritime Cliff and Slopes	-	-
Orchard	-	-
Scrub	64% (14)	100% (1)
Semi-improved (scrub)	-	-
Semi-improved grassland	55% (29)	100% (2)
Semi-improved grassland (wet)	-	-
Urban	-	0% (1)
Waterbodies	0% (2)	100% (1)
Woodland Rides	-	-
Across habitats	22%	11%

2.6 Transferability

2.6.1. Temporal transferability

There are a number of points to consider in relation to the Norfolk Living Map data and transferability. The first of these relates to transferability of the Norfolk data over time. There are clearly challenges in identifying some habitats from EO data, depending on the season that the image was taken, but a real change in habitat could occur between the date that the imagery on which the map was derived and the process of volunteer-based validation. In this situation, the Norfolk Living Map data and the volunteer may both be correct, but could still assign the same habitat parcel to a different habitat. Whilst determined by how regularly the Norfolk data is updated from new satellite imagery, it might be expected that for most habitats change would be small, but that this would be habitat-specific. For example Urban is unlikely to change to another habitat, although with an increasing pressure for housing and increasing infrastructure to support a growing population, particular habitats may be more predisposed to change to Urban. In addition, other habitats like Felled woodland are inherently short-lived, and change in these might be expected to be greater than some other habitats. It is important to consider that this will inflate to some degree the apparent error rate of any validation work, and may relate to real change rather than the accuracy of the Norfolk Living Map. It should be noted that some habitat categories such as Felled woodland and Bracken were defined as separate categories because it was possible to distinguish these, rather than necessarily a particular need to distinguish these. Further advice is needed regarding whether to retain these in the validation process.

Recommendation: Ideally the imagery and map would be updated at regular intervals. Whilst it may be difficult to distinguish habitat change occurring at some point since the timing of the satellite imagery, repeating the same validation process over time would allow for trends in error rate to be identified, and so allow those habitats where change is greatest to be identified.

2.6.2 Regional transferability

The habitats used in the Norfolk Living Map (see Table 2.1) do not map directly onto other lists. They are not based on a single defined habitat classification, but instead include some habitats from BAP Priority, Annex 1 and some other non-priority habitat descriptions. The Norfolk Map includes some habitats, such as Coastal and Floodplain Grazing Marsh (low, medium and high productivity), which were split into a number of new categories on the basis of what was straightforward to do from an Earth Observation perspective. Whilst it is difficult to pre-define what additional habitats could be split like the Coastal and Floodplain Grazing Marsh in other areas of the country, it is important to try and ensure as far as possible that all Living Maps can be combined and be directly comparable at some scale. There is a concern that if maps are produced independently, that a habitat map for any one area will not be directly comparable with another. For understanding the transferability of validation methods this is particularly problematic as we cannot foresee which habitats and subclasses will be used in the future.

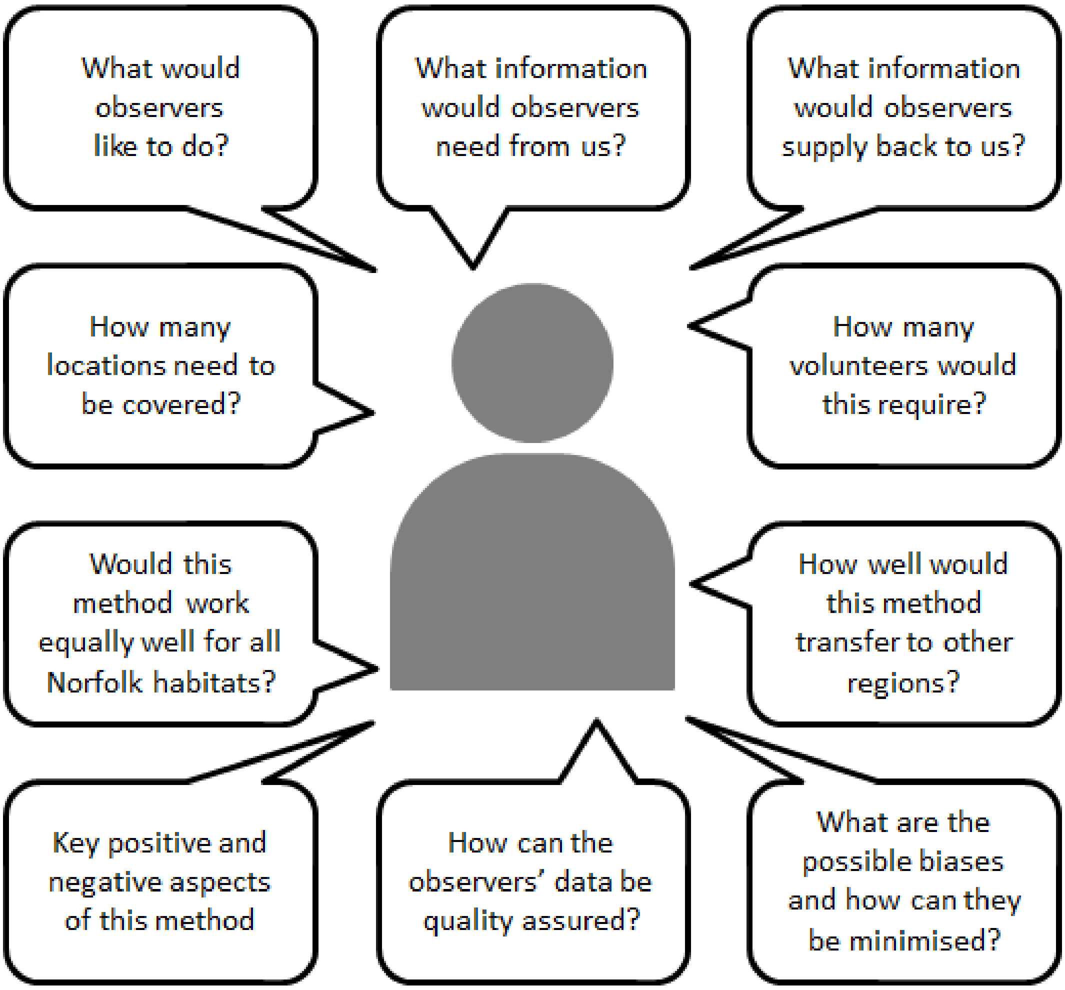
Recommendation: To facilitate the transferability of validation methods we recommend that best practice guidelines are produced to ensure consistency in habitat classification across future Living Maps.

2.7 Conclusions and Recommendations

- Professional desk-based work is required to clean the Norfolk Living Map data and remove obvious artefacts before the data are used in volunteer-based validation.
- Clarity regarding naming of habitats and unambiguous definitions of their target coverage is critical.
- Best practice guidelines are developed to ensure comparability of habitats between existing and future Living Maps.
- Accuracy metrics may be needed for each classification zone.

- Validating some non-priority habitats such as Gardens and Urban in built up areas, where there can be a large number of habitat parcels would be impractical in the field, but should be straightforward through desk-based validation.
- The Norfolk Living Map parcels are heavily pixelated which may make their identification in the field difficult, and limit the Map's appeal to community groups.
- Considering the rarity of some priority habitats, some targeted sampling towards these habitats is likely to be necessary.
- Desk-based validation should avoid using as ancillary data the same data sources used in its production.
- A field trial is required to confirm how easily different habitats can be identified by volunteers, and to help develop and test identification material and training.
- Few priority habitats can be identified from EO alone, but ancillary datasets can help with the desk-based validation of some habitats. Given training, field validation is likely to be possible for a wide range of priority habitats.
- Previous accuracy data are not sufficiently robust to guide sampling design.
- Ideally the imagery and Living Map would be updated at regular intervals, to ensure that habitat change does not unduly influence the validation process.

3 Review of survey method issues



Caption: Key questions when designing a field method.

We begin this Chapter with various definitions and key design considerations and discuss their pros and cons before moving to specifics discussion and recommendation for the Norfolk Living Map.

3.1 Structured versus unstructured survey

Unstructured methods lack a formal spatial or temporal basis to data collection (i.e. they are casual in nature) whereas Structured methods have fixed recording locations (e.g. a random sample of squares) and/or predefined recording periods and may have prescriptive in-field requirements (e.g. % cover in a 1-m quadrat).

Unstructured methods rely on people being motivated to choose a particular area or habitat and tend to have higher appeal and uptake because people are released from the constraints of visiting specific location with which they may have no personal association. A key disadvantage of unstructured approaches is that they can lead to spatial, temporal or thematic differences in coverage. The important question to consider here is whether these differences in coverage will lead to a bias in the assessment of classification accuracy. If this question was posed for biological recording, e.g. a bird survey, we might conclude that people may choose to birdwatch in areas which they know are good for birds or particular species, which will clearly lead to biased measures of abundance or trends. In the context of assessing classification accuracy there are few mechanisms by which people will be able to choose their recording locations in a manner that is correlated with classification accuracy. One conceivable way that bias could arise is if people choose to go to complex landscapes because they are inherently more interesting, and that these are also landscapes where classification error is atypical. It is likely that an unstructured method will not give the geographic or thematic coverage needed because, for example, most observers go to the coast so do not get the opportunity to validate inland habitats. There can often be value in combining an unstructured component with a structured component to get the best of both (e.g. *Bird Atlas 2007–11*; Balmer et al. 2013).

Recommendation: structured sampling is probably required to ensure coverage of rare habitats but unstructured sampling may be useful if it leads to much greater participation and volume of data and does not generate strong bias.

3.2 Field-based versus online desk-based validation

Field-based validation is where the volunteer is required to visit an area or habitat. This compares with desk-based validation where validation can be done remotely from the location. Desk-based validation can be done by professionals using GIS software but increasingly, desk-based validation can be done using online platforms. A key benefit of desk-based validation platforms is that they can be opened up to potentially very large numbers of people, allowing independent assessments of data to provide quality assurance metrics. Desk-based methods also expose scientific questions and products to different communities that are not able to participate directly in field-based survey approaches. It is also possible for people to take part at different times of day and year when it might otherwise be different or less convenient to contribute. A desk-based method may be appropriate for some habitats that can be easily identified from aerial imagery but others are likely to require field visits. A concern with desk-based methods is whether users can be constrained to use only the information provided on screen. The independence of the validation exercise could be compromised if observers use additional data and information that was used in original algorithms (e.g. OS data, other imagery layers).

Recommendation: both methods show potential, but more work is needed to assess the ease of identification of habitats remotely.

3.3 Choice of sample unit

Ultimately the project requires individual parcels be checked for the validity of their habitat classification, but there are several ways in which these could be presented, chosen by and allocated to observers. The simplest option is for observers to validate individual parcels. However, this is unlikely to be feasible for field-based methods owing to the likely travel costs of visiting an individual parcel. This is an important consideration in all regions, but particularly so in upland and remote regions. Parcel-based validation is feasible and probably preferable for desk-based methods where “travel time” is zero although, arguably, time taken to reload aerial and other map imagery onto a computer for each new parcel’s location can be an issue in areas with slow broadband connectivity.

To overcome travel time issues in field-based methods, parcels will need to be bundled together in larger spatial units, which can then be presented to the observer as a target area for validation. Several options are possible for the spatial aggregation of parcels:

- Point and radius - all parcels intersecting a circle described by a fixed radius around a point. Points could be self chosen, regularly spaced or randomly located. The ability to vary the radius allows for varying numbers of parcels per unit area. For example, in cities there are thousands of urban parcels in a 500 m radius circle whereas the same circle in the Fens may have only 10 parcels. A key disadvantage of circular plots is that they cannot be treated in a modular way to enable community groups to “stitch together” samples to validate a larger area (e.g. a parish).
- Transect - all parcels within a fixed distance of a transect. The transect could be a section of a footpath or other right of way. This approach is appealing in its focus around rights of way and may help to broaden participation to communities associated with footpaths (e.g. Ramblers, Council Trails team). There are several disadvantages: 1) footpaths follow linear features so sampled patches may be biased towards those associated with linear features; 2) footpaths are not present at the same density throughout Britain, making an approach based around footpaths less transferable; 3) long transects and their associated parcels

could be more difficult to represent and navigate through on a smartphone app than the same number of parcels in a more contained circle or square arrangement; 4) although transect sections could be treated in a modular manner to incrementally validate the whole length of a footpath or long distance path, this modularity does not extend so well in a two dimensional manner to enable coverage of parishes.

- Square - all parcels in a square of particular size. Squares could be self-chosen, regularly spaced or randomly located. The size of the square can be varied or optimised to accommodate differing parcel densities. Parcels can be allocated to squares on the basis of their centroid, meaning no patches are duplicated in adjacent squares. Squares can be stitched together to provide validation of larger areas of interest (see Appendix A for examples on a 500-m grid).

Whichever way parcels are aggregated, they should be presented to volunteers as complete parcels. That is, the boundary of a square should not be used to artificially clip a parcel. This is because the remote sensing procedure will have utilised all pixels within a parcel to assign its identity, and validation should follow the same logic to be comparable.
Recommendation: parcel-based sampling is advised for desk-based validation and square-based bundling of parcels is advised for field-based sampling. Field-validated squares have an advantage over other approaches in that they can be stitched together to validate larger areas of interest. The optimum size of grid square is considered in Chapter 5.

3.4 Targeted, regular and random sample selections

Samples of parcels or squares for coverage can be generated in several ways. We might target specific rare habitats that are unlikely to appear in random selections owing to their rarity. We might use a regular grid of squares across the county to ensure representative geographic coverage in different regions. Or squares could be selected at random, potentially in a stratified manner to ensure coverage of key regions, landscapes or classification zones. If grid squares are selected on the basis of containing rare habitats it is likely that this will also achieve coverage of many other common and widespread habitats, thereby capitalising on the time invested visiting the interesting and rare habitats. Further thought to consider the implications of this approach may be needed, if the classification accuracy of common and widespread habitats sampled close to rare habitats could be different in some way to accuracy for these habitats in the wider countryside.
Recommendation: different sampling techniques should be trialled in simulations to assess the likely coverage of different habitat classes and the power to detect error with particular precision.

3.5 Spatial independence

Sample units may contain multiple parcels of the same habitat class. We cannot assume that classification accuracies of nearby or adjacent parcels of the same class are independent. For example, nearby parcels will have been characterised using the same earth observation data, so may be similarly impacted by errors (e.g. cloud cover, edge effects). Also, the classification of one parcel could impact on the classification of its neighbour if pixels along parcel boundaries are incorrectly assigned to one parcel rather than the other.
Recommendation: Final estimates of classification accuracy and error should consider spatial non-independence. It may be better to validate many small areas than a few large areas. Overall error estimates may be produced using a mixed modelling approach incorporating random effects for grid square identity.

3.6 What to present and what to ask volunteers

For validating habitat parcels, a decision needs to be made on how much information is given to volunteers. Perhaps ideally, for an independent assessment the volunteer would be given a complete list of habitats from which to assign a habitat parcel to without prior knowledge of the existing classification. There is a compromise here between limiting the amount of work on the part

of the volunteer, and providing too much information. For example, one approach might be to provide a series of questions. For example, “we think that this habitat is arable farmland, do you agree?”. If the participant answer no, the next question is “is the habitat A, B, C or D”, which are the most likely confusion habitats, or “Other”, which brings up the complete habitat list from which to make a choice. It is important for the volunteer to be given a “do not know” option, or to provide a ranking of possible habitats. They should also be able to indicate that they were not able to assess a particular parcel (e.g. due to access constraints).
Conclusion: more work is needed on the ideal strategy for presenting questions to observers. This would be best achieved using pilot fieldwork.

3.7 Can the validation task be supplemented with data from existing schemes

For efficiency savings it is important to consider whether habitat data that is already being collected through biological recording schemes could be used in the validation task, and so reduce the sample size of new data required for the validation process. For this to be possible, three main requirements of the recording scheme would need to be met, that (a) habitat recording is carried out across a large spatial scale and collected according to a standardised monitoring protocol, (b) that the habitats recorded are comparable to those used in the Living Map or a lookup created in order to utilise existing data, and (c) that it is possible to spatially match the habitat recorded through the scheme to the Living Map habitat assigned at the parcel level. Below we consider the potential for three national monitoring schemes for birds, butterflies and plants to do the job.

For birds, volunteers taking part in the national BTO / JNCC / RSPB Breeding Bird Survey (BBS) record habitat within each of ten walked 200-m transect sections according to a hierarchical coding system (Crick 1992). BBS recording, which surveys c2,000 1-km squares annually, focuses on structural elements relevant to birds, which are unlikely to map well to priority habitats[2], and as such most of the habitats that would be needed for the Living Map would not be recorded. In addition, the primary habitat is recorded at the scale of the walked 200-m transect section (roughly 200 × 200 m), which would be impossible to match to the habitat at the parcel level except in very large areas of uniform habitat. This is illustrated with an example in Figure 3.1.

Figure 3.1. Map showing an example BBS square, where numbered 200-m transect sections, the spatial scale at which BBS habitat recording is carried out, is overlaid on Norfolk Living Map habitat parcel data. This illustrates the difficulty of spatially matching habitat recorded by BBS surveyors with Living Map data.

[2] http://www.bto.org/volunteer-surveys/bbs/taking-part/survey-methods/habitat-recording

For butterflies, the BC Butterfly Monitoring Scheme (BMS) is based on a similar design to the BBS, in that volunteers walk line transect and record butterflies. Butterflies and the primary habitat is recorded in up to 15 sections of different length, split when the primary habitat changes. The BMS focuses on habitats that are likely to be good for butterflies, but also now incorporates the Wider Countryside Monitoring Scheme which is run in collaboration with the BTO, where butterflies are recorded on BBS squares. The habitat is recorded 2.5 metres each side of the walked transect line. Habitat is recorded according to the EUNIS habitat classification, which is a comprehensive hierarchical pan-European system to facilitate the description and collection of data across Europe. The habitat categories include priority habitats, which may make comparisons with the Living Map possible. However, in the same way that walked BBS transects would be difficult to compare spatially with the Living Map parcel data, it would be difficult to spatially match habitat recorded through the BMS. There are c1500 BMS sites in Britain but they are not randomly selected and may be biased towards sites that are "good" for butterflies. As such they could provide a means to assess certain priority habitats but spatial coverage may be limited.

For plants, the National Plant Monitoring Scheme (NPMS) is a recently established habitat-based plant monitoring scheme designed by BSBI, CEH, Plantlife and JNCC to monitor the abundance of sets of species within fixed plots (squares or linear plots), in 28 habitats. Plant recording is carried out within a minimum of five 5 × 5 m or 1 × 25 m plots. It may be possible to match plant recording within some habitats to priority habitats, and considering the small plot sample areas to have a reasonable spatial match. Some work would be needed to match plant recording to the Living Map parcel level data, but with only three sites surveyed to date in Norfolk (see https://data.nbn.org.uk), this is not a feasible option for Norfolk. Whilst the number of sites surveyed through the NPMS may increase in the future, less than 90 x 1-km squares have been surveyed to date through this scheme.
Conclusion: It would be difficult for existing biological recording schemes to provide supplementary data to help validate the Living Map. It may be worth looking at the potential for the National Plant Monitoring Scheme to provide some level of validation if the level of survey coverage were to increase significantly in the future, but at the current time and for data from other schemes it would be impossible or very difficult to use data that is already collected through biological recording scheme to help validate Living Map data. It is unlikely schemes would consider adjusting their existing methods, which either have an alternate purpose or long-term comparability.

3.8 Useful metadata for assessing data quality

A key problem with Volunteered Geographic Information (VGI) is that it can be of variable, and typically unknown quality. However, the quality control of volunteer data may be facilitated if some procedures are considered during the data collection process, such as:

- Collecting information from multiple volunteers validating a sample of the same habitat parcels, at least where this is straightforward for desk-based validation, which would enable the checking of the consistency of the results. Foody *et al.* (2013) has provided an analytical approach (latent class analysis) to characterise the relative performance of each volunteer in terms of habitat class-specific and overall classification accuracy, which could be used. Asking volunteers for a confidence rating. This is very subjective, but asking volunteers to flag whenever they are not confident may help with the interpretation of data quality.
- Keeping the date of validation, which would allow information on which to help interpret discrepancies that relate to the timing of visit (e.g. seasonal difficulty in habitat identification, difference between years).
- Collecting additional information on whether the volunteer used instructions or consulted training material. This may provide indirect information about the confidence of the volunteer in assigning a habitat.
- Relevant to desk-based validation, control information for sites where the habitat has been validated by experts or selected volunteers, which could be used to assess the contributions of each volunteer (see e.g. See et al. 2013).

- Use information on where the volunteer is located, with an assumption that the closer a volunteer is to the location of the data, the more reliable the data will be. An example of the value of local knowledge is shown in De Leeuw *et al*. (2011).

Conclusion: It will be important to quality control the volunteer data. This should include using multiple volunteers to validate the same habitat, asking volunteers for a confidence rating, keeping the date of validation, recording additional information on volunteer experience or training, including control sites in desk-based validation where the habitat has been validated by experts, and using information on where the volunteer lives in relation to where the habitat is validated.

3.9 Transferability

These are mostly generic considerations and should be transferable to any region.

3.10 Conclusions and Recommendations

- Both field- and desk-based validation show potential, but more work is needed to assess the ease of identification of habitats remotely.
- Structured field-based sampling will be required to ensure coverage of rare habitats, but unstructured field-based sampling may be useful and unlikely to generate strong bias.
- Parcel-based sampling is advised for desk-based validation and square-based bundling of parcels is advised for field-based sampling. The optimum size of grid square needs further work.
- If grid squares are selected on the basis of containing rare habitats it is likely that this will also achieve coverage of many other common and widespread habitats.
- When calculating classification accuracy and error, it may be important to consider spatial non-independence of parcels.
- It would be difficult for existing biological recording schemes to provide supplementary data to help validate the Living Map.
- It will be important to collect additional metadata (see section 3.8) to be able to quality control the volunteer data.

4 Capacity of local communities to undertake different survey methods

Caption: A volunteer (Credit: David Tipling)

The potential of local communities to undertake the validation task was assessed in two ways. First, interviews (or questionnaires) were undertaken with communities and organisations thought to have an interest in the Norfolk Living Map (Section 4.1). Secondly we analysed spatial data on membership and volunteering to understand how volunteer capacity is likely to vary geographically to assess transferability (Section 4.2).

4.1 Views of potential recording communities

There are a wide range of users of the countryside who could help to validate the Norfolk Living Map. Table 4.1 lists 46 groups identified at the start of the project who represent particular interest groups, administrative authorities, charities, landowners and recording communities. A total of 257 contacts were emailed with a request for a short interview on the potential of communities or volunteers to act as habitat validators. The email requested their time for a short face-to-face or phone interview, at their convenience, over a period of 2–3 weeks set out in a Doodle Poll, for which they were required to complete their availability. All respondents were contacted and times and format for the interview were agreed. Roughly a week after the email for interviews was sent it was agreed by the project team that a questionnaire would also be sent out, asking the same questions, so that those who did not want to be interviewed or did not have the time could answer the same questions at their leisure. Thirty-eight individuals representing 21 of these groups responded either by agreeing to be interviewed or to complete a questionnaire containing the same set of questions. Full interview and questionnaire responses are available in Appendix C. The remainder of this section provides responses to each question of the interview/questionnaire.

Table 4.1. Communities and volunteer groups approached for interview, the individuals who agreed to be interviewed (I) completed a questionnaire (Q) with the same questions.

Communities, groups and NGOs	Person responding	Position, affiliation or regional representation	I/Q
All Local Environmental Records Centres (LERCs) in the eastern region, plus selected others	Steve Whitbread	Northamptonshire Biodiversity Records Centre (NBRC) Manager.	I
	Simon Pickles	North & East Yorkshire Ecological Data Centre (NEYDC)	I
Anglers/Boaters/Canoe clubs/Sailing clubs/Boat houses	Tim Gatti	Burnham Overy Boathouse	Q
Anglian Water and Northumberland Water			
Association of Local Environmental Records Centres (ALERC)	Tom Hunt	National Coordinator	I
British Trust for Ornithology (BTO)	Ben Darvill	Scotland	I
	Dawn Balmer	UK	Q
District Councils: Broadland, Breckland (Capita), Norwich City and the Fringe Project, Great Yarmouth, North Norfolk, South Norfolk, Kings Lynn and West Norfolk	Paul Holley	Natural Areas Officer (Norwich City Council)	I
Broads Authority	Beth Williams	Volunteer Coordinator	Q
Butterfly Conservation			
Centre for Ecology & Hydrology (CEH)			
Centre for Environment, Fisheries and Aquaculture Science (Cefas)			
Campaign to Protect Rural England (CPRE)	Helen Leith	Branch Manager, CPRE, Norfolk	I
Eastern Inshore Fisheries and Conservation Authority (IFCA)			
Easton College			
Environment Agency			
Essex Wildlife Trust	Dr Lorna Shaw	Biological Records Officer	I
Forestry Commission	Neal Armour-Chelu	Ecologist	Q
Hawk and Owl Trust			
Internal Drainage Boards (IDBs)			
Landowners / estate managers, including Holkham	Jake Fiennes	Estate Manager, Raveningham	I
	Ross Haddow	Senior Adviser for Stody Estate Ltd	I
Local conservation and community groups and friends of groups	Lindsey Bilston	Quaker Wood Community woodland	I
	Marya Parker	Norwich Community Green Gym	Q&I
	Tim Angell	Litcham Common Conservation Group	I
Marine Management Organisation (MMO)			

Ministry Of Defence (MOD)			
National Farmers Union (NFU)			
National Trust			
Natural England			
Norfolk and Norwich Naturalist Society, marine and terrestrial recorders.	Stuart Paston	Norfolk & Norwich Naturalists' Society	Q
	Tony Leech	Chairman, Norfolk and Norwich Naturalist Society (NATSOC)	I
	Doreen Wells	County Ant Recorder	I
Norfolk and River Waveney Trusts	Geoff Doggett	Vice Chair, River Waveney Trust.	I
Norfolk Association of Local Councils			
Norfolk Bat Survey	Stuart Newson		Q
Norfolk Biodiversity Partnership	Anne Casey	Coordinator	I
Norfolk County Council, including the North Norfolk Coast Area of Outstanding Natural Beauty (AONB)	Dr Gerry Barnes	Former Head of Environment	I
	Tim Venes	Norfolk Coast Partnership Manager (AONB)	I
	Russell Wilson	Senior Trails Officer (infrastructure)	I
	Adam Hinchliffe	Assistant Trails Officer	I
	Martin Horlock	Senior Biodiversity Officer	I
	David White	Senior Green Infrastructure Officer	I
	Edward Stocker	County Ecologist	I
	Jenny Chamberlin	Norfolk County Councillor for Diss & Roydon	I
Norfolk Farming & Wildlife Advisory Group (FWAG)	Heidi Thompson	Business Manager	I
Norfolk Wildlife Trust	David North	Head of People and Wildlife	I
	John Hiskett	Senior Conservation Officer	I
	Emily Nobbs	Conservation Officer	I
Parish Councils	Simon Bower	Clerk to Snettisham Parish Council	I
Plantlife			
Royal Society for the Protection of Birds (RSPB)	Katy Froud		Q
	Tim Strudwick	Mid Yare Reserves Site Manager	I
Suffolk County Council			
Suffolk Wildlife Trust			
The Wildlife Trusts (National)			
University College London (UCL)			

University of East Anglia (UEA)	Sarah Taigel	School of Environmental Science	I
Wildfowl and Wetland Trust			
Woodland Trust			

4.1.1 How many active volunteers do you have in Norfolk?

The interviews and questionnaires identified 3,563 active volunteers in Norfolk, although an unknown number of these may be duplicated across groups. The average number of volunteers per group or organisation was 159 (median 41). When asked how this number of active volunteers compared with other parts of the UK, the largest proportion of respondents did not know (58%) or that it was higher than in other parts of the UK (21% of respondents).

4.1.2 How many members or supporters do you have in Norfolk?

Over 41,468 members or supporters in Norfolk were identified, although an unknown number of these will be duplicated across groups. The average membership was 4,430 and median 200. This provides context to the capacity available in Norfolk and shows that the median and total active volunteer numbers are less than 10% of the overall membership. When asked how this number of members or supporter compares with other parts of the UK, the largest proportion of respondents did not know (74%) or thought it was high (16% of respondents).

4.1.3 Where are your volunteers' or members' activity distributed in Norfolk?

When asked how their volunteers' or members activities' were distributed in Norfolk the responses were: mainly widely distributed across the county (12 respondents); mostly in and around Norwich or other large towns (9 respondents), or close to where they volunteer or survey (9 respondents a local site or reserve, and 7 respondents a local parish). Some groups also commented that they had few volunteers in the west of the county where fewer people live compared with the rest of the county.

