JULY 2016

Options for a New Integrated Natural Resource Monitoring Framework for Wales

Project Document

Briefing note: The Potential of Citizen Science Data for Monitoring in Wales

Welsh Government Contract No. C147/2010/2011

Agreed Additional Work Requirement Dated 8th March 2016

NERC CEH Project: NEC05945

Canolfan Ecoleg a Hydroleg
Centre for Ecology & Hydrology
Natural Environment Research Council

Prifysgol Bangor University

BTO Looking out for birds

WRc

JNCC Joint Nature Conservation Committee

How to cite this document: Siriwardena, G., Cheffings, C., Smart, S., et al. (2016) **Options for a New Integrated Natural Resource Monitoring Framework for Wales; Phase 1, Project Document - Briefing note: The Potential of Citizen Science Data for Monitoring in Wales;** Report to Welsh Government (Contract reference: C147/2010/11; Agreed Additional Work Requirement Dated 8th March 2016). NERC/Centre for Ecology & Hydrology (NERC CEH Project: NEC05945)

Options for a New Integrated Natural Resources Monitoring Framework for Wales

Project Document - Briefing note:

The Potential of Citizen Science Data for Monitoring in Wales

Gavin Siriwardena (BTO)
Simon Smart (CEH)
Chris Cheffings (JNCC)

With input and assistance from David Robinson (CEH Bangor), Dylan Lloyd, Dave Allen, Liz Howe (Natural Resources Wales), Jeremy Biggs (Freshwater Habitats Trust), Katie Metcalfe (Environment Systems), Colin Chapman (Welsh Government), Kate Lewthwaite (Woodland Trust), Andrew Davey (WCr plc), Steve Lucas (Bat Conservation Trust).

Acknowledgements We would like to thank Dawn Balmer, Andy Musgrove, Rachel Taylor, Kelvin Jones, David Noble, Nick Moran (BTO), Cath Shellswell and Hayley New (Plantlife), Oliver Pescott (CEH Wallingford), Dylan Williams (NRW) and Tara Froggatt (DCWW) for kindly providing additional information and input to this note.

July 2016

Intentionally blank

Briefing note:
The Potential of Citizen Science Data for Monitoring in Wales

Definition/purpose

- Here, we define citizen science as biodiversity and soils recording conducted by volunteer or unpaid observers.
- We recognize five relevant types: (i) structured national surveys designed to collect particular evidence (e.g. for biodiversity, schemes such as the BTO/JNCC/RSPB Breeding Bird Survey), (ii) the collation of records collected independently of any survey structure (e.g. the harvesting of biological records for monitoring), (iii) local monitoring projects conceived and conducted by amateur naturalists, (iv) recording activity designed primarily to encourage public engagement, (v) "blind" sample collection (recordings or physical samples) for professional analysis. These types are most developed for the biodiversity evidence category, but are applicable more widely.
- (i), (ii), (iv) and (v) are typically managed professionally and therefore incur costs for administration, database and project management as well as analytical and reporting effort.
- In Wales, Local Environmental Record Centres often function as a clearing house for the professional storage and dissemination of data from all four categories of activity, although some societies and local schemes work directly with the Biological Records Centre.
- The Wales Biodiversity Partnership (http://www.biodiversitywales.org.uk/Citizen-Science) also acts as a delivery mechanism, publicising and facilitating use of citizen science in Wales, recognising its importance in public engagement and contributions to official statistics.
- We do not consider (iv) any further because data collection may lack rigour or be prone to spatial and temporal biases, but the databases produced can be used as data of type (ii). Examples include OPAL surveys, BioBlitzes, the New Year Plant Hunt, the RSPB Big Garden Birdwatch, RSPB Starling Survey and GWCT Farm Bird Count. Note that type (iv) is increasingly considered to cover all "citizen science" in parts of the global conservation movement, e.g. Kobori et al. (2016): " we define citizen science as engaging the public in a scientific project, a definition that is gaining general acceptance among citizen science researchers and practitioners".
- It should be noted that the process of synthesising disparate datasets and the diverse components of different schemes, including type (iv) onto useable databases and then dealing analytically with variation in quality and quantity can incur substantial downstream costs.
- Table 1 below allows quick comparison of different types of citizen science survey data.