4.1.4 If you have active volunteers, what time or times of the year are they active?

When groups with active volunteers were asked about what time of year their volunteers are active, 21 responded in the spring–summer, 17 responded all year round, and 9 in autumn–winter. In terms of volunteer activities, the autumn–winter period mainly involved reserve or habitat management, such as scrub bashing, and the spring–summer mainly survey work. Some respondents noted that different people volunteer on weekdays compared to weekends. Specialism is also important. For example few botanists will be out surveying in November, which is the main time for lichenologists.

4.1.5 Are your volunteers engaged in structured recording, and if so does this include habitat recording?

Many (23 of 39) of the respondents felt that there were at least some volunteers who undertake structured recording, either through their organisation or through being volunteer surveyors for other organisations. Very few respondents indicated that all their volunteers or volunteer work was structured. Typical examples were almost exclusively birds (e.g. BTO Breeding Bird Survey and Nest Record Schemes), butterflies (Butterfly Conservation transects) and plant monitoring; most other groups were involved in unstructured recording.

Norfolk Wildlife Trust gave evidence of one couple who do 80 hours a week structured recording, but that probably only 20–30 out of all their active volunteers do any. However a new County Wildlife Sites Action project should generate over 100 volunteers active in structured recording. A number of wildlife trusts and LERCs now undertake projects that involve some structured recording.

Councillor Jenny Chamberlin felt that there was a need for a questionnaire to see if volunteers have done this type of recording before, not necessarily as a volunteer. A number of people felt that although this was true, volunteer coordinators and experts will know their volunteers and how they tend to record. A number of respondents explicitly said or inferred that “weeding out” unreliable volunteers was important.

Tim Strudwick (RSPB) felt that structured recording probably accounted for a third of site volunteers and that many are already involved in annual monitoring and some site-specific monitoring e.g. Fen Raft Spiders and Swallowtail monitoring, via a timed search method of randomly selected points for abundance and occupancy.

In general, limited habitat recording is currently completed by volunteers, such as Local Wildlife Sites monitoring, elements of simple habitat recording for Pondnet and bird surveys (BTO) or ditch flora and sward structure surveys (RSPB). Where respondents said yes to this (12 of 39 respondents) they detailed only limited examples. It has been quite clear that few volunteers are carrying out habitat surveys, almost exclusively if they are undertaking management and, or species recording.

4.1.6 Would they consider making additional visits to known or new areas to undertake map validation?

On the whole respondents felt that it may possible for their volunteers to make additional visits to known or to new areas, but for a limited or short timeframe so not to distract and/or put people off from taking part in existing and well established surveys. A separate initiative or survey was recommended, or setting up new volunteer roles specifically for habitat validation. When respondents were asked whether they thought their volunteers would consider making additional visits to areas that the volunteer knew (which were usually local to them) to undertake map validation, 67% said yes. A number of respondents commented that it would be important to have staff time to coordinate volunteers to do this. A number of respondents (FWAG, Stody and Raveningham estates and River Waveney Trust, amongst many others) emphasised how important landowner and land or estate managers could be in this process, who all know their land or their clients’ land and could assess large areas very quickly. Catchment partnerships in the Broads and Waveney areas were seen as examples of how landowner communication and working together for the common local benefit could maximise large-scale validation.

Asking volunteers to go to new areas was regarded as less appealing, although it would depend on the travelling time involved. 58% of respondents thought that their volunteers or supporters might consider making visits to new areas. Access to transport, particularly for individuals reliant on public transport, was seen as vital in assigning locations to people.

Sarah Taigel (UEA) suggested that results from her PhD pilot work indicated that people tend to over-report in familiar areas, but are considerably less willing to record in areas that they do not know, resulting in less data for these areas and spatially bias coverage. It is essential to have clear and simple messages in terms of what is expected of the volunteer.

Tony Leech (NATSOC) felt for the Naturalists’ Society a coordinator would be needed to keep people interested and encourage them. Providing a worthwhile and structured activity would be really useful for the society.

A number of volunteer groups undertake roving surveys or roving management of different sites; it was probably a 50/50 split for those groups as some would be excited to be pointed to new sites, but many green gym or workout groups would not want to stray from the sites they know.

4.1.7 Is your community familiar with the region's habitats?

Given sufficient training or supporting material, it was felt that almost all volunteers would have no problem with the broad, common, easy to identify habitats. Almost all respondents 28 of 33 respondents replied that this would be possible. This is most likely to be true in the area local or of particular interest to the volunteer. Training would build confidence, giving volunteers the skills and interest to get more people involved. Training a small number of individuals so that they could do this on their own and then pass on to others in their group, known as hierarchical training, was seen as important to many groups.

Training cannot be overemphasised as this was a constant theme in responses. People will do a job more enthusiastically if trained, and if they can see the results of their work. Clarity of what the information is being used for and providing feedback is vital, such as knowing how important the work is in a UK context, the idea of contributing to a greater whole and sense of purpose is important.

Tester or sign-up sessions could bring together potential groups of volunteers to see who is ready to commit and assess knowledge levels; these could then be followed up with formal training or workshops. Combining this with working in groups could work for some types of volunteers or volunteer groups. Some people would want to go into greater habitat classification detail and it was recommended that this should not be discouraged from this if they want to record more.

Some respondents (especially TCV and Norwich City / Fringe project) felt that knowledge varied enormously, often along socio-economic factors or according to an existing specialism (whether a birder or botanist). Often knowledge can differ as the motivation for volunteering differs; some for example just want to see their local area clear of litter.

Some respondents felt that the level of knowledge for validating habitat was low (such as NWT and Green Gym or workout groups), and guessed that at about 5% would be familiar (50 volunteers in total). RSPB felt that although their survey and conservation management volunteers would be knowledgeable, their people focused volunteers would be roughly 50%.

4.1.8 What proportion of your community would consider validating the Norfolk map in their local area?

As detailed in previous questions, volunteers are most likely to be interested in their local area. This is the area that people know and feel engaged with. However involvement in habitat validation would depend on how much effort and time the process of habitat validation would take up, the flexibility of what is involved, where the locations are and the volunteer commitment. The local area needs to be defined, e.g. within 2 miles or walking distance. Geoff Doggett (River Waveney Trust) felt that 80% of volunteers would want to validate habitat within walking distance. Local groups are more interested in their local area and have a lot of good local knowledge and appreciate being asked to be involved. It is very important how you ask the questions though. Some respondents felt a survey, questionnaire or trial of volunteers was needed to properly answer this question.

Some felt this may appeal more to general wildlife trust type volunteers, or NATSOC members that are not typically species recorders and would be interested in a different focus. Many felt that this may appeal to a set of completely new volunteers who are not already signed up for current schemes.

Sarah Taigel (UEA) felt that Geograph is a good approach to be followed with completing the picture for your local area (filling in the gaps). Citizen science tools and workbooks can be very useful in incentivising and showing volunteers on a website what they are working towards is very important. Sarah suggest that you could have the following grid square categories for your sign up locations:

Unallocated, allocated and completed. If people are allocated to a square but they have not completed their square, they would feel pressure to finish and be the last piece of the puzzle.

Again the importance of landowners and managers was emphasised. Ross Haddow (Stody) could do 10 parishes in an evening, and so could a handful of land managers across Norfolk and you could have a large proportion of what you need validated without the need for much other input. Ross has several hundred parcels and depending on the amount of parcels required for checking he could do this very quickly.

4.1.9 What proportion of your community would consider validating the Norfolk map in an area of our choosing?

There was little interest in this offer, especially in Norwich and other towns where people's home range is that much narrower. Three respondents felt that lack of transport or requiring transport would preclude this type of work. Although some felt if it was only 5–10 miles from home or up to 15 minutes drive, then there would be an interest, but others felt that would only be the case if they were on their way to somewhere anyway.

The BTO and others felt that it would be much better to allow volunteers to choose from a list of randomly selected areas to visit, rather than assigning people locations.

A few respondents felt that although small in number, some may be stimulated by going outside of their local area, or would do this as part of their working day. Doreen Wells felt that is was not age related, but personality, these type of people want to learn about new areas, pushing their boundaries and enjoy the challenge. Although small in numbers these type of volunteers should be cherished as they are likely to undertake huge amounts of validation due to their drive for this.

4.1.10 What proportion of your community would consider (online) desk-based validation of the Norfolk map?

The response to desk-based validation was quite negative, whereas maps on the wall and group discussions or going out in the field were seen as much more appealing. TCV groups have large proportion of volunteer who do not have computers or access to the internet and so this would not be possible. However there are some key exceptions to this:

It was suggested that people with reduced mobility, particularly those with an interest in conservation or IT, could be a good source of volunteers for the desk-based work. Engaging with such audiences might require different formats (e.g. online tutorials) and through different channels (e.g. charities such as MIND). Many thought that there might be new or different online audience and potential volunteers were out there.

Another exception is a group of currently desk-based, graduate or HQ office volunteers for organisations such as LERCs, Wildlife Trust and RSPB. Although the majority of their volunteers would not be interested in being involved in desk-based validation, the small number could get through a large amount of validation in a very short period. Again, landowners and land managers were seen by a number of respondents as fitting into this category - it is in their own interests to identify the best wildlife features, plus most farmers (99%) are in the new Countryside Stewardship scheme. Schools and universities would be another key exception to this.

4.1.11 What proportion of your community would be happy to use a well-designed smartphone app to do the validation?

Many questioned would not be interested in a smartphone app or the interest would be limited compared to a website version where maps could be printed. This may be due to high proportions of volunteers being retired (or over 50s), who in few cases have smartphones. Tim Strudwick (RSPB) said that apps alienate a generation, e.g. 55 years+, which is c75% of RSPB Strumpshaw volunteers.

Most respondents felt that the younger people would be happier to use Apps. However this was not without exception as some felt things were changing and that it was more 50/50 now that older people are getting used to smartphone technology.

Sarah Taigel (UEA) said from her research it was more about the design of the app rather than the technology putting people off. Geoff Doggett (River Waveney Trust) gave a similar response. Design was seen as very important for all respondents and making sure it is simple to use and understandable. Signal problems, sunlight on the screen and battery life were seen as big drawbacks for apps on smartphones. Sarah Taigel (UEA) suggested that it is known that GPS resolution gets worse as battery dies. Despite this, most people felt you should have both apps and a website platform: if both exist you are maximising chances of contribution. There was not completely glowing praise for websites that are the main platform as there were some points made about the potential data input issues and staff resource required where surveyors undertook the survey on printed out maps, but expected others to input to the website for them.

Making sure volunteers get feedback through whatever technology is very important. Gamification was seen as a potential feedback and incentivising tool. Certification could be used to motivate volunteers.

Several people felt that cross verification is important both for field-based and desk-based methods – e.g. If all volunteers do their own squares and then check a proportion of other volunteers' squares. Having a group of verifiers and requiring people to attach a level of confidence against their validation work were seen as important for quality assurance.

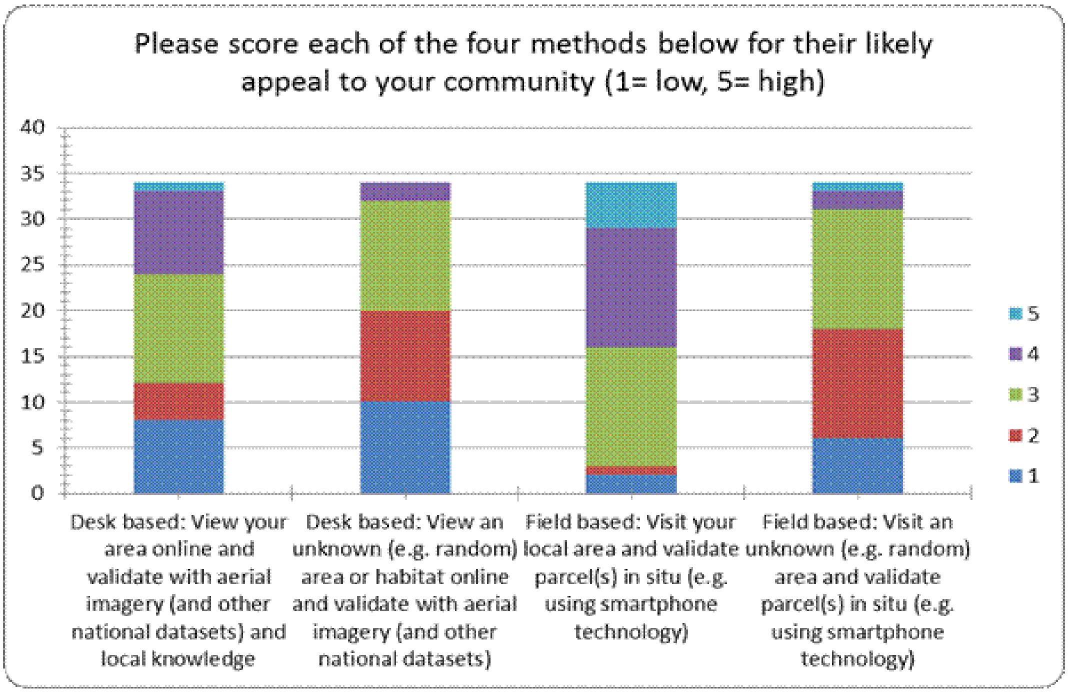
4.1.12 Appeal to volunteers of different survey approaches

Respondents were asked to score on a scale of 1 to 5 each of the four survey methods in relation to their likely appeal to their community. These included desk-based validation of their local area, desk-based validation of an unknown area, field-based validation of their local area, and field-based validation of an unknown area.

A field-based method in their local area was most attractive. The other options were of less interest to volunteers, with perhaps more interest in desk-based local validation. In addition, the following comments were made:

> may be necessary to put aside time and money to pay people to cover areas of map that volunteers do not want or would be difficult to cover; there is likely to be distance decay effect with a field-based random approach; there are a lot of things that stop people going out, even if they are enthusiastic; volunteers may be interested in recording additional information such as habitat condition as part of the part of the validation process.

Figure 4.1.1. Frequency distributions of scores assigned to four different survey methods.



4.1.13 How might your community benefit from the Norfolk Map data?

Although a handful of respondents (all species recorders) felt that the habitat validation would be of no benefit to their volunteers, a number of potential uses and benefits were identified:

> may encourage recorders to visit a wider area than at present; better quality remote-sensed data would be of value to all sorts of analyses; for monitoring change in priority habitats; to just produce a map, for the local group or parish / village website or wall; general health benefits of being outside; to feed into parish neighbourhood plans, green space planning, green infrastructure plans, parish asset management; encourages development or discourages development in other areas; may influences responses to planning applications or issues; could link in with old maps or aerial war time photos; for strengthening links between the natural environment and people in local area; would give the group more cohesion and another reason for the group to meet and a new dimension to encourage new members; provides a sense of ownership after validating the map; being able to contribute such a project; to help direct habitat creation or stewardship if they manage these kind of areas; for ecological network mapping; to support grant applications; benefits to farmers to assess their land and make sure that the map is not misused in the planning process.

4.1.14 Would your organisation be happy to help promote and support this initiative, either locally or nationally?

Without exception, people were happy to promote and support this initiative, where it was felt that requests for help were more likely to be better received if they came from the organisation or group that the volunteer supports. However, time resource was noted as being required for some respondents: For example a coordinator for NATSOC was seen as important, as coordination through the society would have more appeal and success in getting sign up from members. Local Environmental Records Centres involved in the project would be helped if some staff time were covered. Currently, most groups have no spare capacity to do any of this work without funding. Staff are needed to cover and coordinate volunteers.

4.1.15 What would be the best way to promote this activity to your community to get significant uptake?

When groups were asked what the best way to promote habitat validation to their community would be, most respondents thought that promotion should be via all possible routes and methods to maximise reach, but should be completed in a hierarchical fashion where promotion is tailored to the specific needs of the volunteers. Many thought that success will be down to how the initiative is sold, and ease of taking part. The following were suggested:

> Social media (Facebook and Twitter); conservation organisations internal communication via a group email directly to volunteers; blogs; promotion through talks and volunteer open days; organisation web sites; RSPB regional newsletters and local newsletters for reserves; Openstreet map; flickr groups that share maps; forums; crowdsourcing platforms to appeal to a new volunteers; group training and workshops; Norfolk Biodiversity Partnership community directory; local press, radio and TV; Norfolk Coast Guardian; NWT *Tern* newspaper; Parish newsletters and magazines; Norfolk and Norwich Naturalists Society magazine *Natterjack* and mailings throughout the year; LERC E-bulletins; advertise new volunteer vacancies if needed for intensive desk-based validation; parish and town council meetings and open days; include in Norfolk Bat Survey pack – have you considered validating habitats while you are doing bats?; Wild about Norfolk and other local wildlife events; promotion at visitor centres; University of East Anglia and Easton College promotion; farm walks; emails from FWAG; Farming pages in EDP; *Farmers weekly*; via Anglian farmers groups; NFU and CLA.

4.1.16 More generally, can you suggest any ways to promote this activity to new volunteers, or the general public

When respondents were asked of ways of promoting this activity to new volunteers or to the general public, the responses were very similar to those in 4.1.15. Additional points noted include:

> that it could build in social incentives to rank surveyors; could tap into school leavers before going to university, which is something that the RSPB has done with some success; A-level groups; promotion at Libraries, Community Centres, Village Halls via leaflets or posters; BBC SpringWatch; holiday makers; tourist information centres; Caravan sites; ALERC promotion; Defra biodiversity news.

4.2.17 Are there activities that your community already undertake regularly that habitat validation could fit into?

For most groups there were activities which habitat validation could fit into, although some respondents felt that it would be a distraction from those activities and reduce their impact, quality and longevity. Specifically, it could fit into biological recording, walks, BirdTrack and other surveys that people are doing regularly such as Wetland Bird Survey counts, wintering bird surveys. It may be possible to fit in with reserve managers works on RSPB reserves; school visits; river guardians in Norfolk and Essex; FWAG and UEA, potentially with MSc students working in validation to their projects; volunteers that already report on stretches of trails, Living Landscape Surveys; to include in current or new HLF projects; HQ volunteers who are regularly desk-based - it would be possible to advertise for some volunteers to do some desk-based validation at the Norfolk Wildlife Trust HQ; residential volunteers; walking festivals.

4.1.18 Can you think of other groups of people locally, regionally or nationally that you think we should also be interviewing?

When respondents were asked what other groups should be interviewed as potential sources of volunteers, the following suggestions were given:

> Orienteering groups; HLF projects; Geo-cachers; Voluntary Norfolk; Helen Ollet-Nash at Norfolk Association of Local Councils (Parish Councils); Greenlight Trust; U3As; Water

authorities; Local Nature Partnerships (Wild Anglia); Schools; Field Studies Council for different locations; Friends of the earth; Wensum Alliance; CLA; groups involved in landscape history and old maps; Norfolk archaeological trust (NAT); East of England apple and orchards project; Countryside restoration trust; new sorts of volunteers – e.g. AVIVA; Reserve managers.

4.2 Analysis of UK-wide volunteer capacity

4.2.1 Analysis of membership and volunteering data

The effectiveness of any volunteer-based survey when transferred to another region will depend on a number of factors, but primary among them will be the number of likely participants which, in turn, is likely to be a function of the number of residents. Although volunteers can be encouraged to travel to survey locations remote from where they live, participation is highest in the vicinity of the home. For example, around one third of volunteers for *Bird Atlas 2007–11* submitted data only for their home 10-km square, with a further 19% submitting data for their home square plus its immediate neighbours (Balmer et al. 2013). Therefore, understanding how the number of volunteers per 10-km square relates to the total resident population of a square is a useful step in understanding the likely volunteer capacity in areas outside Norfolk.

This requires spatial information on the density of volunteers and of the wider public. In Britain, human population data are collected as part of the National Census and typically presented in small area units which do not directly relate to grid squares. However, the 2011 National Census data have been converted to a 1-km grid and combined with data from other countries to produce a Europe-wide gridded human population dataset[3]. We extracted these data, reprojected and summarised them to give estimated human population per 10-km of the British and Irish national grids. Then, for several organisations and survey schemes (Table 4.1), we used supplied postcodes to assign all supporters and survey participants to grid squares. This enabled us to relate the number of supporters and volunteers per square to the human population. We estimate the average volunteering rate (expressed as number of volunteers per 10,000 residents) as the slope of a linear regression of volunteer numbers on resident numbers. The statistical residuals of this relationship (how far each square lies from the best fit line) indicate whether volunteering in a particular square is relatively high or low given the square's resident population. We map these residuals to reveal areas of relatively high or relatively low volunteering in Britain.

The graphs in Figure 4.1 show the relationships for each organisation/scheme. Note that for individual organisations, some volunteers are members, and some participate in more than one scheme/survey. In general, activity was correlated with human population, with membership and volunteering rates (volunteers or members per 10,000 residents) varying between 0.2 and 3.3. Differences in participation between BBS (a structured survey) and BirdTrack (an unstructured survey) are revealing, but note that participation in Bird Atlas, which had both structured and unstructured components was almost as high as for BirdTrack. The comparison between the National Bat Monitoring Programme and the Norfolk Bat Survey is interesting, with a twelve-fold difference in uptake, suggesting that surveys with the same aim or focus can have vastly differing levels of uptake depending on aspects such as site selection, skill requirements, technology and engagement. Of note is that the engagement in the latter is local in nature

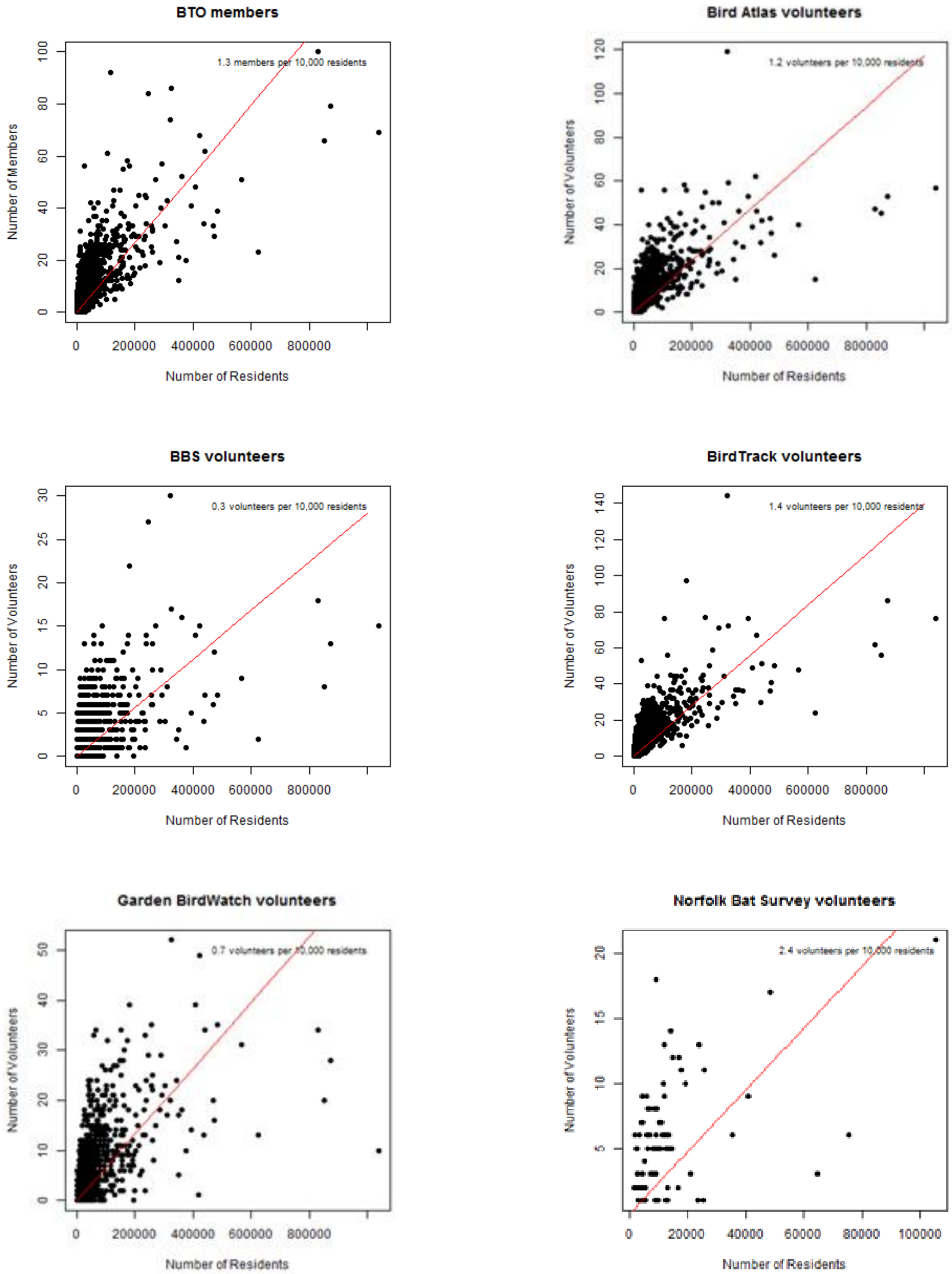
[3] EuroStat; GEOSTAT 2011: http://ec.europa.eu/eurostat

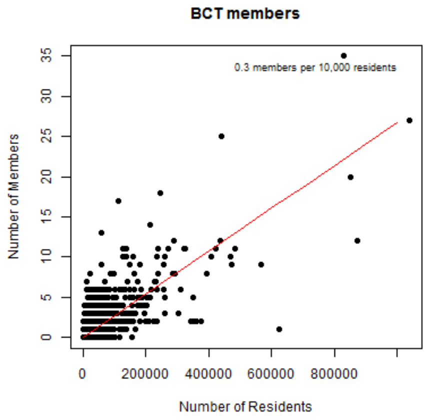
Table 4.1. Organisations and schemes for which we extracted supporter and volunteer locational information.

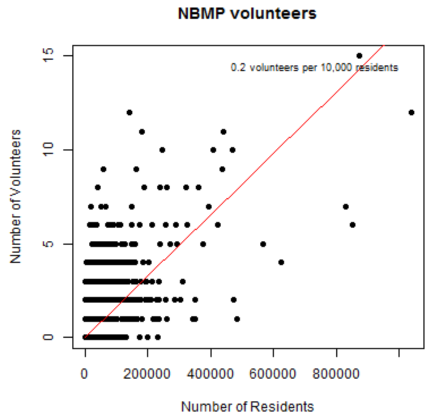
Organisation	Type of interaction	Period
BTO	Membership	2016
	Bird Atlas volunteer	2007–11
	Breeding Bird Survey volunteer	2015
	Garden BirdWatch volunteer	2016
	BirdTrack volunteer	2010–15
	Norfolk Bat Survey	2013–15
Butterfly Conservation	Membership	2016
	UKBMS volunteers	Recent
Bat Conservation Trust	Membership	2016
	National Bat Monitoring Programme	2010–15
Plantlife	Members	2016
	Wildflower Count	Recent
	National Plant Monitoring Scheme	2015

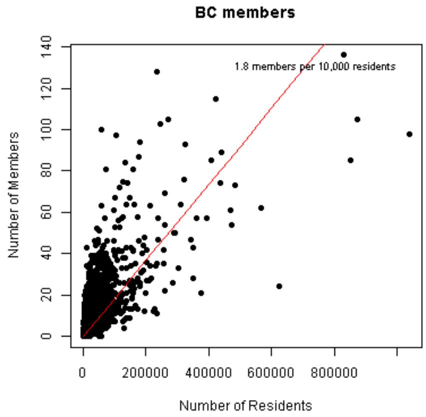
.

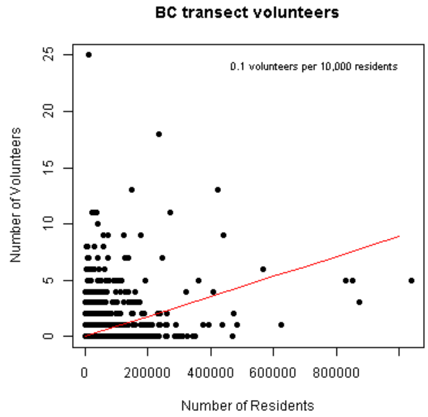
Figure 4.1. Relationships between the number of people living in a 10-km square and the number supporting an organisation or participating in a citizen science scheme.

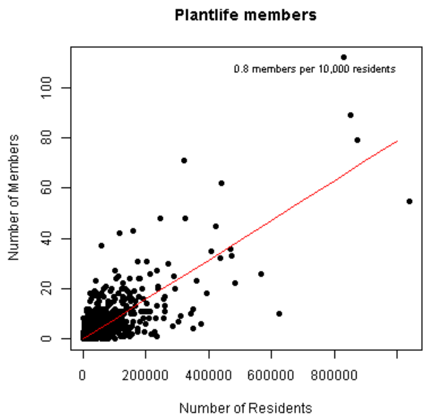


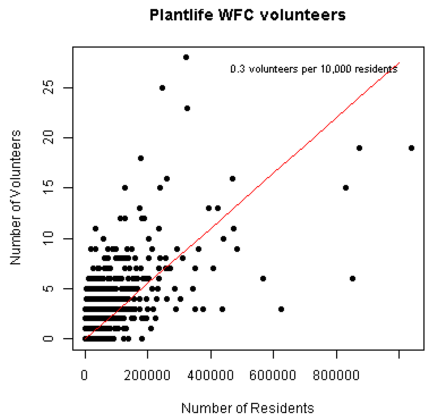

BCT members
0.3 members per 10,000 residents
Number of Members
Number of Residents
0 200000 400000 600000 800000
0 5 10 15 20 25 30 35


NBMP volunteers
0.2 volunteers per 10,000 residents
Number of Volunteers
Number of Residents
0 200000 400000 600000 800000
0 5 10 15


BC members
1.8 members per 10,000 residents
Number of Members
Number of Residents
0 200000 400000 600000 800000
0 20 40 60 80 100 120 140


BC transect volunteers
0.1 volunteers per 10,000 residents
Number of Volunteers
Number of Residents
0 200000 400000 600000 800000
0 5 10 15 20 25


Plantlife members
0.8 members per 10,000 residents
Number of Members
Number of Residents
0 200000 400000 600000 800000
0 20 40 60 80 100


Plantlife WFC volunteers
0.3 volunteers per 10,000 residents
Number of Volunteers
Number of Residents
0 200000 400000 600000 800000
0 5 10 15 20 25

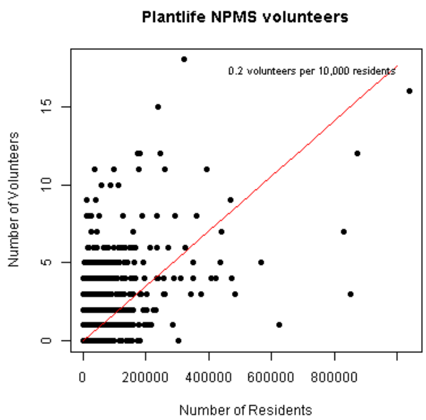

Plantlife NPMS volunteers
0.2 volunteers per 10,000 residents
Number of Volunteers
Number of Residents
0 200000 400000 600000 800000
0 5 10 15

Figure 4.2. Maps of relative supporting and volunteering rates. Blue areas are those where supporting/volunteering is low relative to the area's’ resident population size whereas red areas are those with relatively high rates of supporting/volunteering. White areas show supporting/volunteering at the national average or where there is no resident population.

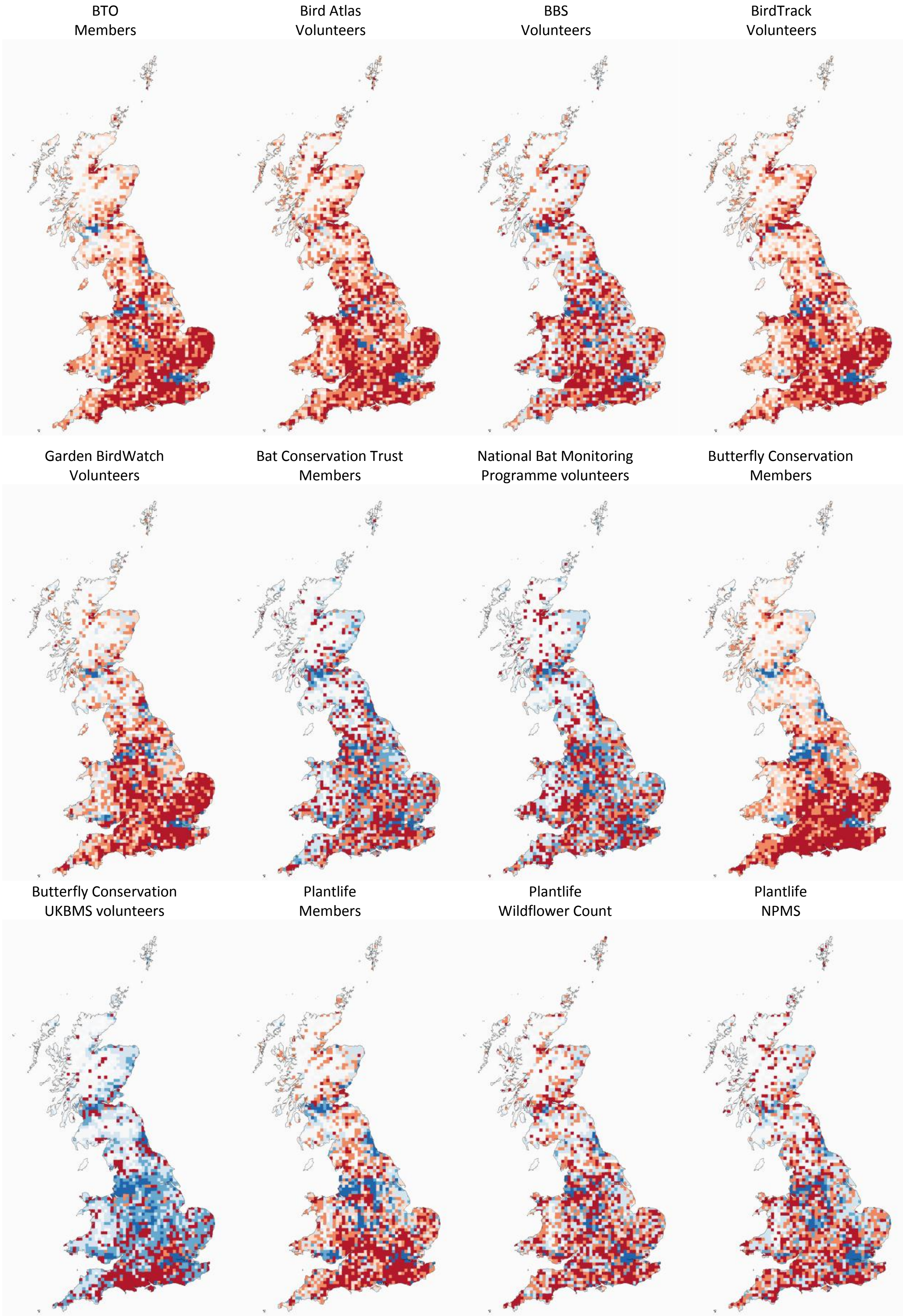


Figure 4.2 shows maps of the deviations in levels of membership or volunteering relative to the national average level per head of population. Red areas are those where membership or volunteering is higher than expected and blue areas show lower than expected. Virtually all organisations and schemes show lower than expected uptake in major urban centres. For butterflies and bats low uptake is also apparent in the Midlands and Yorkshire respectively. The maps are helpful in highlighting where the validation task could be most difficult. It is no surprise that the uplands present a major problem owing to low or no resident human presence. Perhaps most striking is the southerly bias to participation, particularly in botanical activities. Botanists are among the most qualified groups to undertake the validation task owing to their knowledge of indicator species of particular priority habitats, but these maps suggest their potential is limited in eastern and northern England and in Scotland. This reinforces the view stated in Section 3.7 that plant monitoring data will have limited applicability to the validation task.

4.3 Transferability

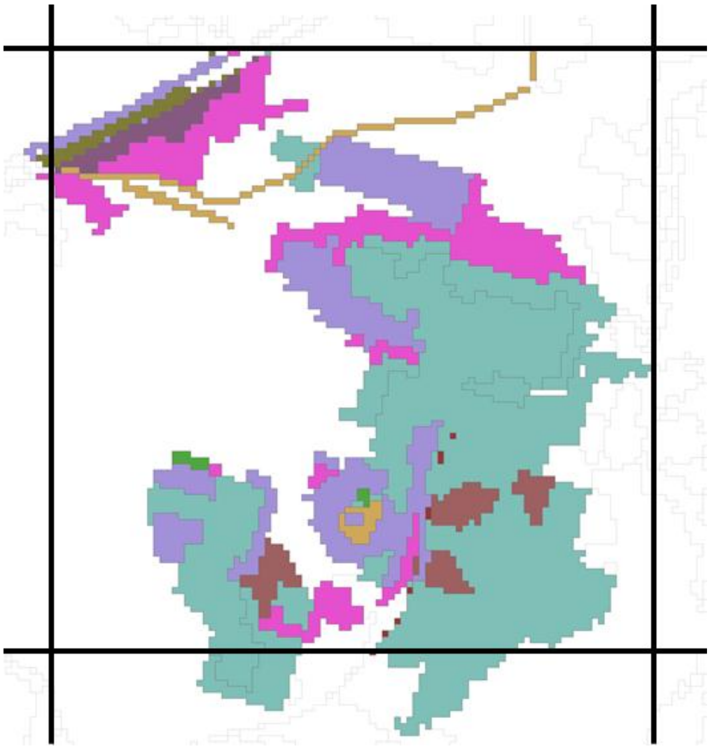
Interviewees were unsure of the transferability of the Norfolk workforce, or thought it was probably above average. Analyses of volunteering data show major regional variation for some taxonomic groups. If the project is successful in recruiting entirely volunteers from untapped populations (e.g. hikers, green gyms) it is impossible to predict the likely geographical variation in resource.

On a separate note, following a recommendation by Richard Alexander, Sam Neal (NBIS) interviewed Simon Pickles (North and East Yorkshire Ecological Data Centre) with respect to his negative experience of volunteer-based habitat validation. Simon summed up the situation from a recent project and which involved short timescales and difficult, time consuming and complicated survey methods. Simon's experience was in difficulties getting volunteers to identify habitats from scratch, rather than confirming/checking habitat type. His experience is that recording habitat is less attractive than other offerings to volunteer groups that are species based or habitat management orientated. However, he suggested that it may be possible to incorporate as part of other activities that volunteers are already motivated to do (this differs from the view expressed in Norfolk).

4.4 Conclusions and Recommendations

- The interviews and questionnaires identified 3,563 active volunteers, which generally comprises just under 10% of the total membership of the volunteer groups.
- Volunteer activity is likely to be biased towards areas where there is more people.
- Most groups thought that at least some of their volunteers were involved in structured recording, but currently very limited habitat recording.
- Respondents thought that their volunteers may be willing to make additional visits to known or to new areas, but that this would need to be limited so as not to distract or put volunteers off taking part in established surveys or activities.
- Asking volunteers to visit new areas was less appealing than going to sites known or local to the volunteer, but it would depend on the travel time required.
- Given sufficient training, it was felt that most volunteers could identify broad, common or easy to identify habitats, but training would be required to cover more than these.
- The response to desk-based validation was quite negative, but a number of groups were identified as a potential source for a small number of "high input" volunteers.
- A field-based method in the volunteers local area was thought to be most attractive.
- Many groups interviewed thought that their volunteers probably would not want to use a Smartphone app but others suggested that a well-designed app would be used.
- Smartphone allow greater control of data collection e.g. collection of GPS coordinates and controlled vocabulary lists, which reduce the risk of transcriptions errors.
- Analysis of existing schemes suggests structured scheme capacity of 0.5 volunteers per 10,000 residents, rising to 1–2 volunteers for unstructured schemes. Based on the current Norfolk population (878,000) we could expect 44 volunteers for a structured scheme or 88–176 for an unstructured scheme.

5. Evaluation of feasibility, practicality and efficacy of different methods and sampling approaches



Caption: Parcels in the TL6728NW 500-m square showing pixelated boundaries.

5.1 Do we need both desk-based and field-based methods?

It could be argued that there is no need for a desk-based component because the entire validation task can be done from the field, especially given more habitats are likely to be identifiable in the field than at the desk. However, visibility and accessibility of individual parcels may be more difficult in the field. Also, as shown in Chapter 2, when all habitat types are considered, the number of parcels per grid squares can be very high. If grid squares are made small enough to ensure parcel-rich squares can be covered, the opposite problem arises: some squares have too few parcels to warrant sending an observer to them. Fortunately, the parcel types that are so numerous in the parcel-rich squares are also ones that should be identifiable from the desk with relative ease. Hence we recommend a twin approach, involving desk-based validation of certain habitat types to facilitate a field-based approach of the remainder of the habitat types. We consider the design of each component separately in the following sections. In all cases we should strive for structured recording. Whether there is merit in an additional unstructured element is a secondary question.

5.2 Desk-based study design

Our assessment of the ease of identification of habitats from the desk considered two variants: a) identification by highly trained volunteers using a system that presents additional data layers and b) identification by a larger pool of crowd-sourced volunteers where the ability to achieve a high level of training is more difficult, and the provision of multiple data layers is likely to significantly reduce participation.

It is our view that the effectiveness and habitat coverage of the crowd-sourced option (Table 2.3) is so limited as to be not worth the investment. Furthermore, a large crowd-sourced solution might cause communication problems: a) the number of parcels and hence number of observers needed may be relatively modest and less than provided by the crowd and b) engaging with a large crowd could raise high expectations for participation in areas that do not yet have a Living Map and for detected errors to be corrected more rapidly than is likely to be the case.

On this basis, and following feedback from local groups on the availability of "high-input volunteers", we recommend a desk-based solution as follows:

- targeted at a relatively small number of high-input volunteers; enough to get the job done and manageable to train and invest in; enough to ensure some duplication of coverage of parcels to allow quality assurance assessment.
- they can be provided with detailed training material/one-to-one training
- they focus on a subset of habitats; in particular focusing on super-abundant habitat classes, plus any that are best identified using additional data layers
- sampling would be random, stratified by habitat class and classification zone and the unit of coverage would be the parcel (zero "travel time" for desk-based methods removes need for clustering parcels within squares).
- the technological solution would need to provide additional data layers to aid identification of particular habitats.
- whereas parcels would be selected at random, it may ease workflow for users to select a particular habitat type and work through its random parcels. This would save them repeatedly turning on/off data layers as each new parcel is presented. An argument against this is presenting randomly ordered habitats could reduce fatigue/complacency.
- The range of habitats to be covered by a desk-based task is likely to vary among regions of the UK as each may have a different suite of habitats that can be identified remotely. However, a subset of habitats will be common to all regions, including Gardens and Urban habitats which are the key superabundant habitats that need to be processed at the desk to ease the task in the field.

A key disadvantage of this approach is that it requires a bespoke technological solution that will be used by relatively few people at any one time. In addition, this approach is only likely to be cost effective if it can be certain that the validation task will be rolled out to other regions in due course.

In the following section, we look at the number of parcels that would be needed to be validated per habitat class to achieve adequate statistical power for quantifying error.

5.2.1 Simulations and power analysis

A power analysis was undertaken to determine the relationship between sample size (the number of parcels checked) and the ability to correctly quantify error. Simulations were performed in the R computing environment using bespoke code. For each sample size the number of errors (or events) was simulated using a binomial probability function (R binom) with a per-trial probability of 0.15. For example, if 75 parcels were checked, this R code will produce a random number which is drawn from the frequency distribution of events that can arise from 75 trials, where the probability of an event is 0.15. This is equivalent to tossing a loaded coin (probability of a head = 0.15) 75 times and counting the number of heads. Repeating this process yields a series of numbers (13, 13, 9, 5, 12...) which are independent estimates of the numbers of errors that could be detected, from which error rates can be calculated (17%, 17%, 12%, 7%, 16%...). From these values we can determine how often the true error rate of 15% is identified under different sampling conditions. Based on an earlier view from the Steering Group that we are looking for accuracy of e.g. 85 ± 5%, we graded each simulation run as having successfully identified the true error rate if it returned an error rate in the range 0.1–0.2 (i.e. 15% ± 5% points). We also evaluated precision of ±2% points and ±1% point. We then find the sample size at which 95% of error estimates fell within these precision bounds. Simulations were run 10,000 times.

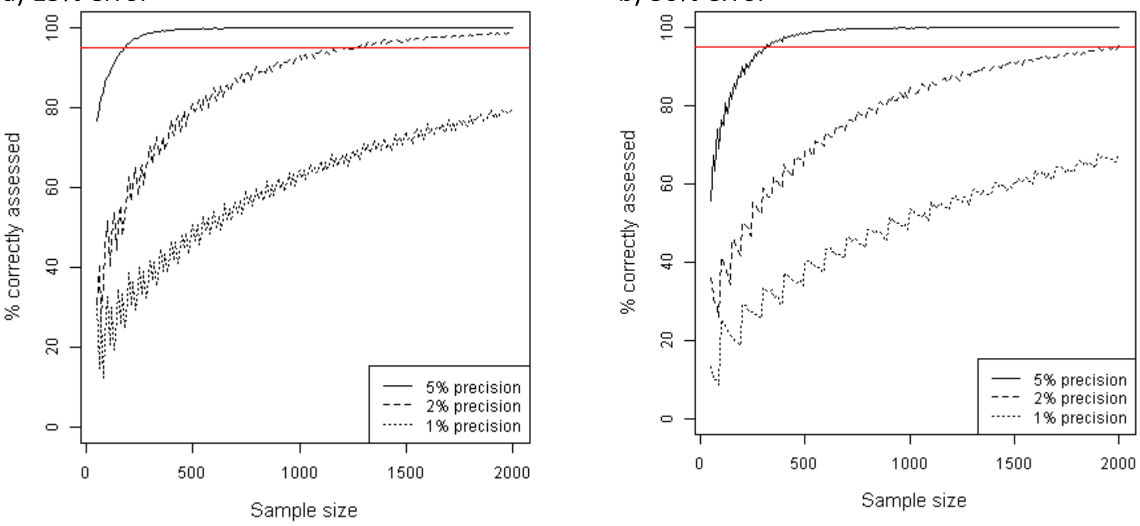
The results of this simulation are shown in Figure 5.1a and reveal that c180 parcels of a habitat type would need to be checked in order to have a 95% certainty of detecting 15% error rate with ±5% point precision. For precision of ±2% points minimum sample size rises to c1200 parcels and for ±1%

precision to c4800 parcels. Five rare Norfolk Living Map habitats have only c180 or fewer parcels in total: Improved grass with scrub (78 parcels), Humid dune slacks (152), Coastal Dune Heathland (170), Felled Woodland (181) and Orchard (183). For these habitats, sampling to yield 5% precision would result in complete coverage and an absolute error figure. If ±2% precision was required, three more habitats fall below the threshold number of parcels: Maritime Cliff and Slopes (308), Lowland Heathland (Scattered) (352) and Coastal Sand Dunes (scrub) (853).

Repeating the analysis using a higher error rate (30%) increased required sample sizes to c320, c2000 and c8000 (Fig 5.1b). For very high error (60%) sample sizes were increased to 360, 2300 and c9000. More generally, uncertainty peaks as error approaches 50% as shown in the contour plots (Fig 5.2).

Figure 5.1. Power analysis of parcel-level sampling with error rate of a) 15% and b) 30%. The three black lines show how the percentage of instances where simulations identified the correct error rate varied with sample size (number of parcels checked). Results are shown for three different precision levels. The red reference line is at 95% of cases.

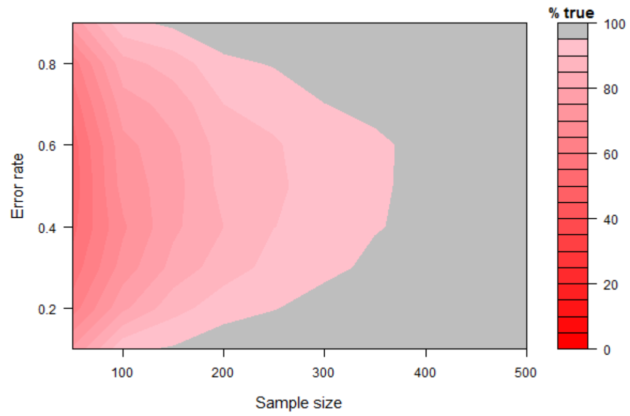
a) 15% error b) 30% error



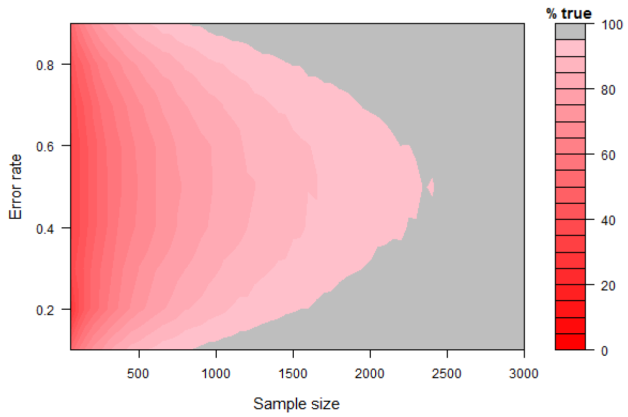
These simulations and graphs consider only the false positive error rate (commission error: a different habitat is classified as the focal habitat). A different approach is needed to ascertain false negatives error rates (omission errors; a focal habitat is mistakenly classified as a different habitat). For example, the Living Map documentation indicates that Orchards may be confused with Lowland Mixed Deciduous Woodland. Only 183 parcels of Orchards exist compared to 68,698 parcels of woodland. An error rate (false positive) of 15% for woodland would mean 10,305 parcels incorrectly classified as woodland. If 1% of these errors were in fact Orchards, the number of Orchard parcels would be increased to 286. If the number of false positives is calculated for each habitat type, and then the errors distributed among the remaining habitats in proportion to their initial abundance in the Norfolk Living Map, most habitats gain c6% more parcels. Only Urban, with 1,936,692 parcels initially would reduce in extent (by 5%). These figures are artificial because not all habitats are confused to the same degree, nor is it likely that the error rate will be constant across habitats.

Figure 5.2. Contour plots showing the relationship between sample size, underlying error rate, and the ability to correctly identify that error rate under different levels of precision: a) ±5% points, b) ±2% points or c) ±1% point. The coloured scale indicates the % of simulation runs providing the true error estimate within these bounds. Grey areas show combinations of sample size and error for which error was correctly estimated 95% of the time. Note differing x-axis scale between figures.

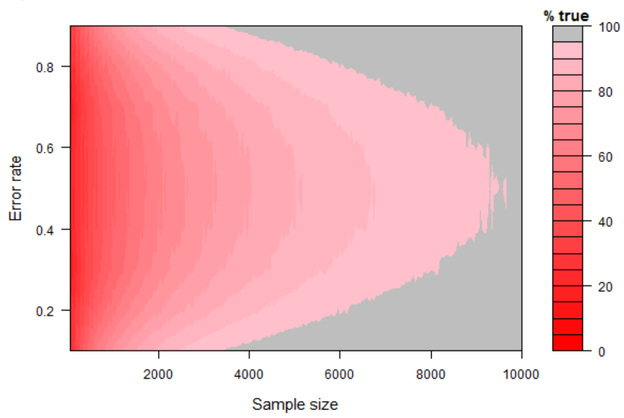
a)



b)



c)



5.2.2 Transferability

This method can be transferred to any region; in each case the habitat classes would need to be divided into:

A. those that are superabundant and need to be sampled at the desk to make the field validation task practical
B. those that are less abundant and are easily identified at the desk without additional data
C. those that can be identified at the desk only using locally or nationally available digital data sources
D. those that cannot be identified remotely and can only be identified in the field
E. those that cannot be identified at the desk or in the field by volunteers

The only transferability issue arises if there are habitat classes that would fall in A on abundance grounds but falls in D on desk-based identifiability grounds. Pasture fields in western Britain may be an example.

5.2.3 Recommendations

On the basis of ±5% point precision and anticipated error estimates of 15% we recommend a minimum sample size of 200 parcels of each habitat type. If there is a strong expectation that error rates could vary among classification zones, this sample size would be needed in each zone (subject to availability of parcels). If error rates are expected to be nearer 50% the sample size should be increased to 400 parcels per habitat (optionally also per zone).

5.3 Structured field-based study design

As discussed in Chapter 3, there are a number of advantages of structured over unstructured data collection, which mean that structured recording is most likely to result in representative habitat and geographic coverage. Following the recommendations in Chapter 3, there are good reasons to base field-based validation on grid squares, within which all parcels of most habitat types are validated. There is clearly a compromise to be made between maximising the number of habitat parcels that are validated and what is reasonable to ask a volunteer, with the further consideration of spatial non-independence potentially diluting the benefit of large squares containing multiple parcels per habitat class.

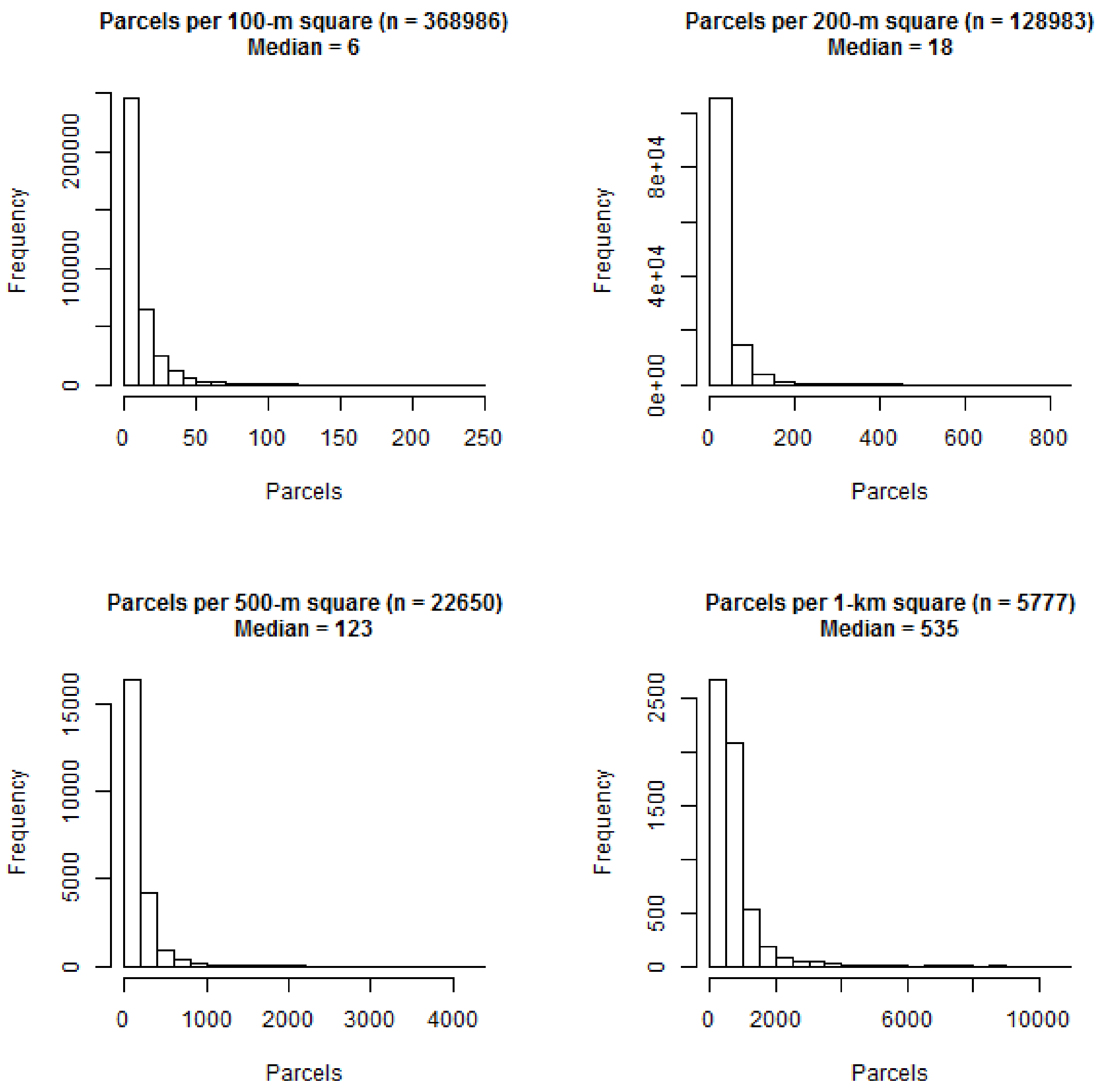
To consider what could be achieved with different size grid squares, initial GIS analysis of the Norfolk Living Map involved calculating the easting and northing coordinates of the centroid of each of the 4.5 million parcels. Using these coordinates and the geometry of the British National Grid, each parcel was assigned to grid squares of three sizes: 100-m, 200-m and 500-m. Note that 1-km squares had already been discounted owing to the excessive numbers of parcels for volunteers to cover, but are shown here for completeness.