Advantages/Disadvantages of citizen science approaches to monitoring

Advantages:

- Cheap with regard to survey effort (including identification training) compared to professional survey.
- Tapping into an otherwise unused skill set among the general public, although the potential for increasing both skills of individuals and the numbers of skilled people is limited.
- Potentially education of the public (although this should be secondary because learning on the job conflicts with quality control), including supporting longer term policy outcomes, increasing

awareness of environmental issues, and influencing behaviours that contribute to mitigating or adapting to environmental change.
- Can increasingly be tied to remote-sensed data to provide habitat context and thus to reduce the need for complex sampling protocols; can also potentially validate or ground-truth remote-sensed data.
- Quantity of information collected compensates for lack of quality control for individual records.
- New technologies increasingly allow more sophisticated data collection by untrained observers; together with automated verification, this may help to attract new cohorts of observers.
- Highly committed surveyors may fund their own survey/sampling equipment, although there will be limits, of course.
- Fully structured surveys, especially those with randomized site selection, provide data equivalent to those from professional monitoring when protocols are well-designed (although survey intensity is unlikely to be as high).
- Local survey intensity restrictions mean that citizen science data are typically most valuable for inference at large geographical scales (regional or national), although this value increases with sampling structure and geographical biases can be a problem.
- Coverage of greener locations within easy reach of people tends to be good, meaning that data can be representative of lowland farmland and suburban areas, but uplands and city centres are more challenging. However small schemes and societies can help to drive delivery of *ad hoc* data for many specialised taxa.
- Current WG policy restricts professional survey data collection to areas for which express access permissions have been obtained, but volunteer observers are free to survey from all rights-of-way, so can potentially cover locations that professionals could not.
- Type (iii) surveys are entirely volunteer-driven, so require no *necessary* central funding or control, but collation, storage and dissemination of data, e.g. via NBN or LERC, incurs costs.

<u>Disadvantages:</u>
- Considerable investment in recruitment, reporting-back and engagement activities for surveyors, including continual replacement of older observers (although overall the costs are always likely to be far lower than those associated with purely professional surveys of the same size).
- Data collected still need to be input, checked, processed and analysed professionally, or by volunteer organizers, such as LERCs (although online and automated systems are increasingly performing some of these functions). Resources for these activities need to be provided for any scheme.
- Not all taxa or monitoring activities can be assumed to be equally attractive to volunteers, even maximizing uptake. Taxa well covered by volunteer schemes are strongly aligned to taxonomic groups of higher public interest; capacity to increase coverage of others will always be low and more structured surveys are likely to interest fewer people. Type (v) surveys require a different approach to make them attractive if they are to be conducted by volunteers.
- Sampling from precise locations (and repeat visits) cannot be guaranteed, so targeting specifically for local data applications may not be possible.
- Repeat monitoring in the absence of change may cause observer fatigue, limiting long-term consistency, although this should be minimized by sophisticated sampling designs, as well as well-chosen survey frequency and a balance between spreading effort spatially and temporally.

- Data that can be collected are restricted, in particular the collection of contextual data is often resisted by surveyors with a strong interest in a particular group, for example.
- Complicated protocols both turn off some observers and may not be followed closely by those who do take part, although sophisticated designs can be successful and there are examples of approaches with high uptake and high data quality among BTO surveys, for example.
- Restrictions on recording effort (available spare time limits survey duration/sampling effort) limit protocols to less complex or low intensity designs.
- Avoidance of unpopular locations, even with randomized site selection (e.g. avoiding inner cities, remote places or areas of low biodiversity interest, or the need to secure land owner's permission or to follow biosecurity protocols to gain access), can cause geographical bias and volunteer drop-out, and limits representativeness at large scales. Unpopular locations are also likely to include habitats perceived as boring, such as arable fields, conifer plantations and improved grasslands, despite the fact that such areas are often a focus for agri-environment scheme funding. These biases can be measured, but not necessarily corrected for.
- Responsive recording may cause bias (e.g. collection of samples only when problems are perceived or submission of records of common species only when rare species are also present).
- Quality control is limited (record verification and training or certification of volunteers may be impossible or costly, although this is highly variable between taxa); it may also be impossible to control how closely observers follow protocols. The required level of quality may, however, differ with the purpose of the scheme.
- Particularly with respect to type (ii) and (iii) data, ownership of the raw data often lies with multiple individual recorders and societies, which complicates how permissions for uses of the data are obtained and managed, and may create logistical difficulties with reporting.
- Particularly considering type (ii) data, people sufficiently interested to put effort into sampling may be intrinsically biased, e.g. self-reporting of environmental impacts by farmers or conservationists highlighting policy priorities, although some such problems may involve perception and credibility, rather than being real.
- Wider cultural change could mean that current interest in monitoring is not reflected in future generations, although it is also possible that interest will rise and education could play a role.
- Individual observers survey fewer locations than professionals, introducing more observer variation per unit sample size, and potentially requiring greater investment in equipment.
- Type (iii) surveys lack central direction and, potentially, rigour, while data supply to central monitoring processes cannot be guaranteed, may not be free and may leave significant processing work to be done professionally.

I. Applications and current state of development

Long-running volunteer surveys (type (i)) in the UK underpin much of the monitoring of biodiversity in the UK, particularly with respect to birds and butterflies. Historically, the norm was for structured, detailed surveys of user-selected locations (e.g. the Butterfly Monitoring Scheme, Common Birds Census and Rothamsted moth recording), but there has been an increasing drive to replace or to augment these schemes with surveys based on random site selection, to avoid geographical or habitat biases and to produce results representative of national populations. Thus, the Breeding Bird Survey has reported on bird (and some mammal) populations since 1994. Newer schemes are now

aiming to do the same for butterflies (Wider Countryside Butterfly Survey http://butterfly-conservation.org/113/wider-countryside-butterfly-survey.html) and plants (National Plant Monitoring Scheme http://www.npms.org.uk/). The National Bat Monitoring Programme (NBMP) has annual structured transects as well as targeted recording of roost sites and hibernacula, while new technologies are also facilitating the development of new, standardized and structured approaches. A further new scheme is also likely to be launched in the near future for pollinators. All of these surveys are particularly strong in revealing large-scale variation in abundance of widespread species; they are less useful for rarer species, because their habitats are unlikely to be covered by random sampling. Statistical approaches to make the most of these data have been in development for more than 20 years and established procedures now exist for index production, separation of long-term trend from inter-annual fluctuation, dealing with spatial and temporal autocorrelation, spatial generalization ("gap-filling"), estimation of precision and the investigation of causes of change. New, more refined approaches continue to be developed, however, as this is a live field of research. An important area of research is in determining how many plots and locations is enough. This requires first specifying the monitoring question requiring evidence and then determining, often using simulation methods, whether a particular number of records with particular spatial coverage is sufficient for answering the question. Questions differ with policy-driven needs for evidence. Monitoring of simple temporal change requires fewer data points than analysis to diagnose causes of change. Notwithstanding the analytical power of newer techniques such as Bayesian modelling, more records surveyed with minimum spatial and temporal bias are likely to give more reliable answers.

In recent years, there has been a new focus on extracting information from unstructured biodiversity recording, especially for taxa and regions where sufficient amateur survey effort to support structured surveys has traditionally been difficult to find. There have also been various drives to increase the collection of such data, often using online and smartphone/tablet technology to provide user-friendly interfaces for data collection and basic verification, with the added value (from the user's perspective) of central data collation and storage. Unstructured data from record-harvesting notably have the potential to provide critical information about scarce taxa that are not surveyed efficiently by randomly located sample sites. New developments to enhance the value of unstructured information have taken two principal directions, first, to encourage the recording of spatial and temporal recording effort and, second, developing new analytical approaches. Each approach attempts to overcome the effects of biases in recording effort and thus to allow the extraction of reliable information on (especially) temporal change. This is also a live research area, but there will always be a hard limit to data quality for target taxa and areas where recording effort is negligible. It is in these areas where there may be no choice but to deploy fully funded professional surveyors.