The frequency distribution of the number of habitat parcels that a volunteer may be given to validate if assigned grid squares of 100-m, 200-m, 500-m and 1-km are shown in Figure 5.3. Using these distributions (and calculating the median and 90% percentiles), it is possible to determine the number of habitat parcels that 50% of surveyors, and at the upper end, 10% of surveyors will experience given a grid square of that size:

- For a 1-km square, 50% of surveyors will experience squares with 535 or more habitat parcels, whilst the upper 10% of surveyors will be assigned squares with at least 1,366 habitat parcels.
- For a 500-m square, 50% of surveyors will experience squares with 123 or more habitat parcels, whilst the upper 10% of surveyors will be assigned squares with at least 379 habitat parcels.
- For a 200-m square, 50% of surveyors will experience squares with 18 or more habitat parcels, whilst the upper 10% of surveyors will be assigned squares with at least 75 habitat parcels.

- For a 100-m square, 50% of surveyors will experience squares with 6 or more habitat parcels, whilst the upper 10% of surveyors will be assigned squares with at least 28 habitat parcels.

Figure 5.3 The frequency distribution of the number of habitat parcels within grid squares of four sizes: 100-m, 200-m, 500-m and 1-km.



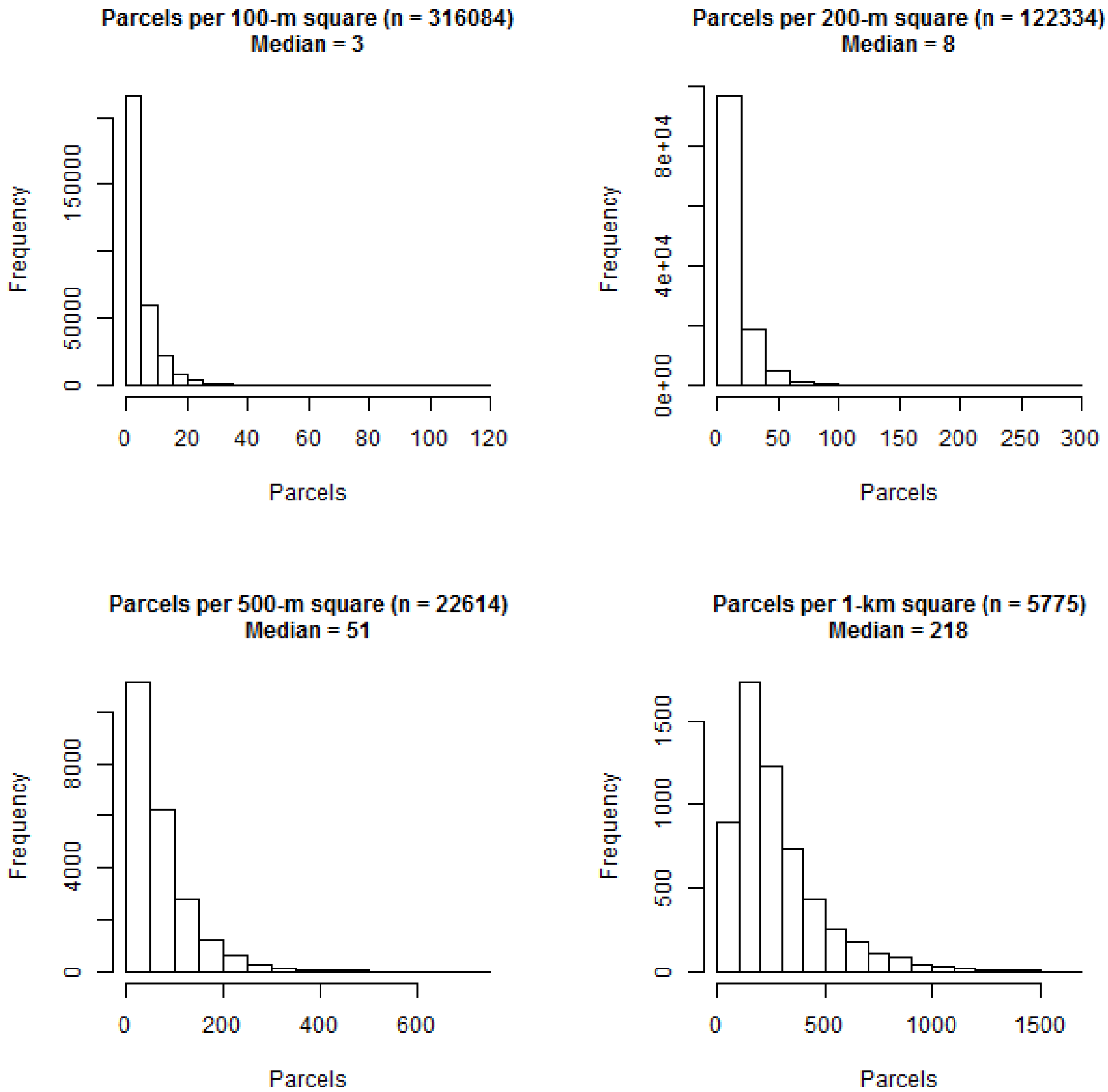
If we consider the frequency distribution of parcels of each habitat that would need to be surveyed in different size grid squares, it is clear that Gardens and Urban parcels increase considerably the number of parcels that would need to be validated. This is illustrated for a 500-m grid squares in Table 5.1. Gardens and Urban parcels can be validated through a desk-based approach, so one option might be for field validation either not to record these, or for squares containing a large proportion of Garden and Urban parcels to be removed from the field validation process. Figure 5.4 looks at the frequency distribution of the number of habitat parcels that a volunteer may be given to validate if assigned grid squares of 100-m, 200-m, 500-m and 1-km, but now having excluding Gardens and Urban parcels.

Table 5.1. Number of 500-m grid squares containing or not containing one or more parcel of each habitat and the frequency distribution of number of parcels of each habitat that are likely to be encountered in a 500-m squares.

Habitat	No. squares with habitat	No. squares without habitat	Frequency distribution of habitat parcels / 500-m square				
			10th percentile	Lower quartile	Median	Upper quartile	90th percentile
Arable	18,449	4,201	1	1	2	3	5
Bare Ground	345	22,305	1	1	2	5	13
Beach	210	22,440	1	1	3	10	18
Bracken	2,449	20,201	1	1	2	3	5
Coastal and Floodplain Grazing Marsh (Coastal)	300	22,350	1	3	8	16	35
Coastal and Floodplain Grazing Marsh (high productivity)	1,596	21,054	1	3	8	17	29
Coastal and Floodplain Grazing Marsh (low productivity)	1,493	21,157	1	2	7	17	32
Coastal and Floodplain Grazing Marsh (medium prod.)	1,815	20,835	1	3	14	32	51
Coastal Dune Heathland	28	22,622	1	2	3	10	17
Coastal Saltmarsh (established)	439	22,211	1	3	7	17	34
Coastal Saltmarsh (pioneer)	392	22,258	1	3	6	13	22
Coastal Sand Dunes	307	22,343	1	2	4	9	20
Coastal Sand Dunes (scrub)	125	22,525	1	1	3	10	19
Coastal Sediment	446	22,204	1	2	4	10	22
Coniferous Plantation	3,475	19,175	1	1	2	4	7
Dune Grassland	176	22,474	1	2	4	9	26

Felled Woodland	122	22,528	1	1	1	1	3
Fen, Marsh and Swamp	1,021	21,629	1	1	2	5	11
Gardens	16,606	6,044	2	4	12	32	116
Hedgerow or Field Margin	20,421	2,229	3	7	17	34	57
Humid dune slacks	70	22,580	1	1	1	2	5
Improved (scrub)	57	22,593	1	1	1	2	2
Improved Grassland	11,138	11,512	1	1	3	9	20
Lowland Heathland	1,381	21,269	1	1	2	4	7
Lowland Heathland (Scattered)	192	22,458	1	1	1	2	4
Lowland Mixed Deciduous Woodland	12,329	10,321	1	1	3	7	13
Maritime Cliff and Slopes	70	22,580	1	1	2	6	9
Orchard	116	22,534	1	1	1	2	3
Scrub	12,988	9,662	1	2	4	9	18
Semi-improved (scrub)	1,349	21,301	1	1	1	1	2
Semi-improved grassland	16,355	6,295	1	3	7	16	30
Semi-improved grassland (wet)	1,310	21,340	1	1	1	3	5
Urban	20,925	1,725	8	23	52	98	180
Waterbodies	15,176	7,474	1	2	9	37	77
Woodland Rides	2,018	20,632	1	2	6	17	32

Figure 5.4 The frequency distribution of the number of habitat parcels within grid squares of three size: 100-m, 200-m and 500-m, having excluded parcel classes as Gardens and Urban.



- For a 1-km square where Gardens and Urban parcels are not validated, 50% of surveyors will experience squares with 218 (down from 535) or more habitat parcels, whilst the upper 10% of surveyors will be assigned squares with at least 560 (down from 1,366) habitat parcels.
- For a 500-m square where Gardens and Urban parcels are not validated, 50% of surveyors will experience squares with 51 (down from 123) or more habitat parcels, whilst the upper 10% of surveyors will be assigned squares with at least 154 (down from 379) habitat parcels.
- For a 200-m square where Gardens and Urban parcels are not validated, 50% of surveyors will experience squares with 8 (down from 18) or more habitat parcels, whilst the upper 10% of surveyors will be assigned squares with at least 31 (down from 75) habitat parcels.
- For a 100-m square where Gardens and Urban parcels are not validated, 50% of surveyors will experience squares with 3 (down from 6) or more habitat parcels, whilst the upper 10% of surveyors will be assigned squares with at least 12 (down from 28) habitat parcels.

5.3.1 Simulations and power analysis

We ran a series of simulations to assess the level of precision delivered by different sampling strategies, grid sizes and sample sizes (numbers of squares). Simulation of errors at parcel level within squares followed a similar procedure to that used for parcel-based sampling (see Section 5.2.1). An error rate of 15% was assumed for all habitat types - i.e. we assumed that 15% of the parcels classified as Arable, for example, had been incorrectly classified. During the simulations, squares were identified at random, or according to certain rules depending on the habitats they contained (see below). As in Section 5.2.1, a habitat was said to have been adequately assessed under a particular sampling strategy if at least 95% of error estimates fell in acceptable bounds. Note that we do not consider spatial non-independence in these simulations. If this is judged to be a major concern then all recommendations on sample sizes should be considered minima.

Four sampling strategies were tested: A) Simple random sampling; B) Sampling focussed on rare habitats; C) Sampling focussed on all/most habitat and D) Random sampling stratified by classification zone. The form of each simulation and key results are detailed below.

5.3.1.1 Simple random sampling

Objective: to achieve coverage of all habitats via randomly selecting *N* grid squares.
Simulation structure: *N* grid squares are chosen at random from the list of squares present in Norfolk. Variants were grid resolution (500-m, 200-m, 100-m) and the number of squares selected (100–2000, increments of 100). Procedure repeated 1000 times.

Figure 5.5 shows the type of results generated for each habitat, indicating for a particular grid resolution how different sample sizes (numbers of squares surveyed) yielded varying numbers of parcels checked and varying estimates of error. Results for all habitats and grid resolutions are given in Appendix D.

FIgure 5.6 shows the summarised results of the number of habitats for which the correct error rate was detected 95% of the time with respect to varying sample size. With sample sizes ranging to 2000 squares, none of the grid sizes resulted in all habitats achieving the required level of error precision in 95% of simulations. Surveying 2000 500-m squares would yield robust error estimates 95% of the time for 23 habitats, but not for: Improved (scrub) (error correct 29% of time), Coastal Dune Heathland (33%), Humid dune slacks (40%), Orchard (42%), Felled Woodland (43%), Maritime Cliff and Slopes (53%), Lowland Heathland (Scattered) (57%), Coastal Sand Dunes (scrub) (77%), Beach (86%), Dune Grassland (90%), Bare Ground (91%), Semi-improved (scrub) (94%).

Doubling the absolute error rate (from 15% to 30%) slightly reduces the ability to confidently detect errors for a given sample sizes (Figure 5.7) and this phenomenon is likely to peak when errors approach 50% (see Section 5.2.1).

Figure 5.5. Example of power analysis results using simple random sampling of 500-m squares showing a) variation in the number of parcels of Dune grassland checked with squares sampled and b) variation in estimates of error rate; the solid red line shows the 15% error rate that was used in simulations and dashed red lines show ±5% points indicating the level of acceptable precision. Figures for all habitats are shown in Appendix D.

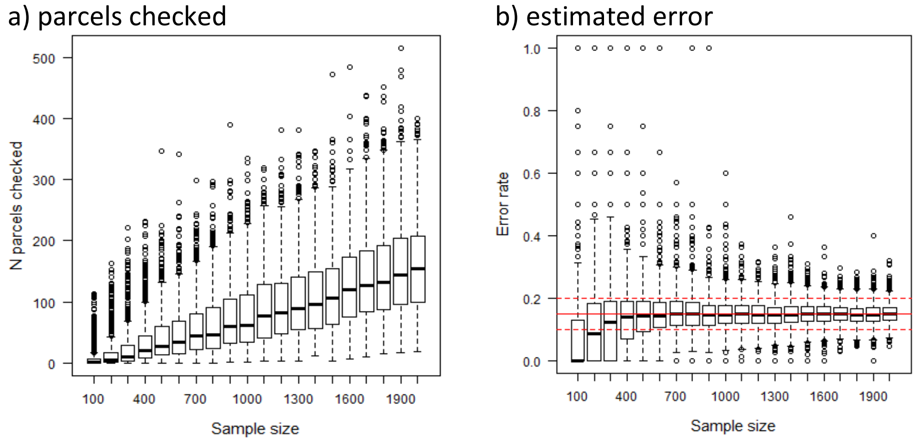


Figure 5.6. Power of simple random sampling to detect errors with differing levels of precision (a–c). Graphs show how the number of habitats for which error is correctly estimated varies with increasing sample size with separate results given for three grid sizes. Here correct assessment is taken to be when error is estimated within acceptable bounds 95% of the time (e.g. in a) 95% of simulations return error ±5% points of known error rate).

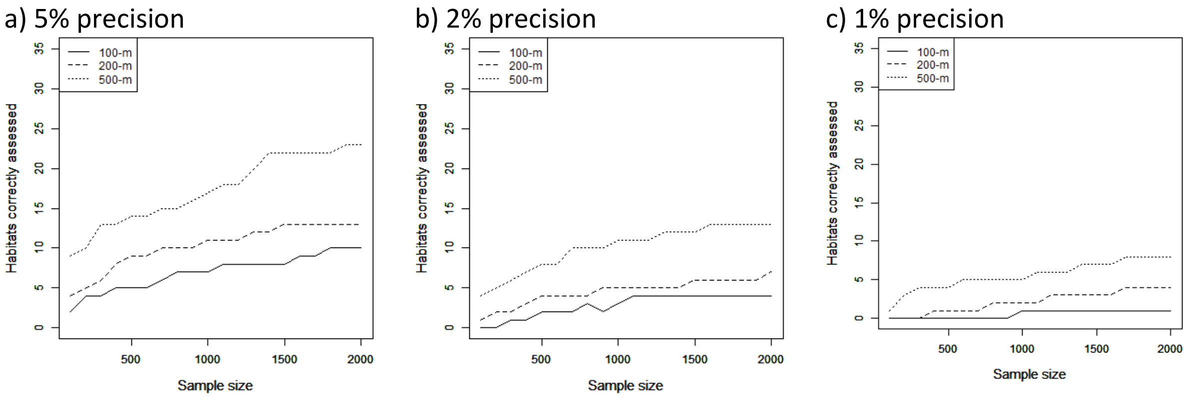
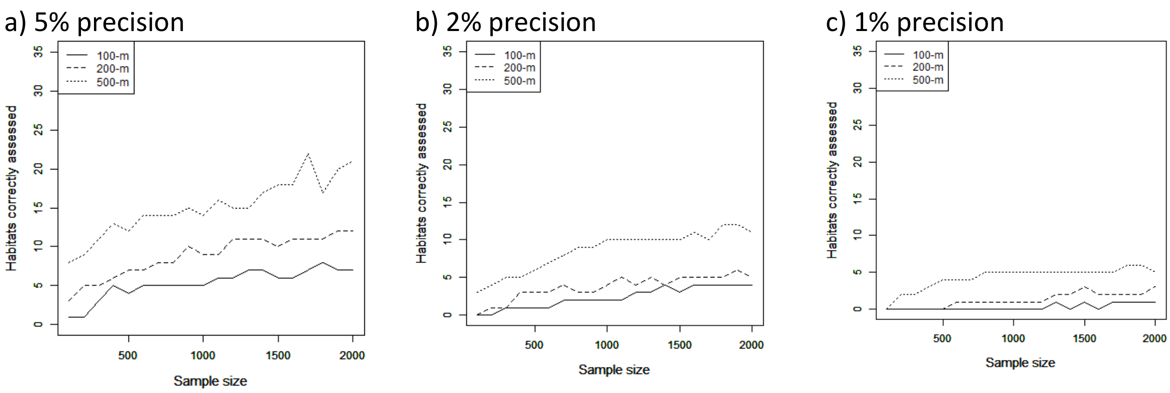


Figure 5.7. As Fig 5.6 except based on a 30% error rate.



5.3.1.2 Sampling focussed on rare habitats

Objective: to achieve coverage of all habitats by selecting grid squares that contain rare habitats in the expectation that common habitats will be covered as a by-product.

Simulation structure: Habitats are listed in order of rarity (rarest first); N squares are chosen at random from those that contain the rarest habitat, followed by N more squares that contain the 2nd rarest habitat, and so on for the first R rarest habitats. Variants were grid resolution: (500-m, 200-m, 100-m), the number of squares selected (50, 100, 200) and the number of rare habitats used for sampling (2 to 10, increments of 2). Procedure repeated 1000 times.

For the Norfolk Living Map data the 10 rare habitats were:

1. Improved (scrub) - 78 parcels (in 57 500-m squares)
2. Humid dune slacks - 152 parcels (in 70 500-m squares)
3. Coastal Dune Heathland - 170 parcels (in 28 500-m squares)
4. Felled Woodland - 181 parcels (in 122 500-m squares)
5. Orchard - 183 parcels (in 116 500-m squares)
6. Maritime Cliff and Slopes - 308 parcels (in 70 500-m squares)
7. Lowland Heathland (Scattered) - 352 parcels (in 192 500-m squares)
8. Coastal Sand Dunes (scrub) - 853 parcels (in 125 500-m squares)
9. Beach - 1598 parcels (in 210 squares
10. Dune Grassland - 1821 parcels (in 176 500-m squares)

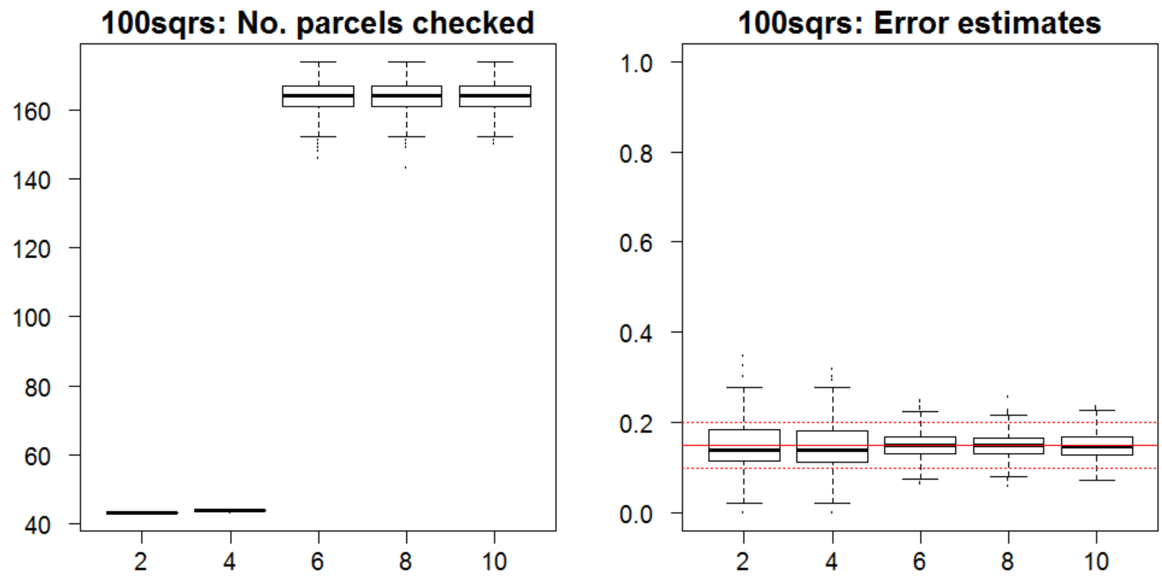
Results suggested using the rare-habitat focus for selecting 100-m and 200-m grids was ineffective because too few common habitat parcels coincide in these small squares meaning these habitats are not adequately assessed as a by-product. Hereafter we only consider rare-habitat focus using a 500-m grid. Figure 5.8 shows sample results for a rare habitat (Orchard) and a common habitat (Arable) (full results in Appendix D). Note at the 500-m scale only one rare Norfolk habitat is present in more than 200 squares, so results for 100 and 200 squares are largely identical. As the fifth rarest habitat, only a few Orchard parcels were sampled when 100 squares were selected for the first four rare habitats. When squares were selected for the first six rare habitats, c170 parcels were assessed (there are only 183 Orchard parcels in whole of Norfolk), giving error estimates approaching but not meeting the desired ±5% precision. For Arable habitat, as the number of rare habitats considered increased, so the number of Arable parcels captured in 100 squares with rare habitats increased to c800 and there was high precision around error estimates.

Overall, sampling based around 100 squares for each of ten rare habitats produced precise estimates for five habitats; extending sampling to 200 squares per habitat increased coverage to 6 habitats. These results are slightly misleading because with 100 squares, all parcels were captured for four habitats so error figures are absolute not estimates so precision is irrelevant (nine habitats for 200 squares). With 100 squares, only two rare habitats did not achieve a precise estimate or an absolute value. For the common habitats, sampling 100 squares for each of ten rare habitats produced precise estimates for 22 habitats, increasing to 23 for 200 squares (see Section 5.3.2).

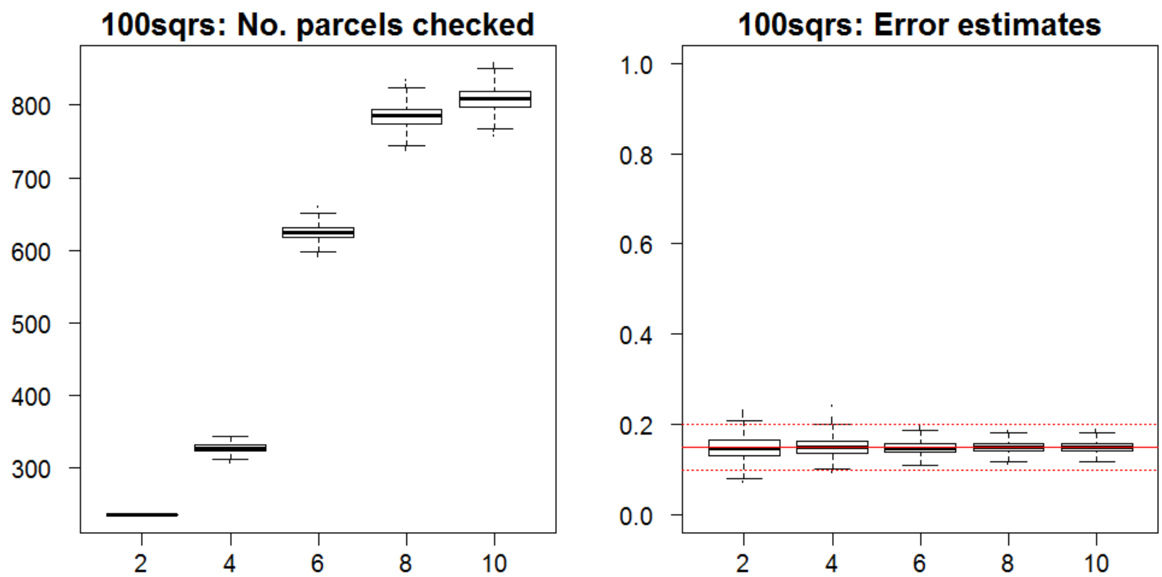
We conclude that selecting sample squares on the basis of rare habitat distribution shows some promise and can achieve adequate precision or absolute error values for 30 habitats if 100 squares are selected for each of the 10 rarest habitats. This requires a total sample size of c630 squares. A concern with this method is that common habitats in the vicinity of rare habitats could have different classification accuracy compared to the wider population, for example because of differing soil type in areas where rare habitats occur or due to pixel misclassification along shared borders.

Figure 5.8. Example results for the number of parcels checked and error estimates resulting from focussing sampling of 100 500-m squares per rare habitat, where the number of rare habitats was varied from two to ten. Graphs are for a) Orchards and b) Arable parcels (all habitats are shown in Appendix D).

a) Orchard



b) Arable



5.3.1.3 Sampling focussed on all/most habitat

Objective: to achieve coverage of all habitats by selecting *N* grid squares containing each type of habitat.

Simulation structure: This follows the same procedure as sampling focussed on rare habitats, but the number of "rare" habitats is set at 34 to include all habitats (here excluding Gardens on assumption they are validated using desk-based methods). Variants were grid resolution (500-m, 200-m, 100-m) and the number of squares selected (50, 100, 200). Procedure repeated 1000 times.

The results of this approach are summarised simply as the percentage of simulation runs on which the correct error estimate was recorded (within ±5% points). Note that for some of the rarer habitats, all parcels were captured by certain sampling designs so errors are absolute and precision can be ignored. Nineteen habitats could be adequately assessed (95% or more simulation runs gave precise answer) for all permutations of grid sizes and the three sample sizes selected (50, 100, 200; Table 5.2). A further seven habitats could be adequately assessed for all sampling strategies except 50 squares of each habitat at 100-m resolution. The best performing strategies were 200 squares of any grid size which returned precise or absolute error values for all habitats. Sampling 100 500-m squares per habitat yielded precise or absolute figures for all but two habitats (Table 5.2).

Owing to the aggregation of certain habitat types, randomly selecting 200 squares with habitat *X* and 200 with habitat *Y* will not necessarily generate a list of 400 squares. For the Norfolk habitats and their level of aggregation, a saving of between 2% and 37% can be made on the sample size depending on grid scale and target sample size per habitat (Table 5.3). For 500-m squares, co-occurrence of habitats in the same square means that the total sample size can be reduced by up to 37%. Nevertheless, sampling 200 squares per habitat still requires coverage of over 4000 squares which is very high for the potential volunteer capacity of Norfolk.

Based on a 500-m grid, and using a combination of 50 squares per habitat for 28 habitats plus 100 squares per habitat for two habitats (as defined in Table 5.2) achieves adequate sampling for 30 habitats. Of the remaining five, Coastal Dune Heathland (28 squares), Improved (scrub) (57 squares) and Humid dune slacks (70 squares) are rare so could be validated by complete coverage. The final two habitats are Orchards (116 squares) and Felled Woodland (122 squares), potentially also coverable completely. Such a strategy could require 1993 squares, but habitat co-occurrence means this level of sampling can be achieved with c1700 squares.

Table 5.2. Power of habitat-based sampling to detect error. Figures show the percentage of simulation runs where a known error rate (15%) was identified with adequate precision (±5% points). Separate results are given for three grid sizes (100-m, 200-m and 500-m) and for differing sample sizes (numbers of squares) per habitat. Habitats where error was adequately estimated for all grids and sample sizes are omitted (*). Cells marked in yellow are those where the method detected the correct error rate less than 95% of the time; those marked in orange are where all parcels were sampled so errors returned are absolute and precision can be ignored.

Habitat	100-m			200-m			500-m		
	50	10 0	200	50	100	200	50	100	200
Arable	85	96	99	99	100	100	100	100	100
Bare Ground	95	99	100	96	100	100	99	100	100
Beach	95	99	100	100	100	100	100	100	100
Bracken	88	97	100	99	100	100	100	100	100
Coastal Dune Heathland	83	93	95	92	94	94	95	94	94
Coniferous Plantation	93	98	100	100	100	100	100	100	100
Felled Woodland	73	87	94	77	90	93	82	93	95
Fen, Marsh and Swamp	86	98	100	95	100	100	100	100	100
Humid dune slacks	79	89	92	83	92	91	92	92	92
Improved (scrub)	69	80	79	73	80	78	78	81	77
Lowland Heathland	85	95	100	96	100	100	100	100	100

Lowland Heathland (Scattered)	73	90	97	79	93	98	89	97	99
Maritime Cliff and Slopes	86	95	99	90	98	98	98	99	99
Orchard	66	83	93	74	87	94	86	94	95
Semi-improved (scrub)	75	87	97	81	94	99	95	98	100

* The following habitats had adequate power for all scenarios: Coastal and Floodplain Grazing Marsh (Coastal), Coastal and Floodplain Grazing Marsh (high productivity), Coastal and Floodplain Grazing Marsh (low productivity), Coastal and Floodplain Grazing Marsh (medium productivity), Coastal Saltmarsh (established), Coastal Saltmarsh (pioneer), Coastal Sand Dunes, Coastal Sand Dunes (scrub), Coastal Sediment, Dune Grassland, Hedgerow or Field Margin, Improved Grassland, Lowland Mixed Deciduous Woodland, Scrub, Semi-improved (wet), Semi-improved grassland, Urban, Waterbodies, Woodland Rides

Table 5.3. Results showing how the total sample sizes across all habitats varies with sampling strategy. Expected sample size is the maximum sample expected with no overlap of habitats (e.g. 50 squares for each of 34 habitats = 1700 squares).