All of the above monitoring effort considers populations or distributions of different taxa, but other properties of key populations, such as timing (phenology) and wildlife health, are also covered with citizen science approaches, generally via collated type (ii) data. F or example, the long-running Nature's Calendar run by the Woodland Trust (WT), which involves volunteers across Britain recording the timing of natural events such as date of first Ash leaf or Swallow arrival in the vicinity of their home. The trends captured by these data have been used in earnest by academic researchers investigating the advance of the growing season in recent years, for example (see Box ? for other citizen science schemes run by WT).

Natural resource monitoring outside the broad biodiversity area has had little citizen science involvement until recently, but growing numbers of approaches now exist. One general technique,

which is being applied to cryptic biodiversity and to other targets, such as soils, is the solicitation of "blind" sample collection by citizens for image or chemical analysis centrally, generally in professional laboratories. Sample collection can be entirely unstructured/opportunistic, or via soliciting records from specific locations. Such methods are in use for soils (mySoil, http://www.bgs.ac.uk/mySoil), amphibians (using eDNA to identify the presence of newts in ponds, http://freshwaterhabitats.org.uk/projects/pondnet/survey-options/edna-for-great-crested-newts/), bats (static detectors recording ultrasonics for sound spectrogram analysis centrally, http://www.batsurvey.org) and the Predatory Bird Monitoring Scheme for surveillance for pesticide effects from analyses of livers and eggs (via corpses submitted to a central laboratory, http://pbms.ceh.ac.uk/). In addition, proven relationships between groupings of organisms, identifiable by non-expert volunteers, and environmental conditions mean that records can be used as proxy indicators of pollution (http://www.apis.ac.uk/nitrogen-lichen-field-manual).

Policy-relevant questions about biodiversity change that have been answered effectively using citizen science datasets include the following:

What has changed over time? A domain of interest is specified such as the UK or Wales and the answer is provided by an analysis of trends in a species or a group of species. Examples include the C4, C5, C6 and C8 UK Biodiversity indicators published by JNCC (http://jncc.defra.gov.uk/page-4233).

What has caused the observed change? This question requires that changes in the biota of interest can be either divided into impacted versus unimpacted groups or arranged along gradients with respect to likely driving variables. Structured survey designs with relatively even yet randomised coverage of the areas in which drivers have operated make it much easier to address this question. Spatial and temporal biases either associated with unstructured data or with variable uptake of structured schemes increase the chances that gradients of various drivers will be unevenly sampled leading to unreliable hypothesis tests about the causes of change.

In Wales, survey coverage has historically been sparser than in England, chiefly reflecting the low density of human observers in the upland areas that make up much of the country. For example, BBS coverage has been lower than ideal and volunteer recruitment drives and mentoring have been undertaken in recent years with the aim of sustainably increasing long-term coverage. This has been very successful with the number of BBS squares in Wales increasing from 245 in 2010 to 330 in 2015. Similarly, the new National Plant Monitoring Scheme (NPMS) began volunteer recruitment and recording in earnest in 2015. Uptake in Wales was low, however, but 2015 was considered a pilot year and a major push has been organised by Plantlife to promote greater engagement, hopefully resulting in more vegetation plots per habitat in subsequent years (see Appendix 1).

Current national reporting and evidence gathering for Wales relies on a wide range of available structured survey data (see Appendix 1), showing the value that NRW place on current and past citizen science data in Wales. There may be additional value in the harvesting of unstructured records to contribute to future recording in Wales, both by increasing biological recording effort and carrying out more central collation and processing. Unstructured survey data should be exploited to its fullest potential because full deployment of structured survey effort may be hampered by low uptake. For example, this may limit the potential of the new National Plant Monitoring Scheme in Wales. Exploiting unstructured survey data requires that the quantity and quality of records are maximised. This means working in partnership with Wales LERC, whose data holdings may often exceed the numbers available via the NBN, and with centres of analytical expertise such as CEH and BTO who can help deal with spatial and temporal variation in recording effort.

Citizen science approaches are well-suited, within limits and with careful use, to a range of surveillance, monitoring and evaluation applications with respect to natural resources (including monitoring towards national or regional environmental targets), but they are not suitable for local-scale regulatory applications (e.g. compliance of farms to statutory regulations). Variation in the ability of differing citizen schemes to provide robust long-term, large-scale evidence for monitoring can be understood in terms of a tension between policy-focused end-users of data and those more focussed on the benefits to scheme participants.[1]

a. Costs

All monitoring schemes are different, so costs are variable. For planned schemes, requirements for data inputting, sample processing and data analysis are variable as well. While online data capture saves costs, the systems required can also be expensive to develop, and have ongoing hosting, updating, user web support and maintenance costs. Simple field survey costs are far lower than those of monitoring programmes using professional observers, but volunteer management (site allocation, dealing with queries, data checking and reporting back) and system development costs are likely to negate much of any savings for the first few years. Record harvesting approaches are less costly, especially if they can make use of pre-existing systems for data recording (which may exist primarily for the personal interest of the recorders), although analyses will certainly be more complex and this may introduce higher associated costs.

b. Key Issues for Implementation (including costs)

- Citizen science schemes of types (i), (ii), (iv) and (v) require professional infrastructure for design, administration, data storage, analysis and reporting, both nationally and to volunteers. This incurs significant costs, but field cost per unit effort is far lower than that of professional surveys.
- Costs vary between monitoring schemes. For structured schemes, requirements for data inputting, sample processing and data analysis also vary. Online data capture systems can also be expensive to develop, and have ongoing hosting, updating, user web support and maintenance costs. Record harvesting approaches are less costly, especially if they use pre-existing systems for data recording, although analyses will be more complex and this may increase associated costs.
- Skills required of contributors are highly variable, from following detailed protocols and specialist identification of difficult groups to simple deployment/collection of sampling equipment, with no skills required at all.
- Citizen science is critically linked to volunteer motivations; what works for a one-off survey might not work for long-term monitoring. Schemes may also compete with one another for a limited pool of volunteer effort.
- The design of schemes has to take motivation into account and with the growing range of schemes and information portals, participant expectations are rising, for example regarding ease of use of the website, timeliness of feedback, etc.. This has implications for costs.
- In Wales, Local Environmental Record Centres often function as a clearing house for the professional storage and dissemination of data from all types of citizen science, although some societies and local schemes work independently or with the Biological Records Centre.

[1] Pocock, MJO et al (2015). Developing and enhancing biodiversity monitoring programmes: a collaborative assessment of priorities. J.Appl.Ecol. 52, 686-695.

- Certain NGOs already organize structured schemes or calls for unstructured data, typically supported by government or agency funding.
- Citizen science field effort and scheme organization do not have national monitoring for Wales as their primary driver, so any implementation of systems to extract monitoring evidence needs to take account of, and may be subservient to, potentially competing priorities at the scheme level.
- New analytical initiatives, e.g. using scheme data to measure management effects, as in Box 1, require funding support, and potentially extra funds for methodological development.
- New monitoring schemes need long-term support for design, pilot projects and infrastructure.
- The UK Environmental Observation Framework (UKEOF) has produced valuable reports on the practicalities of extracting evidence from citizen science projects, considering motivational factors and their interactions with successful project design, and introducing a new tool for evaluating the costs and benefits of new schemes. Aligning scheme design with (potential) participant motivations is critical, as is buy-in to effective evaluation from stakeholders. The cost tool is freely downloadable, in MS Excel format, and aims to take account of monetized and non-monetized factors via cost-benefit, cost-effectiveness and return-on-investment analyses.

c. Current and potential future contributions of Citizen Science biodiversity recording to NRW evidence needs