	100-m			200-m			500-m		
	50	100	200	50	100	200	50	100	200
Expected sample size	1700	3400	6800	1700	3400	6800	1700	3400	6800
Actual sample size (mean)	1673	3267	6082	1627	3076	5520	1454	2565	4301
% saving	2%	4%	11%	4%	10%	19%	14%	25%	37%

5.3.1.4 Random sampling stratified by classification zone

Objective: to achieve coverage of all habitats in each classification zone (defined in section 2.1) by random sampling within zones.

Simulation structure: The procedure essentially follows the same protocol as the simple random sampling procedure except that squares were selected from within each classification zone.

The number of squares available for selection varies among zones and according to grid size (Table 5.4) so for some sampling strategies the available squares will become exhausted. As the distribution of habitats differs among zones we might expect different sample sizes to have differing effects across zones. Table 5.5 presents the percentage of 500-m squares in each zone that had at least one parcel of a habitat type. It can be seen that in Zone A only 12 habitats were present in 1% or more squares. For this simulation we only consider effective sampling for habitats present in at least 1% of squares in each zone. All habitats are covered in at least one zone by this approach.

Table 5.4. The number of squares of different grid sizes per zone. Zone X represents the squares whose centre fell outside a classification zone; these are not used further in simulations but would need locating more accurately in a future validation exercise.

Grid resolution	Zone A	Zone B	Zone C	Zone D	Zone E	Zone X
100-m	10,155	37,502	100,861	212,270	7534	664
200-m	3214	12,756	32,999	76,935	2615	464
500-m	521	2157	5620	13,540	589	223

Table 5.5. Summary of the percentage of 500-m squares in each zone holding parcels of each habitat type.

Habitat	Zone A	Zone B	Zone C	Zone D	Zone E
Arable	91.7	76.1	83.4	84.9	9.3
Bare Ground	0	2.2	2.3	1.1	2.7
Beach	0	0.2	0.2	0	31.7
Bracken	0	0.2	0	18	0.7
Coastal and Floodplain Grazing Marsh (Coastal)	0	0.4	0	0.5	38.7
Coastal and Floodplain Grazing Marsh (high productivity)	0.2	2.9	18.7	3.2	7
Coastal and Floodplain Grazing Marsh (low productivity)	0	2.2	17.3	3.1	9
Coastal and Floodplain Grazing Marsh (medium productivity)	0.2	4.2	20.2	3.7	12.7
Coastal Dune Heathland	0	0	0.1	0	3.7
Coastal Saltmarsh (established)	0	0.6	0.1	0.5	58.9
Coastal Saltmarsh (pioneer)	0	0.6	0	0.3	57
Coastal Sand Dunes	0	0.1	0.2	0.2	45.5
Coastal Sand Dunes (scrub)	0	0.1	0.2	0.1	17.1
Coastal Sediment	0	0.6	0.1	0.3	64.5
Coniferous Plantation	0	13.7	12.7	17.3	19.2
Dune Grassland	0	0.1	0.3	0.1	25.3
Felled Woodland	0	1.5	0	0.7	0

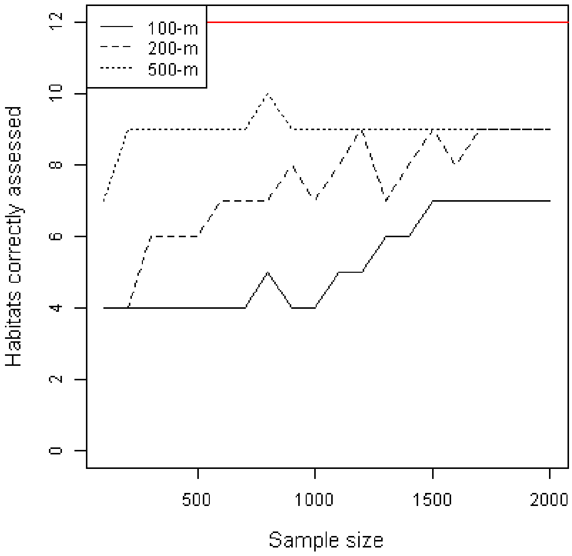
Fen, Marsh and Swamp	0.4	0.4	14.3	1.4	1.2
Gardens	68.7	50	76.5	78.7	23.1
Hedgerow or Field Margin	99.4	92.3	92.8	90.6	41.8
Humid dune slacks	0	0	0	0	11
Improved (scrub)	7.5	0.6	0	0	0
Improved Grassland	46.4	47.3	40.3	54.6	27.3
Lowland Heathland	0	3.6	0.1	9.6	0
Lowland Heathland (Scattered)	0	0	0	1.4	0.2
Lowland Mixed Deciduous Woodland	24.8	42.5	55.1	59.5	9.7
Maritime Cliff and Slopes	0	0.5	0.1	0	8.7
Orchard	15.7	1.3	0	0.1	0
Scrub	41.7	52.7	60.9	58.9	28.4
Semi-improved (scrub)	5	14.1	0.6	7.2	0.3
Semi-improved grassland	66.8	67.5	85.5	69.2	46.7
Semi-improved grassland (wet)	0	1.3	0.3	8.4	20.5
Urban	96.7	92.7	95.4	93.7	38.9
Waterbodies	94.8	76.3	73.8	62.7	48.2
Woodland Rides	0	8.5	2.8	12.3	0

Full results are presented in Appendix D. Here we summarise how many habitats are correctly assessed in each zone for differing grid sizes (focussing on precision of ±5% on a 15% error rate). Figure 5.4 shows that for none of the zones are all habitats in the zone effectively covered by sample sizes up to 2000 squares. As the number of habitats present in zones varies, the target for adequate coverage also varied (varying y-axis scale). For sample strategies of up to 2000 squares, only in Zone C with sampling of 500-m squares were all available habitats adequately validated. In each zone 35–75% of target habitats could be successfully validated with a random sample of 100 500-m squares per region but rarer habitats required very large samples (see Section 5.3.2, Table 5.5). Stratification by zones achieved validation for all habitats in at least one region except Coastal Dune Heathland, Felled Woodland, Improved (scrub) and Lowland Heathland (scattered), although several other habitat required impractical sample sizes in a zone (≥500 squares: Bare ground, Lowland Heathland, Orchard, Semi-improved (scrub)).

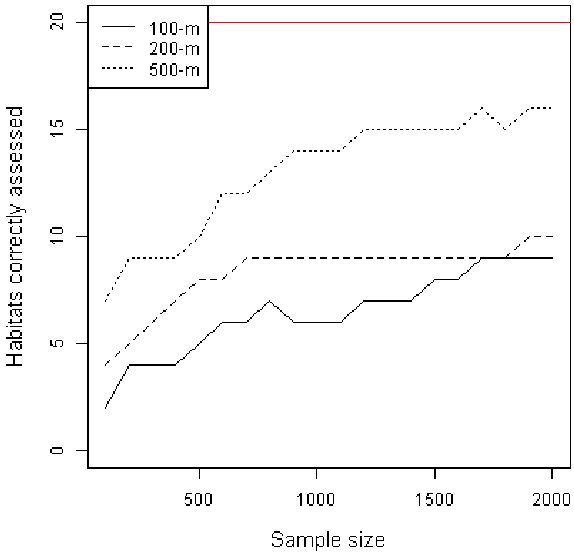
As in the simulation based on stratification by habitat, it may be possible to calculate optimal habitat-specific sample sizes per zone which ensure that each habitat is adequately assessed in the key zones in which it occurs.

Figure 5.4. Power of stratified random sampling to detect errors within classification zones (a–e). Graphs show how the number of habitats for which error is correctly estimated varies with increasing sample size with separate results given for three grid sizes. Here correct assessment is taken to be when error is estimated within acceptable bounds 95% of the time (e.g. 95% of simulations return error ±5% points of known error rate). Note that the number of habitats present in a region varies, hence the y-axis scale varies. The number of habitats present in 1% or more 500-m squares is shown for reference (solid horizontal red line).

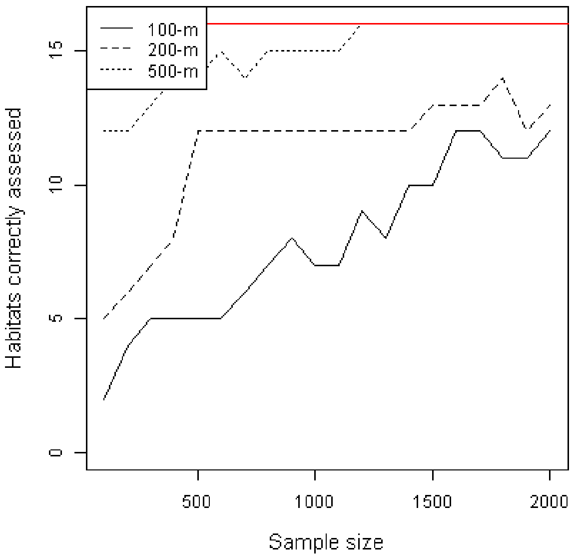
a) Zone A



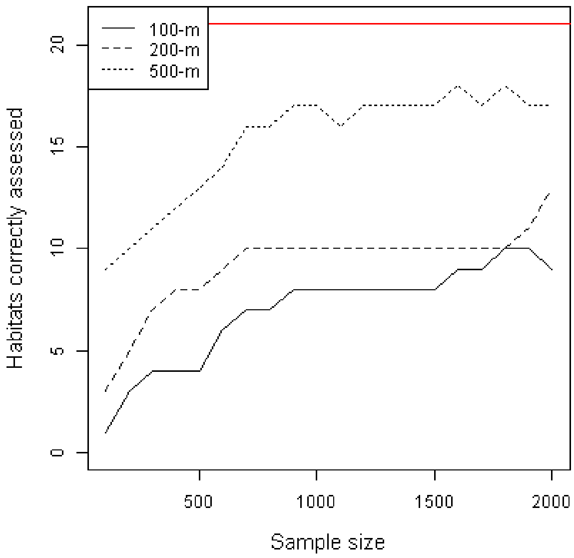
b) Zone B



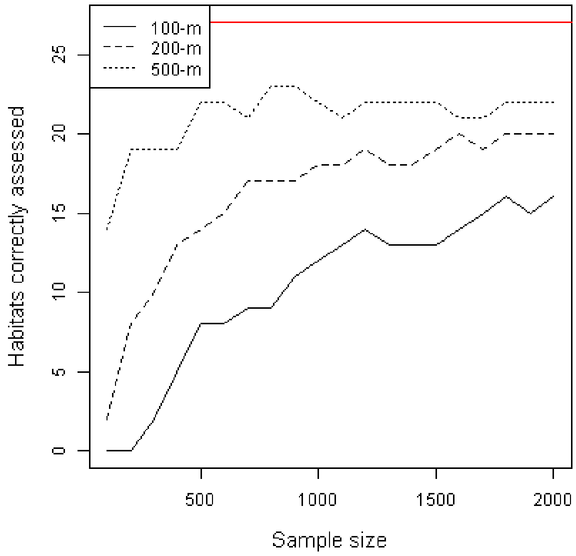
c) Zone C



d) Zone D



e) Zone E



5.3.2 Overview of field-based sampling simulations

Simulations have identified that 500-m squares offer the best balance of number of parcels per square and number of repeat assessments per habitat type. Using this square resolution, the key conclusions for each sampling strategy are presented for each habitat in Table 5.6 and can be summarised as:

- Simple random sampling of 500 squares can achieve effective validation for 14 habitats; increased sampling to 1000 squares gains coverage of two habitats and increasing to 2000 gains a further six habitats. Twelve habitats cannot be validated using simple random sampling with 2000 squares.
- Sampling 100 squares for each of the 10 rarest habitats achieves validation for 30 habitats with an overall sample size of c630 squares.
- Stratification by habitat can achieve effective validation for all habitats using 50–100 squares for common habitats, and all squares for the rarest five habitats and requires a total sample size of c1700 squares.
- Stratification by zone with a focus on all but the rarest habitats per zone achieves similar results to simple random sampling, with effective coverage of many habitats with 100–200 squares per zone but certain habitats are too localised and require substantially higher sampling.

5.3.3 Transferability

Conclusions about the size of grid squares and their impacts on the numbers of parcels observers would need to validate are likely to hold across the UK. Although the precise mix of habitats may vary between Norfolk and other regions, the scale of parcels may be quite similar. For example, the complexity of urban landscapes in parts of Norwich may map well to other cities. Similarly, the size of some fields in the Norfolk fens may be similar to the extent of individual habitat features in many other areas. Some wooded and upland habitats may be larger in extent but future Living Maps may break such large areas into finer habitat divisions. A difficulty may arise in western pastoral systems where improved pastures are very abundant, presenting a problem similar to the “urban” problem in Norfolk of inflated numbers of parcels for field validation. In such cases it will be necessary to shift the burden of validation to the desk-based component, provided such habitats can be identified remotely.

Simple random sampling did not work in Norfolk owing to the rarity and aggregated distribution of certain rare habitats. The nature of Living Maps, with a focus on priority habitats which are rare by their nature, means simple random sampling is unlikely to work in many regions.

Sampling based around rare habitats was effective for rare habitats and produced adequate coverage of common habitats as a by-product only when using 500-m squares. If priority habitats in other regions occupy unusual environmental conditions (e.g. specialist soil types) or limited geographical extents (e.g. arranged along linear features such as coasts), this sampling method is unlikely to be effective in other regions either. Before this approach is adopted in other regions it will require testing and tuning to determine the level of coincidence of rare and common habitats.

The transferability of sampling stratified by habitat type will depend on the number of habitat types in any new region and how often they co-occur in the same square. In Norfolk high co-occurrence significantly reduced total sample size but this level of co-occurrence and consequent saving on sample size cannot be guaranteed in other regions. If this approach is adopted in other regions it will require tuning to determine habitat-specific sample sizes.

The statistical aspect of the results concerning stratified sampling using classification zones are likely to be transferable to other regions. However, if the number of classification zones is large, the total sample size across zones could exceed the volunteer capacity of the region. How habitats are distributed across zones will also determine sample size in an unpredictable way. In Norfolk, several of the rare habitats were focussed in one or two zones, making it easier to target them with random sampling. If parcels of rare habitats are evenly distributed across zones it will be much harder to achieve effective coverage in any zone without stratification by habitat and zone.

Table 5.6. Habitat-specific overview of simulation results. The availability columns summarise the total number of parcels in Norfolk, the number of 500-m squares containing the habitat and the zones in which at least 1% of squares contain the habitat. Random sampling shows minimum sample size needed per habitat, or where this exceeds 2000 squares (grey shading) the level of confidence (% of simulations returning correct answer). Ten rare habitat sampling shows the level of confidence achieved when sampling 100 squares for each of 10 rare habitats; grey shading shows confidence <95%. Stratified by habitat shows the minimum per-habitat sample size needed to achieve precise estimates or absolute errors for all habitats. Stratified by zone shows the minimum sample size needed per habitat, or where this exceeds 2000 squares (grey shading) the level of confidence. No information is presented (black cells) for habitats occurring in less than 1% of squares in a zone. Names of the ten rare habitats are italicised.

Habitat	Availability			Random sampling: Min sample size / confidence	10 rare habitats: confidence with 100 squares per rare habitat	Stratified by habitat: Min sample size	Stratified by zone: Min sample size / confidence				
	n parcels	n squares	zones				A	B	C	D	E
Arable	49,527	18449	ABCDE	100	100%	50	100	200	100	100	>2000 (70%)
Bare Ground	2,724	345	BCDE	>2000 (91%)	86%	50		1900	1200	>2000 (79%)	1900
Beach	1,598	210	E	>2000 (86%)	100%	50					100
Bracken	6,267	2449	D	700	100%	50				400	
Coastal/Floodplain G.M. (Coastal)	4,113	300	E	1400	100%	50					100
Coastal/Floodplain G.M. (high p.)	20,367	1596	BCDE	300	97%	50		800	100	600	500
Coastal/Floodplain G.M. (low p.)	19,432	1493	BCDE	300	99%	50		1200	100	600	200
Coastal/Floodplain G.M. (med. p.)	38,443	1815	BCDE	200	100%	50		600	100	500	200
Coastal Dune Heathland	170	28	E	>2000 (33%)	Absolute	All					>2000 (92%)
Coastal Saltmarsh (established)	6,178	439	E	900	100%	50					100
Coastal Saltmarsh (pioneer)	4,249	392	E	1300	100%	50					100
Coastal Sand Dunes	3,200	307	E	1900	100%	50					100
Coastal Sand Dunes (scrub)	853	125	E	>2000 (77%)	100%	50					200
Coastal Sediment	3,812	446	E	1400	100%	50					100
Coniferous Plantation	11,298	3475	BCDE	500	100%	50		600	400	300	200
Dune Grassland	1,821	176	E	>2000 (90%)	100%	50					100
Felled Woodland	181	122	B	>2000 (43%)	92%	All		>2000 (65%)			
Fen, Marsh and Swamp	4,966	1021	CDE	1000	75%	50			300	>2000 (74%)	>2000 (62%)
Gardens	895,592	16606	ABCDE	100	100%	50	100	100	100	100	100

Hedgerow or Field Margin	536,318	20421	ABCDE	100	100%	50	100	100	100	100	100
Humid dune slacks	152	70	E	>2000 (40%)	Absolute	All					800
Improved (scrub)	78	57	A	>2000 (29%)	Absolute	All	>2000 (77%)				
Improved Grassland	83,270	11138	ABCDE	100	100%	50	200	100	100	100	100
Lowland Heathland	4,109	1381	BD	1100	100%	50		>2000 (92%)		700	
Lowland Heathland (Scattered)	352	192	D	>2000 (57%)	97%	100				>2000 (78%)	
Lowland Mixed Deciduous Wood	68,698	12329	ABCDE	100	100%	50	200	200	100	100	500
Maritime Cliff and Slopes	308	70	E	>2000 (53%)	Absolute	50					500
Orchard	183	116	AB	>2000 (42%)	93%	All	800	>2000 (65%)			
Scrub	100,245	12988	ABCDE	100	100%	50	100	100	100	100	100
Semi-improved (scrub)	1,917	1349	ABD	>2000 (94%)	71%	100	>2000 (58%)	900		1600	
Semi-improved grassland	204,109	16355	ABCDE	100	100%	50	100	100	100	100	100
Semi-improved grassland (wet)	3,467	1310	BDE	1300	100%	50		>2000 (81%)		900	
Urban	1,936,692	20925	ABCDE	100	100%	50	100	100	100	100	100
Waterbodies	398,440	15176	ABCDE	100	100%	50	100	100	100	100	
Woodland Rides	28,153	2018	BCD	300	100%	50		500	600	200	

5.3.4 Recommendations

Structured field-based recording is critical to the success of the validation task. For Norfolk an approach using 500-m grid squares excluding Urban and Garden classes appears to provide an acceptable number of parcels to validation. As a rule of thumb, coverage of 50 squares containing each habitat type achieves acceptable precision on error estimates for the majority of habitats because parcels are aggregated and more than 200 parcels are sampled per habitat. Precise countywide estimates for 30 habitats plus absolute error figures for seven rare habitats can be achieved with a sample of c1700 500-m squares. A smaller total sample size can be achieved if sampling is focussed on rare habitats but with potential bias in accuracy assessment for common habitats.

However, further input is required if Defra/JNCC require zone-specific validation. This is achievable statistically but at the cost of significantly increased overall sampling effort which may stretch and probably exceed available volunteer resources.

A hybrid approach, using focussed sampling of squares containing rare habitats and random sampling stratified by zone may deliver the best balance of ability to produce robust error estimates within minimal total sampling effort. In turn, this is critically dependent on which rare and superabundant habitats can be successfully validated from the desk. Agreement on the ease of identification of Norfolk habitats from the desk remains an issue. Such an approach could look like:

- desk-based assessment of Gardens and Urban (plus any diagnostic rare habitats)
- field-based sampling of 500-m squares selected to contain the c.10 rarest habitats
- field-based sampling of 500-m squares, stratified by zone to cover the remaining c.23 habitats, with 50–100 squares per habitat in each zone in which they occur (more than 1% of squares with habitat)

An optimal number of sample squares could be derived in this way, but with at least 10 habitats needing 50+ squares in each of five zones, plus samples for rare habitats, the total sample size is likely to be over 3000 which is prohibitive.

Further optimisation of the sampling strategy is not possible at the present time without greater clarity on the ease of identification of certain habitats. We recommend further consultation, including identifying which ancillary datasets should be used in desk-based validation followed up by pilot fieldwork to devise habitat identification criteria (for volunteer training). The routines used in the power analysis then provide a framework for deriving an optimal allocation of fieldwork effort. Depending on the objectives of the map and validation task, it could be useful to consider whether to vary the importance of validating particular habitats. For example, this could include reducing the level of validation of habitats for which there is a view that the classification is robust (e.g. due to classification algorithms using ground-truthed data) or habitats with low policy priority (e.g. arable). The danger with the former is that we do not have a robust measure of accuracy from previous work on which to base this decision, and the danger of doing the latter, is that of missing rare priority habitats that have been misclassified as common low-priority habitats.

5.4 Unstructured field-based study design

Interviews with the volunteer and local communities identified an interest in using the validation task to produce maps of local areas. In the sense that Defra/JNCC are keen for Living Maps to be used by communities this could be encouraged. In Section 3.3 we recommended the use of grid squares for the validation task where parcels are allocated to squares on the basis of each parcel's centroid. This ensures parcels are not cut by square boundaries and means that community groups can take on multiple adjacent squares to build a composite picture of their area.

A further benefit of unstructured field sampling is that it may be the most effective way of gaining any information on false negative errors. For example, to understand the error associated with mapping Orchards it is relatively straightforward to target Orchard parcels for survey. But detecting how many Woodland parcels are actually Orchards is very difficult, particularly as the number involved would likely be a very small proportion of the Woodland parcels (but still potentially a large number relative to the number of known Orchards). Community groups mapping their area have the potential to generate such false negative information. Whilst there are few mechanisms by which unstructured sampling will yield biased information (see Section 3.1) it may still be difficult to incorporate this information in a formal manner. It will be crucial that data collected through unstructured surveys are stored in a manner that they can be identified and separated from the formal (I.e. Random square) data. A potential communication issue concerns the likely desire for communities to enhance the map to provide the local contextual information they wish to represent; existing Living Map categories may be a poor fit to local (non-specialist) interests.

A second form of unstructured data collection that could benefit the false negative question is where observers submit information on individual parcels of interest. To continue the example above, this might be where an observer submits the location of an Orchard that they know is absent from the Living Map. This is only likely to be effective for rare or special habitats that are a) distinctive, b) of personal interest, c) sufficiently scarce (or missed) that missing parcels can be highlighted (if a habitat is too common, or rarely missed, observers are likely to give up if each parcel they aim to submit is already on the map).

5.4.1 Transferability

Transferability is heavily dependent on the presence of local communities with an interest and willingness to produce local maps. Whilst it can perhaps be assumed that such communities exist throughout the UK, the juxtaposition of priority habitats to willing local communities could vary, meaning that validation of key habitats by this method will be less effective in some areas than others.

5.4.2 Recommendations

Mapping by local communities should be encouraged. It should use the same size grid squares as the structured field-based recording (see Section 5.3). Data collected by this method should be stored in such way as it can be analysed separately from structured data to provide qualitative information on false negatives.

5.5 Conclusions and Recommendations

- We recommend desk-based assessment of 200–400 parcels of each habitat type, with a focus on superabundant easily identified habitats and any rare habitats that can be identified remotely with the use of additional data layers.
- The more habitats that can be processed at the desk, the easier it will be to produce an achievable sample size for field assessment.
- We recommend a parallel field-based assessment of a sample of 500-m grid squares.
- As a rule of thumb, 50 squares per habitat are needed to derive a robust error estimate; more if few parcels co-occur in a square.
- Countywide estimates could be produced with a habitat-stratified sample of c1700 500-m squares. This figure is likely to be at the upper end of local volunteer capacity.
- Further consideration is required on the need for separate accuracy metrics for each classification zone. However, zone-based estimates would require substantially higher effort which will exceed local volunteer capacity (except for habitat restricted to individual zones).
- Optimising the study design is not possible without further information on the ease of identification of different habitats which should be derived using a pilot survey.
- We recommend local communities be encouraged to undertake unstructured validation of a network of 500-m (i.e. self-selected) squares to produce local maps with the aim of providing qualitative information on false negative errors.

6 Review of technological requirements and solutions

Caption: Smart phone app presenting information relative to user's geographical location

6.1 Introduction

Validating the living map is not a trivial problem, and will involve the collation of validation information into a composite of digital results suitable for analysis and presentation. The level of inefficiency in duplication of recording (on paper) and potential for error in data entry indicates that the most likely route for validation data collection is using a direct digital medium.

There are a variety of existing tools and technologies that may be used to support the validation process but in order to evaluate these, questions at three different levels need to be addressed:

- Operationally - Is there an organisation that can implement, manage and maintain a system running the software?
- Functionally - Is there a software solution that provides sufficient functionality to meet the requirements of the validation task?
- Distribution - Can different functional elements live in different operational environments to provide a single solution?

When considering the functionality of a system, we can break this into a set of interrelated components. We can then evaluate each component offering from a vendor and look at how these may be combined in a variety of ways. Specifically, can a set of components be considered in terms of:

- An existing solution: Can a solution from an existing provider be found that meets all of the functional requirements and provides sufficient operational support to meet scalability requirements.
- Loosely coupled hybrid: Can we take existing elements from different providers and couple these together using process or bespoke developments. An example would be combining a bespoke location allocation application with (for example) Geo-Wiki to capture data. In a loosely coupled hybrid, using this example, there would be no check that an individual had validated a specific location in the location allocation system. This approach would require coordination with multiple vendors but not necessarily access to source code.
- Tightly coupled hybrid: Can we take an existing system and extend it with new functionality in order to provide a single system that meets the requirements of the survey. This approach would require access to source code from a single provider.

- Bespoke: Can a system be built (potentially using open source software) that meets the requirements of the project, hosted in an environment that can provide sufficient operational support to meet the needs of the project.

The advantages and disadvantages of these approaches are outlined in Table 6.1.

Table 6.1. Advantages and disadvantages of different types of solution.

Solution	Advantages	Disadvantages
Existing	● Proven technology implementation. ● Low implementation costs ● Minimal time between contract placement and system availability	● Will require a detailed evaluation of target systems to determine suitability for the methodology. ● Existing portal or branding may differentiate the survey giving lower visibility to potential users. ● Data may not be structured in a suitable manner for downstream use. ● Limits the ability to meet future system specific requirements.
Loosely coupled Hybrid	● Proven technology implementation. ● Lower implementation costs than bespoke. ● Potential to limit the time between contract placement and system availability.	● Will require a detailed evaluation of target systems to determine suitability for the methodology. ● The user is likely to require access to multiple disparate systems leading to limited coherence. ● Potential for user confusion (both in terms of process flow and branding) ● Limits the ability to meet future system specific requirements where commercially available existing solution elements are used.
Tightly coupled hybrid	● Potential for lower implementation costs than bespoke. ● Potential to limit the time between contract placement and system availability. ● Potential to provide a seamless user experience. ● Ability to tailor system to the specific survey requirements. ● Ability to lever existing market presence and user base.	● Will require a detailed evaluation of target systems to determine suitability for the methodology. ● Will require the ability to work collaboratively with existing vendors. ● Potential to require compromise in the detailed survey workflow. ● Potential for user confusion (both in terms of process flow and branding)
Bespoke	● Specifically designed to meet the needs of the survey	● Will require development of user community from the ground up ● Highest startup costs

	• Fully integrated workflow and user experience • Differentiation from existing systems • Potential for development of components to integrate into third party systems	• Longer lead time than other approaches

6.2 Review methods

In order to address the three operational, functional and distribution questions posed in Section 6.1 a set of organisations and potential vendors were identified based on our knowledge of the industry and with suggestions from the Steering Group (Table 6.2). Given the short timescales for the project, the vendor questionnaire was developed prior to the identification of the survey methodology. In order to address the range of methodologies that may have been proposed, the questionnaire was designed to cover as wide a range of scenarios as possible. Each sub-section of the questionnaire was then weighted based on its relevance to the proposed survey methodology once the data had been received from the vendors. This approach allowed the vendors time to answer the questionnaire in parallel with the survey methodology being developed. A two stage evaluation process was performed for each vendor:

1. Present vendors with the opportunity to provide input on the questions posed in Section 6.1
2. Review the information and products available

This two stage process is required as it is not possible to determine the operational capability of a system from an external review. For example, the scalability of a system is determined by aspects invisible to the external user such as numbers of servers, disc space and processing speed. Each functional element was then reviewed and scored based against the requirements in Section 2 of the vendor questionnaire (Appendix E - Vendor questionnaire). The raw scores for each vendor are then weighted against a set of requirements for a proposed solution (see section 6.3 Proposed solution) These overall scores are then reviewed and presented within the conclusion.