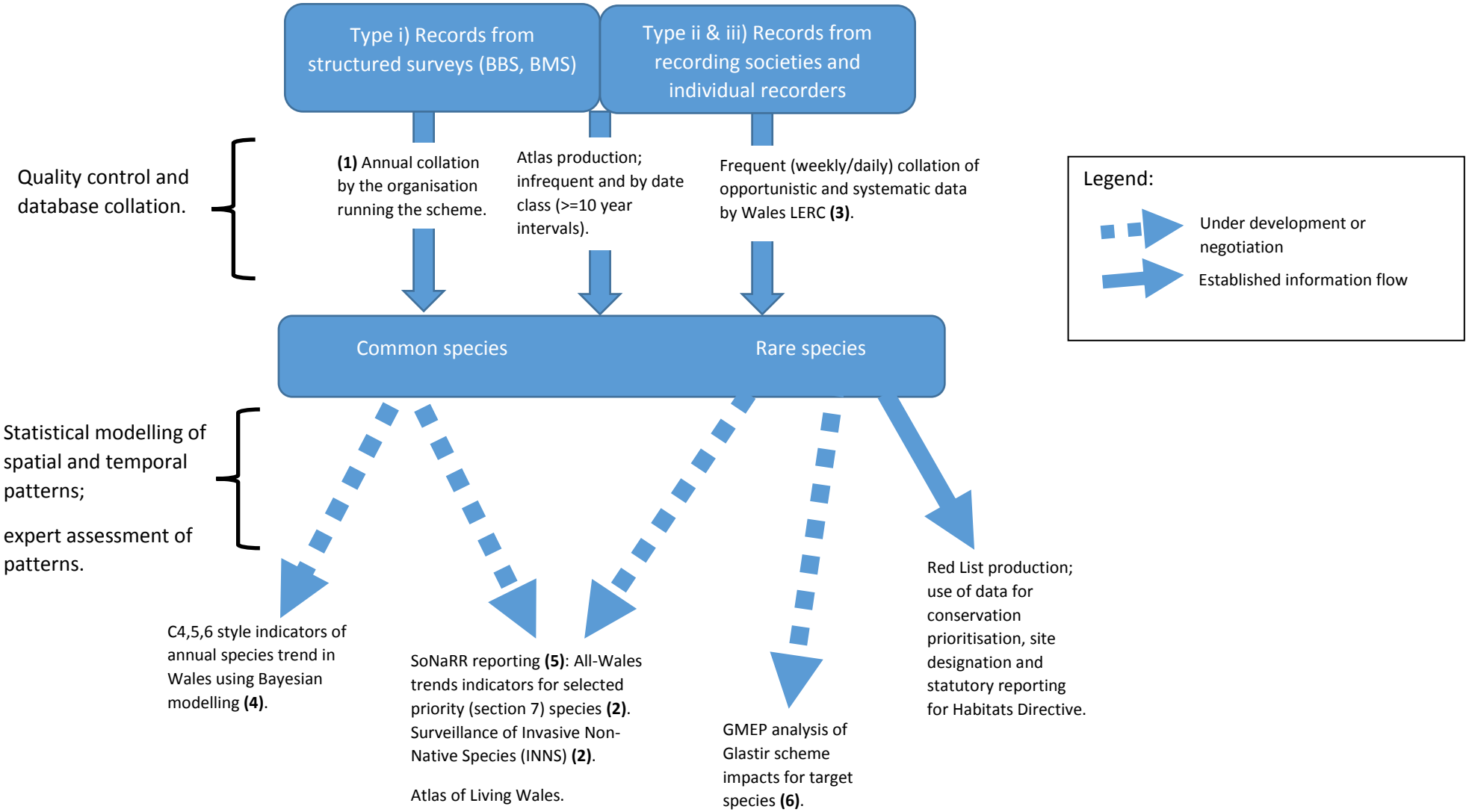
A wide range of citizen science species-recording schemes are relied upon for evidence by NRW (see Appendix 1). Figure 1 summarises the current and potentially future flows of information from citizen science schemes devoted to species observations, from recording through to their use as evidence either internally for operational decision-making or to fulfil reporting obligations for example for SoNaRR and Habitats Directive. Records are included from designed surveys (type i), record collations (type ii) and local voluntary recording projects (type iii). Type iv and v records are also utilised in national evidence when collated into a spatial database and so become referable to type ii.