Table 6.2. Organisations and vendors approached with questionnaires.

System	Web link(s)	Contact
COBWEB	https://cobwebproject.eu/	Dr Crona Hodges, Dr Jamie Williams
Zooniverse	https://www.zooniverse.org/about/contact	Chris Lintott, Ali Swanson
Geo-Wiki / Laco-Wiki	http://www.geo-wiki.org/	Steffen Fritz
E-SMART	http://www.hutton.ac.uk/research/groups/information-and-computational-sciences/esmart	Matt Aitkenhead
Indicia	http://www.indicia.org.uk/downloads	David Roy; John van Breda
BTO	http://www.bto.org	Karen Wright

6.3 Vendor questionnaire responses and reviews

Of the six vendors approached, three provided a full response (BTO, COBWEB, E-SMART), one provided links to the source code (Indicia), and two provided no response (Zooniverse, Geo-Wiki). In order to provide an evaluation of Indicia, Zooniverse and Geo-Wiki an external review was performed on the publically available material for the functional elements of the system. In order to reduce bias in the scoring an external review was also carried out on E-SMART and COBWEB. Both of these reviews provided minimal meaningful data as both of these vendors provide a proprietary solution (i.e. a solution tied to a specific vendor with limited scope for external extension) and the instances of surveys available for public access were limited. One point to note is that all of the vendors who did not provide a response were educational or research organisations. The lack of response may be because of the short timescales provided (2 weeks) or the limited level of detail in the requirements. However one conclusion that could be inferred from this; despite there being a range of functional solutions, there are a much smaller number of organisations willing or capable of providing an operational environment for non-research projects. The following sections provide an overview of each vendor evaluation.

6.3.5 BTO

The BTO provides proprietary products for capturing and reporting taxa observations using a range of survey methodologies. A number of the browser-based products support habitat capture (Wild Surveys, WeBS, BBS) at sites or arbitrary locations identified using a map interface. Individual surveys and schemes include management modules designed for a hierarchical structure with nominated organisers allocating sample squares or sites to individual users. Some products offer the ability to collect data at user-specified locations (BirdTrack) and a number offer reporting and visualisation facilities (WeBS reporting, Garden Bird watch). There are two Mobile applications available on both IOS and Android platforms. One application (BirdTrack) provides the ability to record a variety of taxa data for specific locations uploaded to the device. It also provides the ability to create new locations (sites) and persist this data for viewing on the browser-based application.

Conclusion: The BTO systems could provide components for desk-based user management and desk-based and mobile recording of data. They are proprietary but based on existing open source products. Further investigation would be required to determine an appropriate architectural solution that could incorporate the user management elements.

6.3.1 COBWEB

COBWEB is an EU funded project (http://citizen-obs.eu/) for collecting crowd-sourced environmental data. It focusses on UNESCO BioSphere reserves and ran seven evaluation sub-projects in the Dfyi Biosphere reserve in 2015. The project focus combines citizen science observations with sensor-based data in a common repository. The system includes data cleaning and validation tools (Leibovici[4]) and these are based on open standards, but the vendor offering is essentially proprietary. The software provided is closed source with the Android application freely available for download using the Dyfi Biosphere portal[5] but it is worth noting that at the time of review (March 2016) the portal identifies all of the software as Beta.

Conclusion: COBWEB has the potential to provide a significant proportion of a solution, but is still a relatively immature offering. It is unclear how this may be integrated with elements from other vendors.

[4] Leibovici, Didier. Server Based Data Validation System Design. 1st ed. 2014. Web. 9 Mar. 2016

[5] "COBWEB Dyfi Biosphere Reserve Portal (Beta) - Cobweb". Dyfi.cobwebproject.eu. N.p., 2016. Web. 9 Mar. 2016.
https://dyfi.cobwebproject.eu/geonetwork/portal/eng/catalog.search#/metadata/b42c1ef8-1aed-4deb-92e7-7ffa0a101e33
https://cobwebproject.eu/sites/default/files/COBWEB%20D4.2%20Server-based%20Data%20Validation%20System%20Design_0.pdf

6.3.2 Zooniverse

The Zooniverse solution is open source and the source code is available at https://github.com/zooniverse. The architecture is designed for virtualized deployment using a Docker (Docker[6]) and has a centralised web aware data storage engine http://docs.panoptes.apiary.io. The front end is designed to present users with a series of images (subjects) and to allow the user to classify elements by either answering questions or drawing on the image. The management tools are designed to allow users to set up projects for other volunteers to assess and classify subjects. There is some scope to allow the administrators to present groups of subjects and to compare user results with expert results. The hosted Zooniverse solution is limited to uploading 10,000 subjects of 600Kb each.

Conclusion: The Zooniverse technology provides a suitable component for storage of core classification data. This may be suitable for development of a tightly coupled solution.

6.3.3 Geo-Wiki / Laco-Wiki

The Geo-Wiki platform provides support for a range of different sub-projects, Laco Wiki being the closest match to the current project. The architecture overview (Geo-Wiki: Data Design And Architecture - Technical Background - IIASA[2]) outlines the use of standard open-source components (PostGres and WMS service e.g. Geoserver) to present the projects and their functionality. It is not clear that Geo-Wiki provides an open-source or open-platform approach. Some element of support is available for polygon processing but the majority of the projects focus on raster datasets. Standard classifications are available e.g. CORINE, LISA, Modis but it is not clear that additional classifications are supported. There is no support for a stratified sampling strategy within the product at present although this is outlined for future release. There is support for mobile platforms although this is limited to taking, classifying and uploading geolocated photographs. There is no clear statement of ability to scale or number of users supported.

Conclusion: Geo-Wiki has the potential to support some of the features of the project. It does however appear to be a proprietary platform based on existing open-source products, so using it as part of a hybrid solution may be difficult. One point to note is that the mobile platforms language is currently German and may require English translation.

6.3.4 E-SMART

The E-SMART project provides infrastructure and web and mobile applications for the Scottish environment. Unlike other reviews, limited information is available on the World Wide Web. The initial evaluation was performed using information downloaded from the James Hutton Institute (ESMART: Web Tools And Apps For The Scottish Environment | The James Hutton Institute[7]). The infrastructure supports two apps:

- SIFSS (Soil Indicators for Scottish Soils) is an app that allows the user to find out what soil type is in their area, to explore the characteristics of around 600 different Scottish soils, to discover the differences in soil characteristics between cultivated and uncultivated soils and to examine a range of key indicators of soil quality.
- iDee is an app and website where the user can submit their own records of river conditions in the Dee catchment in the form of a photograph of the river together with simple assessments of water clarity, flow speed and algae cover.

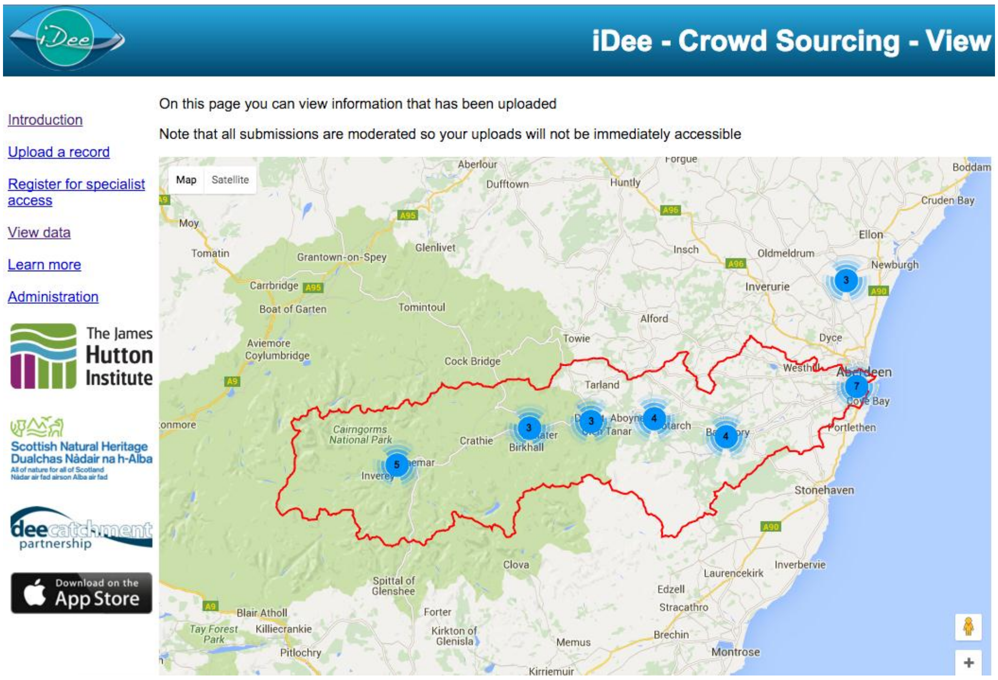
The results page of the iDee project (Figure 6.1) reflects the difficulty in moving a research idea from technology to production given that fewer than 30 records appear to have been entered on the map since 2012.

---

[6] "Docker". Docker. N.p., 2016. Web. 9 Mar. 2016. https://www.docker.com/

[7] "ESMART: Web Tools And Apps For The Scottish Environment | The James Hutton Institute". Hutton.ac.uk. N.p., 2016. Web. 9 Mar. 2016. http://www.hutton.ac.uk/research/groups/information-and-computational-sciences/esmart

Figure 6.1. Results for E-SMART



Conclusions: The E-SMART project appears to be an academic project that has limited potential for reuse within this project given the proprietary nature of the platform and the limited information available.

E-SMART project did provide a completed questionnaire following the external review and the scores from this have been used within the analysis (see Section 6.6).

6.3.5 Indicia

Indicia (http://www.indicia.org.uk/) provides an open source toolkit for wildlife recording that has been used to support a wide range of projects e.g. INNS Mapper (http://ywt-data.org/inns-mapper/), UK Ladybird Survey (http://www.ladybird-survey.org) and Irecord (http://www.brc.ac.uk/irecord/)

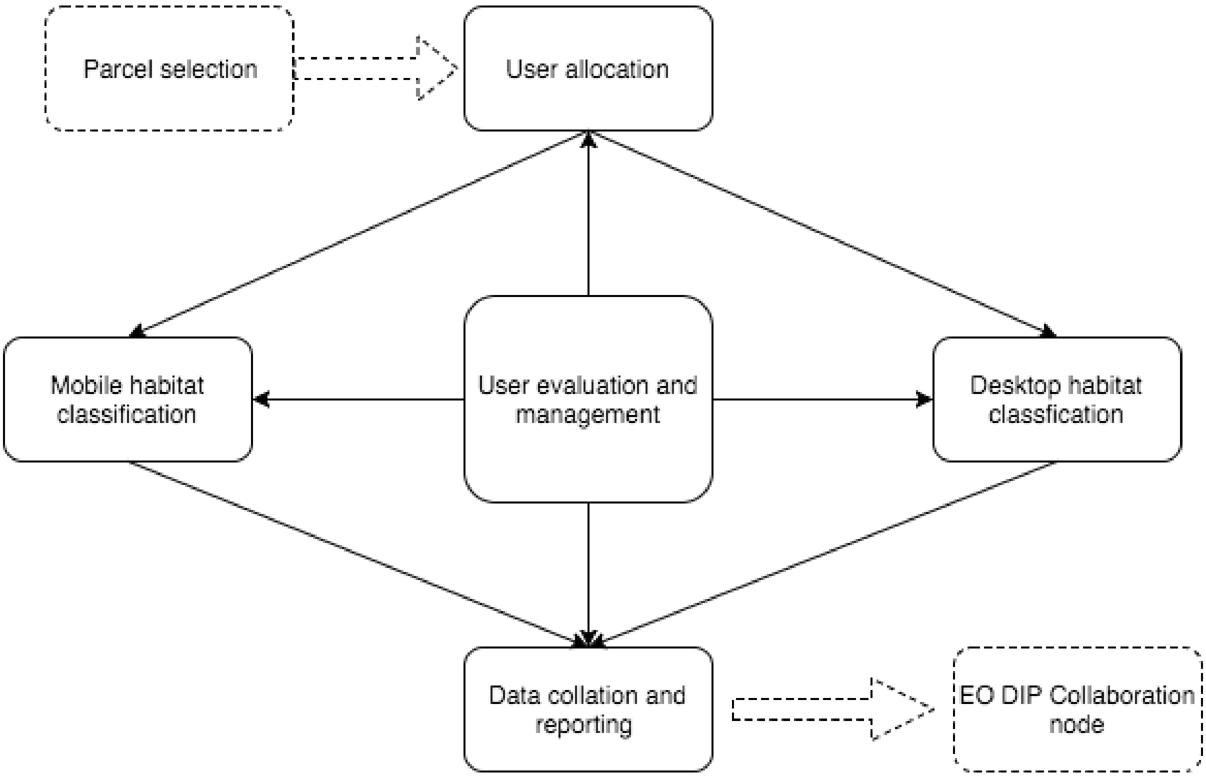
The core of the system is a Drupal-enabled website with a variety of open-source supporting products. This is freely downloadable from their site. The focus of the product is species observation recording and at present provides limited utility for this project. There is no support for a stratified sampling approach, or the ability to select polygons to focus data collection. Users can create their own 'sites' but this is a different paradigm to the approach used in habitat classification. The product could be tailored to support web services allowing the creation of mobile applications, but there is no direct evidence for this within the documentation available.

Conclusion: Indicia appears to be a maturing product for species observation recording and the repository element could be used as the starting point for this project. It does have potential to form a small element within a hybrid project but would require substantial extension to support a mobile architecture.

6.4 Proposed solution

Any software based solution can be broken down into a set of functional elements and the data flows between these elements. Using this divide and conquer approach could allow different vendors to provide individual functional elements that interoperate in order to solve the whole problem. This set of functional elements and their data flows can be considered a high-level architecture for the system. The architecture for the validation task solution will require a number of elements as identified in Figure 6.2 and explained in detail below.

Figure 6.2. Architecture for a proposed technological solution.



6.4.1 Component descriptions

Parcel selection: This module is required to identify the set of parcels for analysis based on some set of selection criteria. The output of this module may be manual or automated, but parcel selection for a given area may be performed multiple times so scalability of the solution should be considered. The output is a list of parcels for allocation and subsequent classification. Parcels may have been selected independently or as clusters (e.g. those parcels falling within each of a set of squares). The output could be as simple as a comma-separated list or be provided as a web service. It may be considered outside of the scope of an online system.

User allocation: This module takes an incoming list of parcels for classification and identifies suitable user / platforms for classification. The selection criteria for this may be automatically generated or may require human intervention. The output from this must be accessible from the mobile and desktop elements in an online manner and is likely to be provided as a web service.

User evaluation and management: User evaluation allows the system to present known parcels to the user and record a classification for comparison against a benchmark result. This may be used for training or to determine the likely error rate for a given user when classifying different types of habitat. The service must be capable of storing the results in the data collation component, and given the need to interact with both the desktop and mobile elements it should present the function as a set of web services. User management allows individual user registration, and storage of preferences. The user management does not necessarily require authentication via a localised registration system, given the ability to leverage federated authentication using OpenId, SAML or OAuth (depending on the specific project requirements). It would also require the ability to store and provide user preferences for user allocation (e.g. maximum travel distance) for both desktop and mobile habitat classification components. Given these requirements this service must be provided in an online manner, most likely as a set of web services.

Desktop habitat classification: This module is intended for use by a small number (up to 100) of skilled users. It should allow the presentation of different geospatial overlays in order to classify randomly assigned parcels. The purpose of this analysis is to process super abundant and/or difficult-to-access parcels, thereby removing them from the pool of habitats and parcels to be surveyed in the field. These would typically be urban parcels such as gardens that may be identified from composite map layers. It may also be used to process rare habitats that are identifiable from other data layers. The storage element of this solution requires integration with the user allocation and user management systems, and must be capable of producing output for the data collation and reporting system.

Mobile habitat classification: This module is intended for citizen scientists (potentially 1000+ users) to enter data for selected parcels and would consider structured and unstructured approaches. The parcel allocation is likely to be gathered into 200-m or 500-m grids in order to provide sufficient 'value' in visiting a location. Parcels could be selected from a list of targets based on a criteria entered by the user. The system should allow the users to classify the parcels offline and upload the results to the data collation and reporting system at a suitable point. This component must be capable of communication with the user allocation and user management systems. Given this requirement each of these components must be capable of supporting web services.

Data collation and reporting: This forms the repository for the data collected for each parcel by each user. This must be capable of outputting the data in a suitable format for further use. Optionally this component may be used to provide reporting statistics e.g. individual user parcels classified, total % parcels classified or % of given habitat classified, which could be made available to the survey organiser and as feedback to the volunteer. This module may be required to support some level of web services depending on how the elements of the system are distributed.

6.5 Requirements weightings

The component descriptions allow grouping of the requirements from the vendor questionnaire. From this we can determine the relative weighting for a given component for a particular scenario. The weighting scale is:

- 0.0 - Not required for this scenario
- 0.25 - Nice to have
- 0.50 - Could be included if time and budget allow
- 0.75 - Should be included if time and budget allow
- 1.0 - Mandatory

Table 6.3. Weightings for different requirements according to the recommended validation solution (see Chapter 5).

Requirement	Component	Weighting	Comment
Data storage	Data collection and reporting	1.0	
User management (Must)	User management	0.75	This may need further refinement as the feedback element is optional
User management (Should)	User management	0.5	Use of federated login may improve uptake for mobile users

Data entry (Must)	Desktop habitat classification Mobile habitat classification	1.0	
Data entry (Should)	Desktop habitat classification Mobile habitat classification	0.25	
Browser based (Must)	Parcel selection User allocation	0.75	Not all of the requirements listed are mandatory
Browser based (Should)	Parcel selection Desktop habitat classification	0.25	
Mobile based (Must)	Mobile habitat classification	1.0	
Mobile based (Should)	Mobile habitat classification	0.75	The requirement to download a list of target polygons may be mandatory
Feedback (must)	Data collection and reporting	0.25	
Feedback (should)	User allocation	0.5	An automated allocation would need information on parcels not classified but allocated
Verification	Desktop analysis Mobile analysis Data collection and reporting	0.5	
Data export	Data collection and reporting	0.75	Filtering of results would be optional

6.6 Analysis

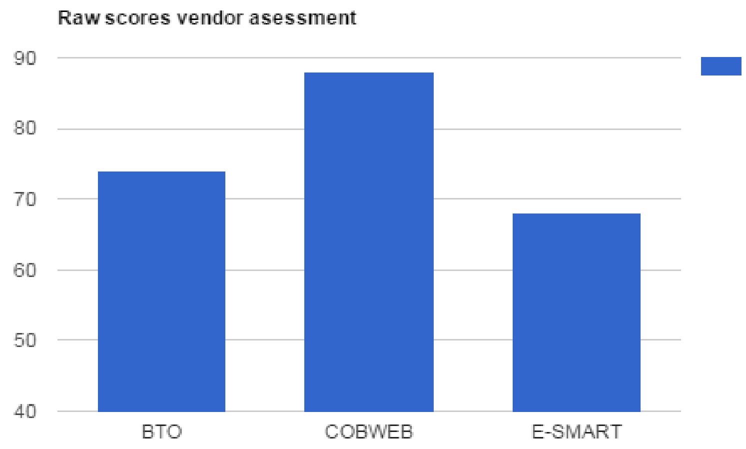
The solutions matrix (See Appendix F Solutions matrix) provides a summary of the vendor scores for the functional elements of the vendor questionnaire. It is useful to note that within the matrix, the three vendors that replied to the questionnaire scored themselves significantly higher than the three that did not reply when assessing the functional fit (Figure 6.2). A number of factors may have influenced this response profile:

1. Vendors who did not reply felt that the product or service they offered was not well matched with the set of requirements presented
2. There is insufficient public evidence to perform an external assessment of a vendor's technology.
3. Vendors that self-assessed provided a 'sales' version of results and that this causes some bias.

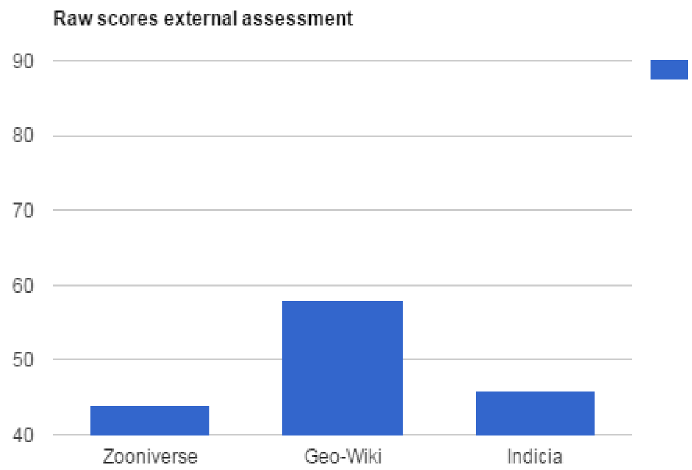
It is likely that some or all of these factors cause the difference in these raw scores.

Figure 6.2. Raw total scores for a) the three vendors who self-assessed and for b) the three vendors that had to be assessed by us

a) vendor assessment



b) external assessment



6.6.1 Non functional requirements

From the three completed responses the following conclusions may be drawn:

- **Licensing:** Ownership of the data being entered in the system varies by vendor and should be considered for any proposed solution
- **Integration:** Two of the three systems provide no opportunity for external integration, the third identifies Open standards for sensor integration, an area not specifically relevant for this survey.
- **Open source:** Of the 6 solutions, only two of the vendors identify the product as open source (Zooniverse and Indicia). In practice, only one of these (Indicia) makes the source freely available as a download.

It is likely that all of the vendors make some use of existing open source products in order to produce a solution. Based on the review these are typically:

- PostgreSQL database server using POSTGIS for spatial data storage
- Apache web server with appropriate modules e.g. mod_php, mod_ssl for web access
- Openlayers for web map rendering
- Geoserver for Web Map Server hosting
- Apache Cordova for mobile development

Scalability and data integrity: Of the three full responses one (COBWEB) identified scalability and data integrity as an issue due to the lack of maturity of the product. The other two responses provided confidence in scalability and data integrity.

6.6.2 Functional requirements

Using the proposed solutions functional components each vendor response was grouped and weighted.

User allocation: This area broadly matches the questions posed in Data entry (must) and Browser Based (Must) in the questionnaire. The combined scores for each vendor in each area is shown in Figure 6.3a. BTO and COBWEB both score highly in the data entry (Must) elements, but COBWEB shows no support for allowing an organiser to allocate area or polygons to individual user or groups. This may be less of an issue when performing an unstructured survey, but a stratified survey is likely to require the ability to manage and direct users to survey certain areas or habitat types.

User management: The combined must and should scores for user management are shown in Figure 6.3b. All vendors except Indicia support the 'must' user management features as being partially addressed but requiring minor modifications. The 'should' user management features score less well with only COBWEB identifying them in the same category. All vendors identify the verification elements as not addressed and requiring high cost modifications. This indicates that while the core of a user management system may be generally available, this would require modification and the addition of a bespoke verification system in order to meet the outline requirements.

Desktop habitat classification: The combined scores for the desktop habitat classification are shown in Figure 6.3c. Both the BTO and COBWEB score the same raw score, however the distribution of the scores in terms of the data entry and browser based requirements is almost the opposite. All of the other vendors identify some or all elements of this section as being not addressed and requiring high cost, major impact modifications.

This may indicate that the specific requirements of this survey are dissimilar to the existing solutions in terms of the data entry requirements for browser based systems. While the core of the solution is likely to be similar (classification of habitat) the way in which this is presented and managed would require significant tailoring in all of the solutions.

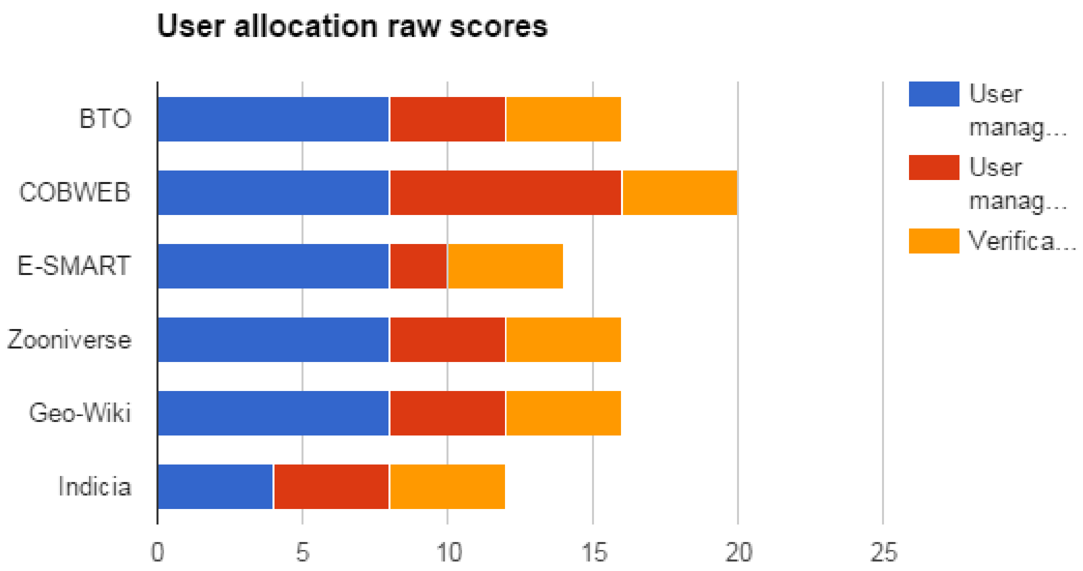
It is also worth noting that because of manner in which the functional sections are grouped this area has significant overlap with the user management scores.

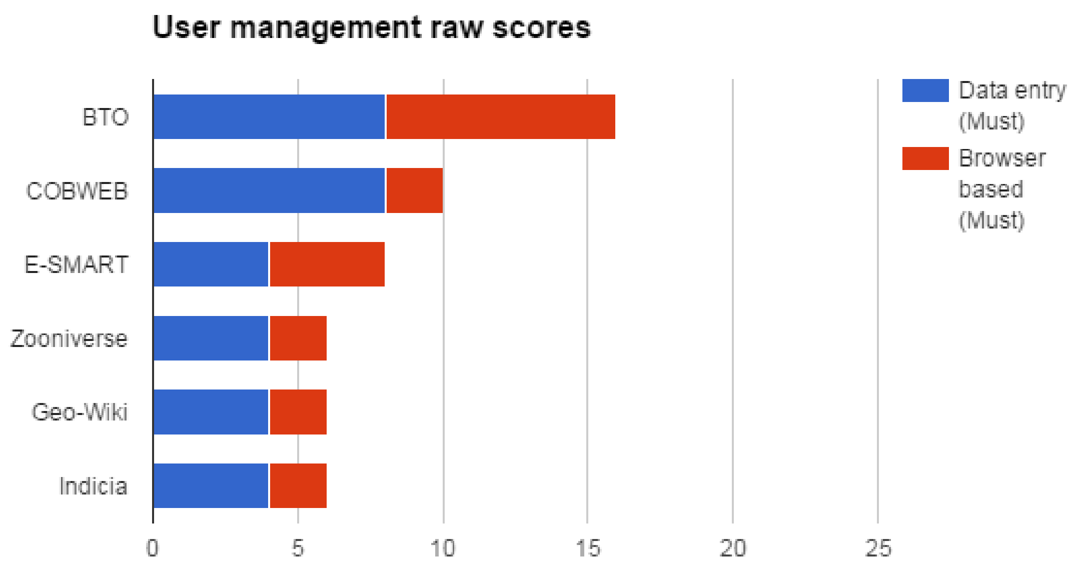
Mobile habitat classification: The combined scores for the mobile habitat classification function are shown in Figure 6.3d. Three vendors provide support for mobile platforms with COBWEB having the closest match to all of the requirements (Including an existing application that meets all of the core requirements)

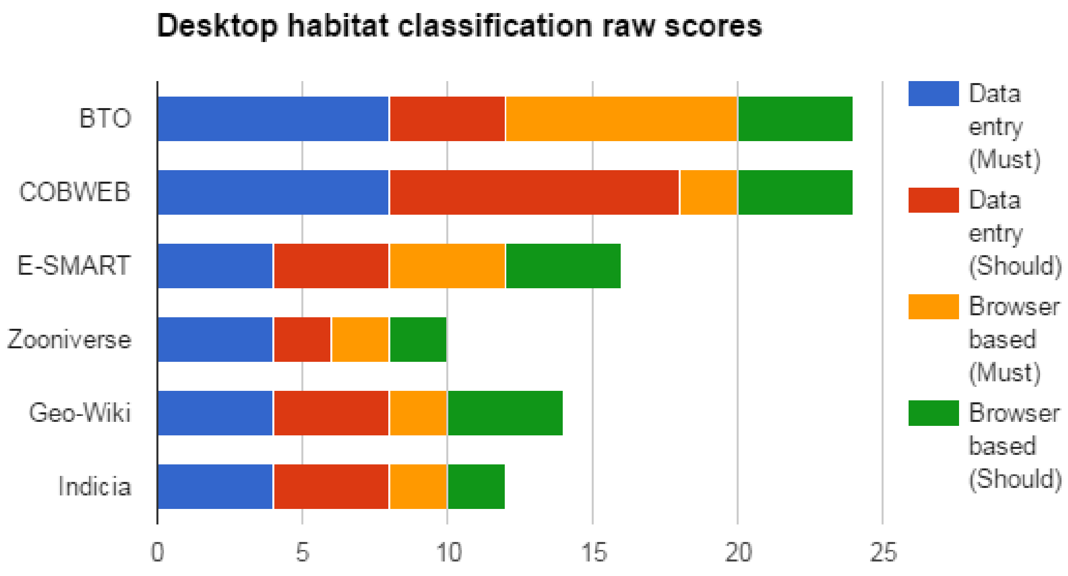
It is useful to note that none of the mobile support is provided as open source. This means that a system to meet all of the should requirements would require vendors to offer a version of their existing mobile solution. It also needs to be noted that a mobile platform needs to be integrated to the storage structure using Web Services, the mobile application alone is insufficient to provide a complete solution.

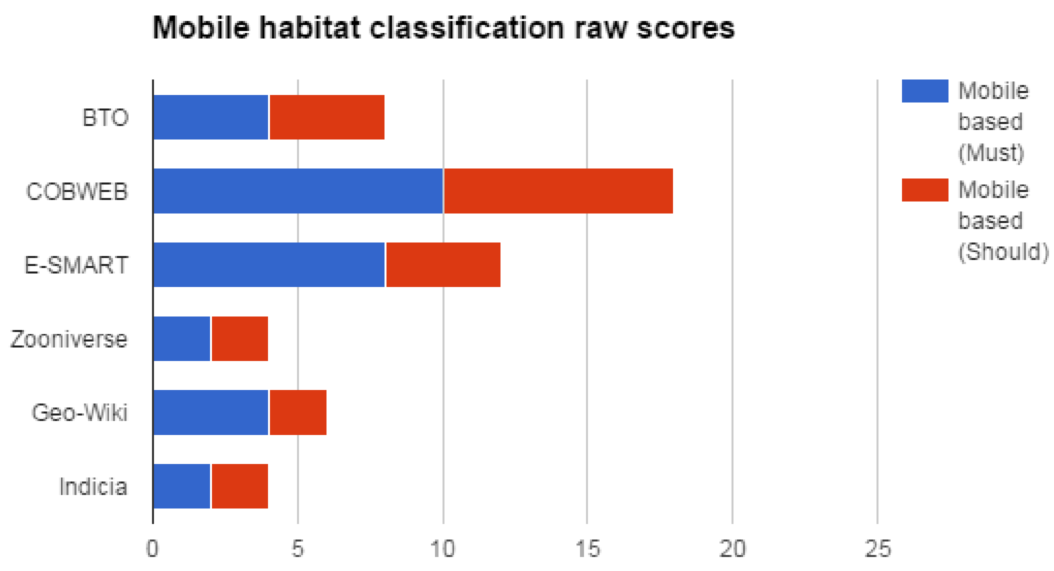
Data storage and feedback: The combined scores for the data storage and feedback functions are shown in Figure 6.3e. BTO, COBWEB and E-SMART all have a similar combined score, however the composition of the overall score differs between each of the vendors. The general weakness in all of the scores is in the user feedback elements. It is likely that the feedback calculations and mechanism would need to be tailored for the project, and whilst not weighted as mandatory (see section 6.4) lack of user feedback is likely to affect the long term engagement of the volunteers.

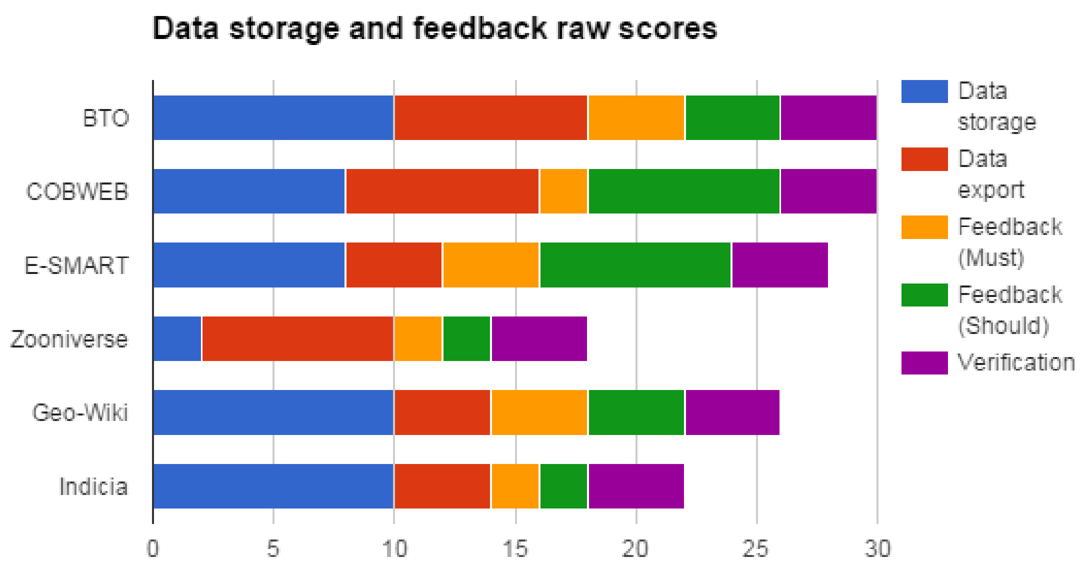
Figure 6.3. Combined scores for different components of the proposed solution for six vendors.






Desktop habitat classification raw scores
BTO
COBWEB
E-SMART
Zooniverse
Geo-Wiki
Indicia
0
5
10
15
20
25
Data entry (Must)
Data entry (Should)
Browser based (Must)
Browser based (Should)


Mobile habitat classification raw scores
BTO
COBWEB
E-SMART
Zooniverse
Geo-Wiki
Indicia
0
5
10
15
20
25
Mobile based (Must)
Mobile based (Should)

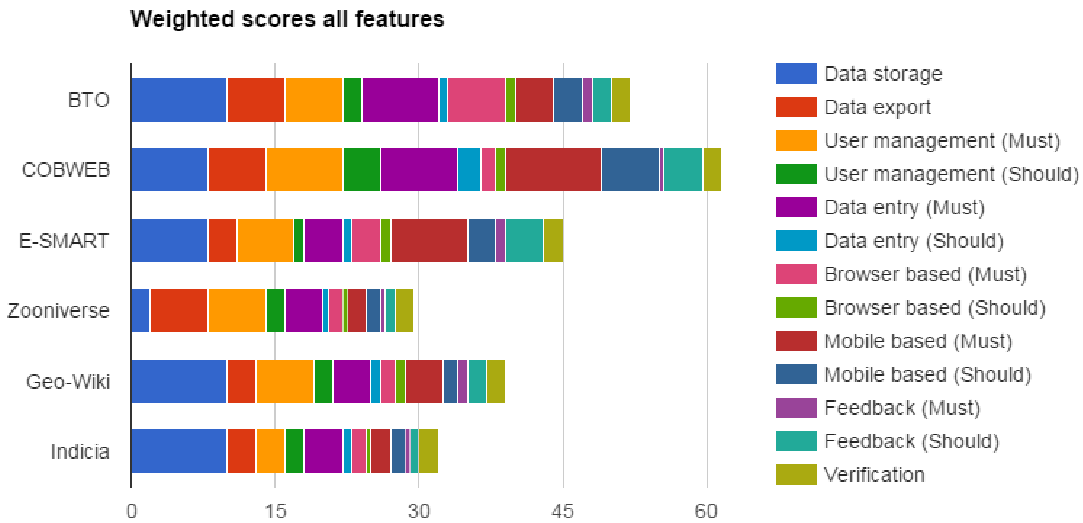


Weighted requirements: The combined weighted scores (weighting taken from Table 6.2) are shown in Figure 6.4. Based on a potential maximum value of 82.5 for all elements in the questionnaire, all vendors fell below 75%. When considering just the 'Must' elements with a total maximum score of 60 a similar result can be seen. This shows that irrespective of the nature of the solution (Proprietary or open source) there is a significant gap between the solutions available and perceived requirements.

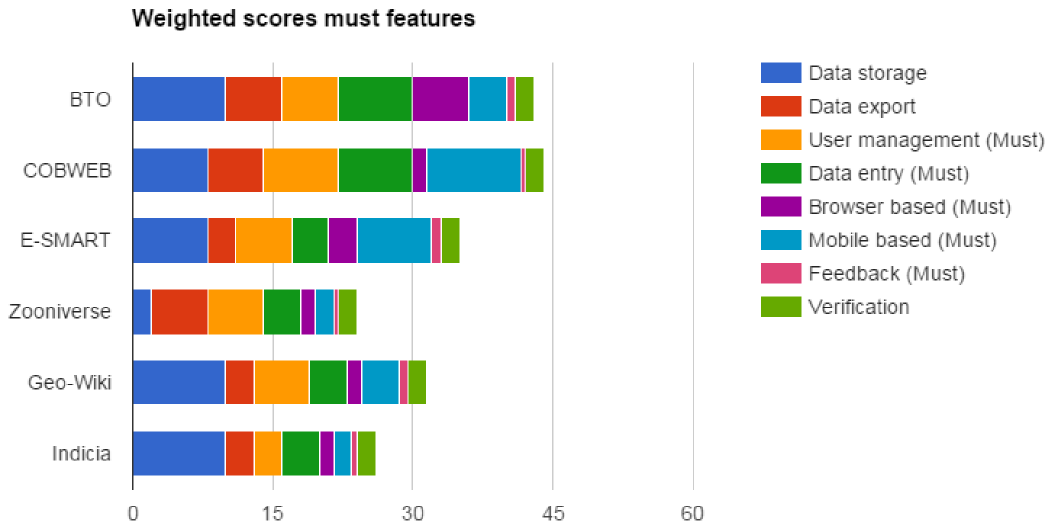
It also shows that the proprietary vendors provide a potentially better match than the open source vendors. This may be due to the fact that open source software is generally released as a generic solution to a problem, with the implementer or systems integrator tailoring to meet the specific requirements.

Figure 6.4. Weighted scores for the proposed solution using the weightings for different components in Table 6.2. Separate figures are shown for all features and for the must-have features alone.

a) all features



b) must-have features



6.6.3 Gap analysis

The gap analysis attempted to identify a preferred solution architecture and indicative costs for building a technology platform to run the project on. From the three completed replies, all three provided a preferred solution architecture and two provided indicative costs (BTO, E-SMART). The third (COBWEB) was approached for additional information on indicative costs but declined to respond. It was not possible to evaluate those vendors that did not respond as the questions relate to their operational costs and preferred solution based on their plans for development of their product

Preferred solutions: All three vendors identified a bespoke solution as their preferred route. While this may be related to the potential revenue for building the technology to run the project, it is also shows that the solutions do not generally provide standards-based (i.e. utilizing common and accepted official and unofficial industry standards) interfaces for exchange of data. This may be due to the nature of the implementers, as all have developed the systems as the basis for their own research programs.

Considering the potential bias for vendors to prefer their own solution, four cost scenarios are considered.

1. Open source GIS solution
2. Bespoke browser solution(s) (Must components only)
3. Bespoke browser solution(s) (All components)
4. Bespoke browser and mobile application solution

Note that Scenarios 1–3 consider only the desk-based part of the validation task to provide initial figures on how an incremental approach to the task could be taken according to financial constraints.

It should be noted that Scenario 1 was added for consideration at a late stage following receipt of volunteer information (Chapter 4) and statistical considerations (Chapter 5). As such it was not part of the formal vendor questionnaire and no comparative costs are available from different vendors.

Only one vendor (BTO) provided sufficient detail to review indicative costs for the remaining three scenarios.

Implementation costs:

Scenario 1 - Open source GIS based solution

It is possible to provide a desk based validation system comprising of a GIS project using an open source product such as QGIS (http://www.qgis.org/en/site/) or GRASS GIS (http://grass.osgeo.org/) This project could be pre-loaded with the appropriate layers for analysis. Copies of the project could be provided to individual users for recording the results of a desk-based validation task.

Having completed the validation task the users would need to export the data and provide this to a central point for consolidation and review.

The solution has no software development cost as it relies on the availability of open source software. Proprietary GIS solutions could also be utilised e.g. ESRI ArcGIS basic edition, and the costs would be dependent on the number of licenses deployed.
This solution would require time for an experienced GIS analyst to produce the GIS project, manage the distribution, validate the results and consolidate this into a final product. Time would also be required to train data entry users in the use of the product, as a general tool such as a GIS platform has a high barrier to entry for non-expert users.

Careful consideration should be made in terms of the costs of providing public information on the progress of the project to volunteers as this offline approach means that results are potentially not available until the end of the project. This shifts the main burden of the cost from technological development to staff costs around implementation.

Scenario 2 - Bespoke browser solution(s) (Must components only)

The cost of building an online system comprising of the Must components within the vendor questionnaire has been provided by the BTO. The indicative costs are:

- Creation of database schema, data import and data extracts: Approximately £20,000
- Provision of online system meeting the Must requirements: Approximately £75,000

This would give an indicative costs of a base online solution at £95,000.

These are an indication of the cost of providing the software and systems to implement the Norfolk Living Map project. They do not include the costs of organising and managing the survey or the cost of systems maintenance.

Scenario 3 - Bespoke browser solution(s) (All components)

The cost of building a complete online system comprising of all components within the vendor questionnaire has been provided by the BTO. An indication of the additional costs over and above scenario 2 are:

- Provision of additional weighted sample presentation, group management features, user preferences, public data pages in Should list: Approximately £25,000

These elements could be provided as extensions to scenario 2. This would give an indicative cost of an online system as £120,000

These are an indication of the cost of providing the software and systems to run the survey. They do not include the costs of organising and managing the survey.

Scenario 4 - Bespoke browser and mobile application solution

This scenario included all of the functionality from scenario 3, but with the addition of a field based validation system for the IOS and Android platforms. Indicative costs for this scenario were provided by two vendors (BTO, E-Smart):

- BTO: Approximately £150,000
- E-Smart: 400 person-days @ £500/day so approximately £200,000

These are an indication of the cost of providing the software and systems for the survey. They do not include the costs of organising and managing the survey, running the systems for the duration of the survey, or costs involved in communication or training.

Lead times for the project following requirements definition varied between 6 month and 1 year for the Norfolk Living Map.

Both of these figures indicate there is a significant gap between the technology currently available and the requirements outlined in the questionnaire.

Given these vendors provided the highest match to the weighted requirements it can be assumed that other solutions are likely to incur higher initial costs because they would require more work to bring them up to the same level.

Scalability: Considering the scalability of the scenarios, Scenario 1 - Open source GIS based solution is very different to the online scenarios 2-4.

Scenario 1 - Open source GIS based solution
This is not a scalable solution. The cost of each area setup would be incurred each time a new area is added to the project. These costs would include the setup of the GIS project for the area, training the users, and potentially housing the systems for the users to work on in some regions.

Scenarios 2–4 - Online systems
The ability to scale the solution across the UK depends very much on how the rollout is performed. The two vendors provided differing costs for the rollout phases:

- BTO: Up to 150 simultaneous users would incur no additional costs
- E-SMART: UK rollout would require additional server and backup support at approximately £30,000

It is worth noting that none of the vendors identified cloud technologies as an option for scalability. This may be due to a number of factors:

- Cost of the solution is elastic with demand. This means that the cost is dictated by the amount of resources (CPU, Disk Space etc) used and these are added on demand without an upper limit. This makes the Total Cost of Ownership (TCO) difficult to calculate. An example is the Amazon Web Services cost calculator (Amazon Web Services Simple Monthly Calculator) in which the monthly bill is based on the data transfer volume in and out of the Web service
- Data ownership and locality of data are perceived as potentially complex issues
- BTO and E-Smart already have support of high bandwidth university data centres

Maintenance costs: Only one vendor provided an indication of on-going system maintenance costs:

- BTO: £2000 per year

This is an indication of the cost of hosting and providing support for the online systems identified in cost scenarios 2–4.

It is worth noting that mobile technology turnover is high, and a project lifecycle longer than 5 years is likely to incur additional costs of mobile application development in order to keep pace with current devices. No vendor provided an estimate for future development work.

Summary: Given the cost scenarios outlined, the eventual solution must balance three factors:

- Systems development costs
- Survey setup and management costs
- Volunteer engagement

The Scenario 1 - Open source GIS based solution has a very low system development cost, but has potentially high setup and management costs. The volunteer engagement is limited, and careful consideration must be made to how manual validation and verification methods are employed.

Scenarios 2–4 incrementally increase the initial development costs of the system. In contrast, the survey management costs may be reduced. This may be achieved as scenarios 2 and 3 provide improved user feedback and automated allocation features aimed at reducing the human intervention required in the validation process. Scenarios 3 and 4 both have the potential to improve the volunteer engagement by providing reporting and progress tracking. Scenario 4 also has the potential to improve data entry accuracy by reducing the chances of transposition error between paper based field notes and the browser based desktop system.

In terms of providing systems to meet Scenario 4 (Vendor based desktop and mobile solution) initial costs would be upwards of £150,000 with potential operational maintenance costs of £2000/year for server hosting, backup and archive etc. This is the cost of providing software and systems and does not include survey management costs. Although extending the validation task to the whole of the UK at any one time is unlikely, doing this would incur an additional £30,000+. This figure is based on the E-Smart response and is incurred in terms of server capacity and additional online storage capacity. If the project were to be incrementally rolled out with a smaller set of regions active at any one time, and the data for completed regions being stored offline, these rollout costs may be reduced. Again, this is the cost of providing software and systems and does not include survey management costs. These costs are significant and there could be a perceived risk that this level of investment is risky if insufficient volunteers can be recruited to the task. Such considerations should be evaluated in a risk log by any parties tendering to run the validation task.

A final additional cost for scenarios 2-4 may be future software updates for technology changes over the lifetime of the project. This is particularly true for mobile technologies where the platforms are still relatively immature and there is a high rate of change.

6.5 Transferability

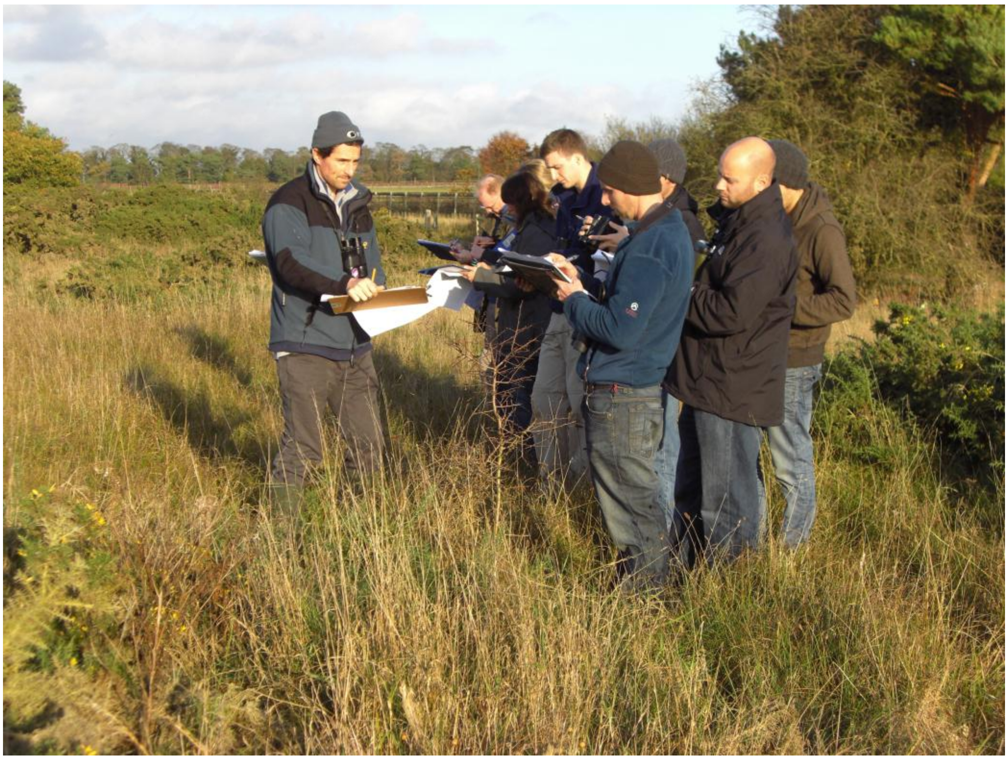
The technological review has considered the general requirements of a solution that is not associated with a particular region. As such the findings are directly transferable to other regions of the UK.

6.6 Conclusions and Recommendations

- There is no off-the-shelf solution that meets the needs of the project. The gap between the perceived project requirements and open-source products is significant and would require considerable systems development to achieve a solution. The gap between the perceived product requirements and proprietary solutions is less, but still considerable.
- Successful delivery of the project would require tailoring of one or more existing systems and allowing these to be coupled together to form a suitable solution. Existing vendors indicate the preferred route would be a bespoke implementation of an existing offering without integration to third party systems.
- Only two of the existing vendors identify their capacity to scale the project from the Norfolk Living Map to a UK solution.
- A number of the solutions, both proprietary and open source, are university research projects. Care would need to be taken that solution scalability is fully assessed prior to implementation.
- Two potential routes to providing a solution are:
    - Use an existing open source GIS solution with trained data entry staff and field based volunteers using paper forms. This has a low cost of development but will

incur significant costs in user training, volunteer management and scaling to other regions.
  - Product an online system based for desk based validation and optionally field based validation using mobile phone technology. This could be achieved in two ways:
    - Approach existing vendors that provide systems capability and work with them to extend their solution to meet the needs of the project. This may include tailoring of both software and infrastructure to meet the needs.
    - Take existing open source software and a vendor with the capability to extend this, as well as the infrastructure to implement the solution for the project, and commission a development and maintenance contract for the project. The resultant implementation may then be moved back into the public domain.
- The costs for producing a GIS based solution are not provided due to the late consideration of the scenario and the fact that overall costs are dependent on the structure of the validation task and fall on staff managing the task rather than on the development of the technology per se.
- Care must be taken on identifying licensed map products to support validation. Without an appropriate strategy for intellectual property management it may not be possible to provide the results as a public resource. It is recommended that source mapping products be evaluated for OGL compliance, and vendors are provided with a set of requirements around intellectual property prior to commissioning the survey.
- The indicative costs for producing an online solution are estimated to be upwards of £150,000 with an annual cost of at least £2000 per year to run the system. Whilst it is unlikely that the task would extend to the whole of the UK, adding multiple regions, or expansion to the UK from the Norfolk Living Map may be upwards of £30,000.
- A formal set of requirements would be required to approach a sub-set of vendors in a tender process.

7 Communication strategy and knowledge exchange



Caption: Recorders on a training course (Credit: Su Gough)

Success of the validation task is dependent on effective communication that extends over the whole project, from the initial planning and recruitment stages, during the survey to the final reporting. In this Chapter we have looked at the issues that should be considered at each stage of the project.

7.1 Key stages of communication

It is important to present very clear messages that outline the aims of the project and explain what it will deliver. These should, for example, include messages on why habitat validation is needed, why volunteers are needed and how the data will be used (*7.1.1 understanding the ask and data uses*). It is important to understand the likely volunteer base and their motivations (*7.1.2 understanding audiences*), so that appropriate communications can be developed, such as those for recruiting volunteers (*7.1.3 recruitment*). An example of a broad communication message (perhaps the central call within a press release announcing the project), might be along the lines of:

> *Knowing which habitats make up a landscape, and how they are distributed, is essential for successful conservation at a landscape scale. High resolution imagery, captured by satellites, has the potential to deliver this information but volunteers are needed to 'ground-truth' the imagery before this new technology can be rolled out more widely. It is hoped that the resulting 'living map' will support farmers, conservation practitioners and landowners as they seek to manage their land and balance the needs of wildlife, agriculture and other commercial interests.*

The communications strategy should include a sufficient lead-in time, with identifiable goals and a clear timeline. It should also identify appropriate metrics by which the success of different approaches can be assessed, thus allowing approaches to be shifted in response to what works best. During the project it will be necessary to communicate what is required of volunteers and to provide material to support this, which may include training and other information (*7.1.4 skill training and support*), and to consider what resources are needed to support this (*7.1.5 supporting volunteer* *preparation*). Feedback will need to be provided to volunteers during the project, to motivate volunteers and encourage further uptake (*7.1.6 feedback during the project*), and at the end of the project to show to the volunteer how their contribution has benefitted the project, data flows back from the volunteer to the database (7.1.7 *data flows from the volunteer to the database*), and to highlight the value of the project as a whole (*7.1.8 feedback after the project*). Lastly, the results of the overall validation task need to be communicated to different stakeholders (*7.1.9 Dissemination of data and results*).

7.1.1 Understanding the ask and data uses

It is important that the project clearly communicates from the outset its purpose and why volunteers can and should get involved. A clear explanation of exactly how the data will be used in the short and long term may be difficult to produce, but it is an extremely important message to get across. For example, in the volunteer interviews (Chapter 4), there was some suspicion from the farming community on how data relating to farmland would be used. Behind this is a concern that more data may lead to more restrictions on farmer's activities. This could affect participation by farmers and their willingness to permit access by other volunteers. It will be important to work closely with FWAG to ensure that the project is able to get the right messages across. Farmers may be keen to participate directly and could validate entire farm estates. This could be useful to unstructured data collection but is less beneficial for the structured sampling. Some may question the impartiality of data supplied by farmers but the same could apply to data supplied by community groups who may be keen to prevent local development. Clear messaging on the importance of accurate impartial data is critical.

It is also potentially quite a complex message to get across to volunteers, that in the long term, results from this project will feed into improving the habitat map data, but that the habitat map probably will not be updated in the short term. Greater clarity is needed on how quickly the validation work will feed into improvements of the map and the longevity of the project. If this cannot be achieved there will be significant impacts upon volunteer engagement with the project and potentially spillover effects into volunteer groups willingness to participate in other projects.

Similarly, communicating the need to validate but not enhance the map will be critically important. Apparently erroneous parcel boundaries and a poor fit between the map's habitat classes and the sort of classification a local community would prefer, are key issues that will require careful management. Volunteers must be helped to understand that the target level of accuracy for a map of this type is not necessarily 100%, and that even after the validation results have been used to improve the map, errors will still exist. This is a significant communication issue arising from the difference in perspective between nationwide (or even countywide) mapping ("99% of all parcels are correct") and local use ("my parcel is wrong").

Communication of field methods is always important but in the context of this project, where engagement may be with people new to biological recording, explaining concepts such as random sampling will be important. Interviews suggest the new volunteer audiences will want to participate in their local area at a site of their choosing, which may not be consistent with a required structured sampling method.

In addition to getting volunteer buy-in, it is important for the volunteer to be able to correctly portray the project to others, which could help promote the project. They will also need to be capable of explaining the purpose of the project to landowners, which will be important for obtaining access to private land. This further emphasises the importance of presenting clear messages on the project and data uses.

7.1.2 Understanding volunteer audiences

In Norfolk, NBIS has a good network of contacts with a broad range of communities and groups, many of which were interviewed as part of this project (see Chapter 4). We would recommend that the first step in any project like this, is to build as broad a contact list as possible of potential volunteers, communities and local groups to approach, and to try and understand the likely motivations and potential interest of these groups for taking part in habitat validation. It is also important to understand what benefits the volunteers will or could get from taking part in the short term and potentially in the long term. This knowledge will help in deciding how to explain the task (Section 7.1.1) and which aspects to highlight to different groups, potentially with subtle tweaks to optional aspects of the task. For example, the current habitat categories in the Norfolk map, may not necessarily be of great interest to the volunteer if they can not, for example, record the ancient wood or meadow in their village. Consideration should be given to the potential for volunteers to record these, even if they are not used in the validation process.

It may be challenging to provide immediate benefits to participation (but see below). In the longer term, there are clear messages around benefit that the completed Living Map will offer, such as understanding flood risk, better protection of important habitats and species. Discussion with JNCC/Defra would be needed to identify these benefits that would stem from the map (linked to planned uses, see Section 7.1.1).

7.1.3 Volunteer recruitment

Recruitment for the desk-based component should be relatively straightforward because it is designed to appeal to, and make use of, "high input" volunteers who are already heavily engaged with local organisations. It may only require c5 volunteers from each of the Norfolk Wildlife Trust and NBIS. Nevertheless, the work is more demanding in some respects so a clear "job description" should be drafted to include the necessary computing and organisational skills required and an indication of the level of time commitment needed. Training will require a major investment on the part of the organising team which would be wasted if volunteers drop out because the scale of the task was not adequately explained to them.

Developing a strategy for recruiting volunteers for the field-based validation is likely to be particularly challenging because its success is contingent on a high level of engagement to achieve likely coverage targets. Experience from other projects suggests local promotion is particularly effective. This is consistent with the likely future roll out being on a county-by-county basis. Once clear messages have been formulated, these should be communicated through local organisations to their members and volunteers. Articles in local press, regional newsletters, on local blogs and social media are likely to be most effective in recruiting wholly new volunteers. However, volunteers in existing schemes should not be neglected, provided the (currently) restricted nature of participation can be communicated satisfactorily. This again calls for clarity on future roll out plans. The local organisers should prepare oral presentations and posters that can be given at appropriate fora to encourage participation. Such publicity can also go a long way in promoting the project to landowners who are already "warmed up" to the idea before being approached to provide access.

For field-based validation, we anticipate that publicity and promotion will direct potential volunteers to take part through an online system, which would coordinate the sign-up process. Most work should be put into encouraging sign-up of the stratified random squares, with the expectation that volunteer uptake to survey an individual's local area or parish will be much easier to achieve. One option would be for people to take on a randomly selected square within a reasonable travel distance of their home, but if this is not possible, they can choose a radius within which to choose squares. A common finding is that rates of sign-up are higher than rates of actual participation because potential volunteers change their mind. This makes assessing progress difficult but could be alleviated by videos or blogs providing a better impression of what is required.

In relation to transferability to other areas, it is likely that uptake will be greatest in areas where there are more people. In Norfolk, the density of people is low in the Fens in the west of the county, which has always been a problem for biological recording. Elsewhere, obtaining sufficient coverage in remote upland areas, is likely to be a greater problem still, and in addition to targeted promotion to particular individuals or groups, for example hill walkers, it may be necessary to consider paid fieldworkers in some of the most different areas for volunteers to visit.

7.1.4 Skill training and support

The validation task will be a new experience for the majority of volunteers so it is essential that they understand what is being asked of them and that they have sufficient information and support to do the job properly. Support should include clear written instructions backed up with tutorials, training videos and clear habitat descriptions with supporting material such as photos. As a starting point, Natural England (unpubl. report) has produced a habitat key for identifying broads, priority and Annex I habitats which could be adapted for this purpose. Targeted training material, with examples where difficulties in identification are most likely, should be provided. A continual learning and self-evaluation process could be built into online systems, and should be integral to the desk-based approach. This would allow volunteers to reflect on the information they contribute. A support email address should be provided but the level of staff resourcing needed to handle email support enquiries should not be underestimated.

In relation to the handling of quality assurance issues, misidentification in biological recording is a very sensitive topic, and it will be important to consider very carefully how to minimise and quantify errors in the volunteer-generated data, whilst maintaining good relations with the different communities taking part. For example, having multiple “high-input” volunteers validating the same habitat parcels through desk-based validation is straightforward. Multiple field surveyors validating the same squares presents a communication issue: observers may expect to see a coverage map updated with their field effort, not see a square still showing as uncovered until someone else has also checked it. There could also be issues related to land access permission if more than one volunteer asks for access for the same area to do the same thing. Publicity material should try to alleviate these problems upfront, perhaps by highlighting how different observers see different features in the landscape and that habitat classes are an interpretation of a continuum of vegetation and land use gradients. But this is a difficult concept to communicate and it may be better to invest resources in one or two ‘floating’ professional surveyors who could provide the repeat visits (and potentially infill coverage gaps in poorly populated areas).

7.1.5 Knowledge exchange with volunteers before and during fieldwork

Preparation before venturing into the field is important for effective surveying and for the morale of the surveyor. In addition to the training material detailed above, preparatory facilities and materials may include the ability to download and print maps of squares, downloading parcel boundaries and definitions to a smartphone, or providing a “letter of support” to help when arranging access to private land. At some stage in the validation process, volunteers will need access to the existing classification of each parcel in their square, but the information may not be presented to them until key actions have been taken. For example, validation independence will be maximised if we do not tell volunteers what “the computer” thinks a parcel is at the outset. If this workflow is followed volunteers should be provided the classification after they have made their assessment, with the opportunity to revise their view if necessary. Importantly, their initial choice should be saved in addition to any revisions. These points assume that all parcels have a unique identifier. Currently parcels in the Norfolk Living Map are merely numbered from 1 to 4,441,282. For future proofing, more robust identifiers are required that indicate the parcels belong to the Norfolk Map and that they are from version 1 of the map (in this case).

To increase the appeal of the validation task we recommend a set of information sheets be created to explain other features of interest that volunteers visiting particular habitats or landscapes could encounter. These could focus on providing information on the importance of the habitats / priority habitats that are being surveyed. For example, volunteers surveying Lowland heathland might be interested to know to look out for Adders and Woodlarks. Such information may be familiar to regular biological recorders but for the validation task to be a success requires engagement with new volunteer groups who may benefit more from such ancillary information. In addition to habitat-specific information, it may be possible to provide location-specific attributes. This could include interesting archaeological features (e.g. standing stones) or biological data. Scripts written by CEH allow the automated extraction of biological records from the National Biodiversity Network database. These could be adapted to provide information for focal squares which could increase uptake and enjoyment for volunteers. Whilst not a direct aim of the validation project, this could in turn lead to recruitment to biological recording and encourage the submission of biological data.

7.1.6 Feedback to volunteers during the project

Feedback to volunteers during the project is as important as feedback at the end of the project; it encourages volunteers to complete allocated tasks, enhances fulfillment and acts as a positive feedback mechanism to encourage further effort. Working with communities and local parishes, it may be important to provide an online version of the whole map to encourage people to sign-up to fill in gaps in survey effort, and potentially for people to download and make use of these data for their area. An online version of the map would need to be a raster image due to OS Mastermap licence restrictions.

If the validation task takes more than one season (or year) to complete, interim results should be provided at the end of the season. Prior to the next season, existing and potential volunteers need to be re-energised, reminding them what has been achieved but making clear what remains to be done and why it is important to complete the job. Explaining to volunteers when it is and is not too late to take part is important.

Multiple channels should be employed for in-project feedback, including social media, a project blog and emails to signed-up participants and community groups. Feedback should include coverage statistics, maps showing success and gaps in survey coverage and potentially a league table of participants. Emails should be as personalised as much as possible, making use of the volunteer's levels of allocation and coverage to strike the right balance of thanks and encouragement to complete the job.

Through this review, we have steered away from focusing too heavily on gamification. The value of gamification will depend on the motivations of the volunteer or volunteer group, but it is felt that if people have a reason anyway for taking part, gamification is less critical. For local areas using unstructured field validation, completing the validation of all parcels in a parish in an area is a game in itself.

7.1.7 Data flows from the volunteer to the database

It is not possible at this stage to fully define the data flows because the final implementation of the validation task is dependent on external factors (e.g. validation goals, funding). However, we can identify certain aspects or possibilities as follows, broken down into the desk-based and field-based solutions.

Two options are suggested for the desk-based solution, use of a bespoke online solution or use of standalone GIS projects. The former would likely comprise a number of related data table:

- Volunteer data table: unique user identifier (primary key), password, contact details, dates of completion of training, contact preferences (data protection requirements).
- Validation data table: unique parcel identifier (primary key), user identifier (to indicate who provided the row of data), date of validation (to cross check with version history and to permit repeat assessments), classification (true/false/unknown), replacement class if

classification judged false, classification certainty (score based on scoresheet for user's certainty of true/false answer), replacement class certainty, indication of which ancillary datasets were visible during validation.

The standalone GIS project approach would require essentially the same data be collected but the volunteer data, and the ability to tie particular validation results to individuals, would require manual processing of GIS files by a volunteer manager.

For the field-based solution we have recommended use of a bespoke online system and the following data flow suggestions apply whether this is implemented as a desktop application, a smartphone application or both:

- Volunteer data table: unique user identifier (primary key), password, contact details, dates of completion of training, affiliations to local groups (e.g. NBIS volunteer, rambler), contact preferences (data protection requirements).
- Visit data table: unique visit id (primary key), user identity, the survey date (to cross check with version history, to permit repeat assessments and to check seasonality of identification), location (500-m square reference), data submission mode (desktop or smartphone), selection method (structured or unstructured).
- Validation data table: unique submission id (primary key), visit id, unique parcel identifier, classification (true/false/unknown), replacement class if classification judged false, classification certainty (score based on scoresheet for user's certainty of true/false answer), replacement class certainty.

The examples above assumes a simple true/false method of validation. For example, having provided the volunteer with the current habitat class they would indicate "Yes it is habitat X" or "No, it is habitat Y". We advise against this approach as it does not afford full independence of assessment. Instead a hierarchical approach could be taken, for example, through a stepwise process the volunteer defines the habitat as Grassland > Wet > Coastal. The application could then present to the volunteer the habitat that the parcel is currently assigned to, and the volunteer given a chance to respond, and change their decision. If this approach was adopted the validation table would require additional fields to record each habitat level provided. Importantly, these would also store any revisions made by the observer. Optional columns may also allow for enhancement to the data to be recorded, for example "I saw marsh orchids" or "it was grazed by cattle".

For all systems, volunteers would be required to accept simple terms and conditions. These would include agreement that all data collected on the validation task would be held under OGL (with the exception of personal data).

7.1.8 Feedback to volunteers at the end of the project

The validation task will involve a lot of effort from a large number of volunteers, many of whom could be new to citizen science. It is important that they feel valued and that their data have, or will be, put to good use. Feedback at the end of the desk-based and field-based validation tasks should focus on the achievements - numbers of parcels that were validated, number of people who took part and any interesting and personal stories, e.g. new sites for particular species found as a by-product. A summary should be given of the key validation outcomes, providing accessible results such as "99% of Arable fields in the Living map were correctly classified but 20% of Bracken parcels were actually Lowland heathland. Lastly, volunteers must be given an impression of what will happen next. Will there be another round of validation following a re-run of the map? They will want to know that their efforts have been worthwhile. Explaining how the data will now be used and what they can expect to see change and on what timescale is crucial. They also need to know if they can take part again in Norfolk, or whether there are other regions where help is needed.

In addition to the successful completion of the validation task, the project has the potential to leave several legacy benefits. It should result in a substantial increase in the number of people who are aware of the existence and value of the Living Map; volunteers must be made aware of how they can use it in the future. Building the volunteer base will have required significant effort; this should be capitalised upon by asking volunteers whether we can contact them in relation to other recording projects. Typically this requires carefully worded questions during the sign-up process so as to meet Data Protection rules without being overly off-putting for novice recorders.

7.1.9 Dissemination of data and results

Once all the data have been gathered and analysed all the underlying data will need to be deposited at the Collaboration Node where they can be accessed by parties responsible for adjusting the remote sensing algorithms prior to the creation of a new version of the map. It is important that any existing users of the map are informed of any significant issues with the map. To facilitate this, prior to use all prospective users of each Living Map should register to get their copy. Although the data would still be free to use, this would ensure a central list of Map users is collated to enable issues and updates to be communicated. We have followed a presumption that any data collected during the validation task will fall under Open Government Licence but if ownership and access to the Living Map data differ, this will need to be clarified so that a data strategy can be developed for data arising from the validation task.

7.2 Transferability

The communication strategy considers the basic motivations and needs of volunteers and necessary data and information flows. As such the majority of these points are directly transferable to other regions of the UK. One difference, however, concerns who will implement the communication strategy. In Norfolk, NBIS has a long-standing involvement in the Living Map process, they have existing volunteers and are well connected with local recording communities. Therefore, in Norfolk NBIS might desire to act as the local organiser. In other regions this organisation capacity or existing investment may not exist. This decision may also have a bearing on how technological solutions are developed. If these are outsourced to technology vendors with no interest in volunteer engagement there is a clear need for a local organising team. In contrast, vendors such as BTO and CEH have a long track record of volunteer engagement and proven capability to develop bespoke technology solutions and it may make more sense to retain these linked tasks in one organisation.

7.3 Conclusions and Recommendations

- Clear message on why the project is needed, why volunteers are required and how the resulting data will be used are critical.
- Vendors tendering for a future validation task would need to provide a risk log considering aspects such as failing to recruit sufficient volunteers to use expensive bespoke technical solutions.
- Explaining the differences between validation and enhancement, and what an accurate map might look like will be important in managing expectations.
- A project such as this will need to explain concepts such as random sampling to justify the chosen methods to new volunteers.
- Volunteers are great ambassadors so providing them with the facts will help promote the Living Map.
- Before the start of the project it is important to build a broad contact list of potential volunteers and groups, to understand their motivations and to identify how the project will benefit them.
- Local promotion is important and should use a variety of media whilst not neglecting national volunteer groups with local presence.
- Promotion should focus on making people aware of the project and how to sign-up and request a square.
- Volunteers should be provided with training and other material to maximise updates and the quality of data collected.

- Volunteers should be supported with information and materials to make the process of taking part as simple and as engaging as possible, including information on features to look for when visiting a location. As a starting point, Natural England (unpubl. report) has produced a habitat key for identifying broads, priority and Annex I habitats which could be adapted for this purpose.
- Support is critical for a project involving new methods and new volunteers.
- It will be important to provide engaging feedback to volunteers during the project to motivate and encourage further volunteer uptake.
- Results should be provided to participants at the end of the project, or end of each field season (and prior to the commencement of the next) if the project runs for more than one season.
- The communication strategy should consider the potential of the volunteer-base to take part in following years, or if carried out over a single season, the legacy of the volunteer-base.
- Upon completion, data would be marked as structured or unstructured and stored in the Collaboration Node ready for production of revised versions of the Map. All known users of the existing map would be alerted of any issues emerging from the validation task.

8 Overall conclusions: feasibility of validating the Norfolk Living Map with volunteers

Caption: Juxtaposition of different habitat classes (Credit: Simon Gillings)

Aspects such as the distribution and aggregation of habitats, density of parcels, pixelated map appearance and ease of identification of habitats make the validation of the Norfolk Living Map a challenge. Volunteer communities are potentially interested in the task but there may be a mismatch between willingness to engage in local validation and the type of structured recording that would provide robust data. There is also a keenness to validate large areas (e.g. parishes, farm estates) whereas from a statistical standpoint it would be better to validate a larger number of smaller areas. There is a danger that some of this interest could be based on a false impression of what the Living Map looks like at fine spatial scales. There may be a keenness to enhance the map which is beyond the scope of the validation task but may be necessary to secure local buy-in.

From statistical and practicality perspectives the validation task can be divided into two components: a field-based task complemented by a desk-based component. The latter would concentrate on providing validation of superabundant habitats plus any priority habitats that are easily identified remotely. This would require a small number of well-trained and skilled volunteers and is in keeping with the results of interviews that suggest a desk-based method would have limited appeal.

The field-based component would deal with the remaining habitats but it is here where the biggest challenge may arise. The priority habitats that are the focus of the Living Map are rare and spatially aggregated. Random squares do not capture them without unfeasibly large sample sizes. We have not tested this but it is unlikely that allowing observers to choose nearby survey locations will provide the necessary coverage of rare habitats because rare habitats do not occur where people live. The most robust validation results can be obtained with random sampling stratified by habitats which will be difficult to sell to volunteer audiences who want to map their local area. Securing sufficient volunteer capacity for a structured field-based sampling protocol will be a major challenge. Extending this approach to provide zone-specific validation is unlikely to be feasible unless a significant number of habitats can be safely validated using the desk-based approach, or using data from existing schemes; we judge the feasibility of the latter to be low in Norfolk at the current time. Pilot deskwork and fieldwork is urgently needed to confirm the ease of identification of Norfolk habitats and to develop identification material.

For the field and desk-based tasks, it will be important to build in procedures for collecting information that will facilitate quality control of volunteer data. This should include using multiple volunteers to validate a sample of the same habitat parcels for both the field and desk-based components, and using control sites in desk-based validation where the habitat has been validated by experts. In addition, it would be useful to record information on volunteer experience and training, for the volunteer to flag when they are not confident in assigning a habitat, and for the field survey to record survey date (considering seasonal difficulty in habitat identification) and where the volunteer lives in relation to the survey (local knowledge).

Technology will be central to the validation task. A desk-based solution is relatively easy to envisage although the provision of additional data layers may require some preliminary work. The biggest issue may be the development cost which will seem high compared to the small number of users. There must be a high expectation that the technology will be reused in other regions. A GIS-based solution would be cheaper to develop but would require higher staff costs in administration, and these costs would be duplicated in each new area. To overcome the perception from volunteer groups that smartphone applications are not useful for fieldwork requires that their benefits are clearly sold over paper maps. To some extent this comment may stem from a miscomprehension of what the Map looks like (a dense array of complex pixelated polygons with no reference points). Whilst any future technology would need to provide printable maps to satisfy demand, we recommend a well-designed smartphone app will provide the ease of navigation and data entry that is required in the field.

Much of the success of the validation task rests on a successful communication strategy. If clarity of purpose can be established, benefits explained and methods justified, recruitment of volunteers for the project should be achievable. Volunteers will require significant support and training otherwise the required field-based sampling targets will not be met. Expectation management will be a major challenge, and reining in the keenness to undertake many large self-selected validation areas in order to redirect a proportion of volunteers to structured recording will be important. On balance, there may be a requirement for one or two paid professionals to undertake repeat visits for quality assurance assessment and as an insurance policy to fill coverage gaps.

It is worth noting that our conclusions derive from the specific of the Living Map solution to habitat recording, where Maps will be produced separately for individual counties or regions in the country, each of which may differ in the habitat categories and methods used, and will therefore require separate validation. As this study has shown, the likely sampling effort required in any one region will be high relative to the potential pool of volunteers. If volunteers are seen as the route by which these Maps are to be validated, then it is worth considering how future Living Maps should be designed in order to optimise the practicality of this solution. Although outside the scope of this report, broadly speaking the greater the spatial area across which common methods, imagery and habitat categories are applied, the lower the density of volunteers required to validate each habitat category. For example, a required sample of 50 squares containing Lowland heathland could be spread across five counties rather than requiring 50 squares in each of five counties, with clear consequences for the number of volunteers required.

In relation to transferability, general considerations relating to field and desk-based approaches, the sampling design and communication strategy should be transferrable to other regions of the UK. However, what will vary are the habitat classes considered in future Living Maps and the precise mix of habitats at the sample level. For any new region, it will be necessary to consider habitat-specific sample sizes when designing a new survey, but there is unlikely to be a problem if the scale of parcels is quite similar to Norfolk. Methods may require adaptation in western pastoral systems where improved pastures are very abundant, presenting a problem similar to the “urban” problem in Norfolk of inflated numbers of parcels for field validation. In such cases it will be necessary to shift the burden of validation to the desk-based component, provided such habitats can be identified remotely. In Norfolk high co-occurrence of different habitats within 500-m grid squares, significantly reduced the total sample size required for the validation task, but this level of co-occurrence and consequent saving on sample size cannot be guaranteed in other regions. The effectiveness of any volunteer-based survey when transferred to another region will also depend on the number of likely participants with an interest and willingness to produce local maps which, in turn, is likely to be a function of the number of residents. Although volunteers can be encouraged to travel to survey locations remote from where they live, participation is highest in the vicinity of the home. We anticipate that obtaining sufficient coverage in remote upland areas is likely to be particularly problematic. Targeted promotion to particular individuals or groups, for example hill walkers may help, but it may be necessary to consider paid fieldworkers in some of the most difficult areas for volunteers to visit. Whilst the communication strategy considers the basic motivations and needs of volunteers and necessary data and information flows, which are directly transferable to other regions of the UK, success may be influenced by whether a local organisation exists with a willingness to implement the communication strategy, or whether this has to be done by a national body more remotely.

Several actions are required before the validation task can be implemented:

- Clarify the overall purpose of Living Maps, the validation task, plans for roll out and map updating to enable clear messaging to volunteers and stakeholders.
- Clarify data ownership and licensing for the overall Living Map products to enable similar data strategies to be developed for data derived from the validation task.
- Clarify whether certain habitat classes can be merged or omitted or from the validation task because they are either low priority or there is existing high confidence in their accuracy.
- Remove obvious artefacts from the Norfolk living Map data before use by volunteers. This will require professional GIS work.
- Confirm licensing requirements for underlying OS Mastermap data and whether the Living Map can be provided to validators in a less pixelated form.
- Clarify the naming and definitions of habitat classes and provide best practice guidelines for ensuring comparability of habitats between existing and future Living Maps.
- Commission a pilot field survey to confirm ease of identification of habitat classes by volunteers, and develop and test identification and training material.
- Build on existing work by NBIS to create a definitive list of supporting datasets for use in the desk-based task.
- Develop a formal set of technological requirements prior to approaching technology vendors in a tender process.
- Ownership of the data being entered in the system varies by vendor and should be considered for any proposed solution.

9 Acknowledgements

Thanks to Tim Angell, Neal Armour-Chelu, Dawn Balmer, Gerry Barnes, Lindsey Bilston, Simon Bower, Anne Casey, Jenny Chamberlin, Ben Darvill, Geoff Doggett, Jake Fiennes, Katy Froud, Tim Gatti, Ross Haddow, Adam Hinchliffe, John Hiskett, Paul Holley, Martin Horlock, Tom Hunt, Tony Leech, Helen Leith, Andy Musgrove, Emily Nobbs, David Noble, David North, Marya Parker, Stuart Paston, Simon Pickles, Lorna Shaw, Edward Stocker, Tim Strudwick , Sarah Taigel, Heidi Thompson, Tim Venes, Doreen Wells, Steve Whitbread, David White, Russell Wilson for completing questionnaires and interviews on volunteer communities (Chapter 4; includes affiliations).

Thank to Katherine Boughey (Bat Conservation Trust), Katie Cruickshanks (Butterfly Conservation) and Hayley New (Plantlife) for providing membership and volunteering information for the spatial analysis of volunteer communities (Chapter 4).

Thanks to Alison Johnston and Mark Miller for provided statistical input in planning the power analysis (Chapter 5).

Thanks to Matt Aitkenhead, Crona Hodges, Jamie Williams and Karen Wright for completing the vendor questionnaire (Chapter 6).

Thanks to the Steering Group members for advice and comments on an earlier draft: Lynn Heeley (Project Manager), Richard Alexander, Colin Chapman, Paul Gilbertson, Paul Robinson, Douglas Scott.

10 References

Aitkenhead, M. J., & Aalders, I.H. (2011). Automating land cover mapping of Scotland using expert system and knowledge integration methods. *Remote Sensing of Environment* **115**, 1285–95.

Arsanjani, J., Helbich, M., Balillah, M., Hagenauer, J. & Zipf, A. (2013). Toward mapping land-use patterns from volunteered geographic information. *International Journal of Geographical Information Science* **27**, 2264–78.

Balmer, D.E., Gillings, S., Caffrey, B.J., Swann, R.L., Downie, I.S. & Fuller R.J. (2013). *Bird Atlas 2007–11: the Breeding and Wintering Birds of Britain and Ireland.* BTO, Thetford.

Clery, D. (2011). Galaxy Zoo Volunteers Share Pain and Glory of Research. *Science* **333**, 173–75.

Comber, A., Van der Velde, M., Perger, C. & Foody, G. (2013). Using control data to determine the reliability of volunteered geographic information about land cover. *International Journal of Applied Earth Observation and Geoinformation* **23**, 37–48.

Comber, A., See, L., Fritz, S. (2014). The Impact of contributor Confidence, Expertise and Distance on the Crowdsourced Land Cover Data Quality. GI_Forum 2014 - Geospatial Innovation for Society.

Crick, H.Q.P. (1992). A bird-habitat coding system for use in Britain and Ireland incorporating aspects of land-management and human activity. *Bird Study* **39**, 1-12.

De Leeuw, J., Said, M., Ortegah, L., Nagda, S., Georgiadou, Y. & DeBlois, M. (2011). An assessment of the accuracy of volunteered road map production in Western Kenya. *Remote Sensing* **3**, 247–256.

Fonte, C., Bastin, L., See, L., Foody, G. Estima, J., (2015). Good practice guidelines for assessing VGI data quality. Proceedings of AGILE 2015, 9-12 June 2015, Lisbon, Portugal.

Foody, G. M. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of Environment* **80**, 185–201.

Foody, G.M., See, L., Fritz, S., Van der Velde, M., Perger, C., Schill, C. & Boyd, D.S. (2013). Assessing the accuracy of Volunteered Geographic Information arising from multiple contributors to an internet based collaborative project: Accuracy of VGI. *Transactions in GIS* **17**, 847–860.

Fritz, S., McCallum, I., Schill, C., Perger, C., Grillmayer, r., Achard, F., Kax-ner, F. & Obersteinder, M. (2009). Geo-Wiki.Org: The use of crowd-sourcing to improve global land cover. *Remote Sensing* **1**, 345-354.

Goodchild, M.F. & Li, L. (2012). Assuring the quality of volunteered geographic information. *Spatial Statistics* **1**, 110–120.

Khatib, F., Dimaio, F., Foldit Contenders Group, Foldit Void Crushers Group, Cooper, S., Kazmierczyk, M., Gilski, M., Krywda, S., Zabranka, H., Pichova, I., Thomspon, J., Popovic, Z., Jaskolski, M., Baker, D. (2011). Crystal structure of a monomeric retroviral protease solved by protein folding game players. *Nature Structural & Molecular Biology* **18**, 1175-1177.

Medcalf, K.A, Parker, J.A., Turton, N. & Bell, G. (2014). *Making Earth Observation Work for UK Biodiversity Conservation Phase 2*. JNCC Report 495B, JNCC Peterborough.

Perger, C., Fritz, S., See, L., Schill, C., Van der Velde, M., McCallum, I., Obersteiner, M. (2012) A campaign to collect volunteered geographic Information on land cover and human impact. In: Jekel T, Car A, Strobl J, Griesebner G, editors. GI_Forum 2012: *Geovisualisation, Society and Learning*. Berlin/Offenbach: Herbert Wichmann Verlag. pp83–91.

Rutzinger, M., Rottensteiner, F. & Pfeifer, N. (2009). A comparison of evaluation techniques for building extraction from airborne laser scanning. I*EEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **2**, 11–20.

Schurmann, N. (2009). The new Brave NewWorld: geography, GIS, and the emergence of ubiquitous mapping and data. *Environment and Planning D: Society and Space* **27**, 571-580.

See, L. , Comber, A., Salk, C., Fritz, S., Van der Velde, M., Perger, C., Schill, C., McCallum, I., Kraxner, F. Obersteiner, M. (2013). Comparing the quality of crowdsourced data contributed by expert and non-experts. *PLoS ONE* **8**, e69958.

Appendices

The following Appendices are supplied as standalone documents.

Appendix A - Distribution maps of habitats in Norfolk Living Map

Distribution maps to aid in consideration of the abundance and aggregation of different habitat classes.

Appendix B - Grid square maps

Examples of 500-m grid squares in different landscapes to aid consideration of the practicality of field recording in grid squares.

Appendix C - Volunteer and Stakeholder interviews

Full details of all responses to interviews.

Appendix D - Power analysis simulation results

Full set of graphs showing results for different sampling scenarios.

Appendix E - Technological Review Vendor Questionnaires

Full details of all responses to vendor questionnaires

Appendix F - Technology solutions matrix

Scores assigned to different requirements for different vendors.

Appendix G - R scripts

R scripts used for Chapter 2 (analysis of the Living Map data) and Chapter 5 (power analysis). These are available at: https://github.com/BritishTrustForOrnithology/eodip5_earth_obs_power_analysis