The collation and analysis of records accumulated by designed survey schemes such as BBS and BMS represents a model for established, state-of-the-art translation of records into evidence in the form of GB spatial maps and temporal trends **(1)**. However variable recording effort in Wales, especially with regard to rare species, raises ongoing issues regarding the acquisition, identity and possible level of aggregation of species records into useful evidence of status and trend. Maximising the numbers of records available for the range of section 7 species and INNS is an obvious priority **(2)**. The Wales Local Environmental Record Centres therefore have a central role as a hub for collating and disseminating up-to-date records and thereby removing from the end-user the sometimes complex task of establishing trust and good working relationships with a diverse range of recorders and societies **(3)**. Moderation of the records by species' experts prior to acceptance for analysis is an additional process handled in-house by the structured national schemes yet also needed for a range of other species particularly section 7 taxa.

Translating species records into reliable spatial maps and temporal trends depends fundamentally on the numbers of records available and the biases in their distribution. Modern statistical modelling methods can help adjust for differences in recorder effort. Applying these techniques and carrying out the research needed to explore how they perform under varying levels of bias and noise, is currently performed at an analytical hub such as the Biological Records Centre at CEH **(4).**

A number of information flows are currently the subject of ongoing development and discussion and will be supported by the Future Options consultation process. This includes the role of citizen science sources in providing possibly new evidence required to discharge the duty to prepare and publish SoNaRR in the new Environment Act (2016) **(5)**. Work is also ongoing to determine the scope of a priority species trend indicator for Wales **(4)**. In addition, ongoing discussion seeks to make available the highest quantity and quality of records for taxa explicitly targeted in Glastir option bundles. These data are essential for testing the hypothesis that species abundance has benefited from Glastir uptake **(6)**.

Figure 1. Citizen science species-recording schemes and NRW evidence requirements



d. Key Issues for Interpretation

- Designed schemes have data collection designed to lead to monitoring outputs and indicators, and so a natural and clear link between data and evidence; this also facilitates their use for other analytical purposes, such as evaluating environmental impacts.
- Unstructured data sources may need considerable scoping work to determine the level of reliable inference supported, and then sophisticated analyses to extract real evidence from the raw data.
- Citizen science sources feature inevitable geographical reporting bias, less with structured data but not zero, because there is bias in the uptake of randomly allocated sites for surveys in practice away from those that are harder to access. This leads to situations like a "black hole" in coverage in mid-Wales. However, given *some* coverage of difficult regions, sampling biases can be corrected, especially within structured surveys.
- New analytical approaches using Bayesian models may increase the utility of unstructured data, for example incorporating prior knowledge of the extent of species' distributions or their trends in England to inform the production of trends for Wales. Recent CEH research (Isaac et al. 2014, Methods. Ecol. Evol.) has compared a range of approaches to developing temporal trends from these data and made recommendations, but tests of approaches for use in the assessment of environmental impacts have yet to be conducted.
- The principal benefit from citizen science data is that large quantities of information can be collected or collated at a low cost, such that possible problems with the quality of individual records are swamped by a more reliable majority.
- Citizen science is best suited to low-intensity, low-effort surveys that require only limited skills (or skills that are common in the population). Hence, the data produced are best for large-scale surveillance intended to detect widespread changes and are less useful for local-scale, short-term impacts of management or environmental change. However, sampling biases can have significant effects on the representativeness of the results.
- There is a very wide range of forms of data, from randomized, structured surveys indistinguishable from professionally-collected data sets to entirely opportunistic and biased sets of records. The options for interpretation of these datasets are similarly broad. *They cannot be considered as a single form of information.* If structured surveys are feasible, they are preferred, but harvesting unstructured records may be the only option. In either case, the extent to which the desired, reliable inference can be gleaned from the best citizen science approach available needs to be assessed objectively before a final decision on survey approaches is taken.

e. Experience to date

- Almost all citizen science data collection and use in monitoring to date has involved biodiversity, and the majority of that has involved more charismatic, diurnal animal groups (although particular amateur experts have contributed hugely in respect of specific other groups). Thus, evidence of the value of such data is heavily biased towards population trends and, to some extent, evaluation of environmental impacts, on birds and butterflies. The use of these data for national reporting is well-established.
- NRW place a high value on current and past citizen science data in Wales and use the information in national reporting and evidence gathering, including monitoring trends in biodiversity via the C4, C5, C6 and C8 UK Biodiversity indicators published by JNCC (http://jncc.defra.gov.uk/page-

4233). The same data are then of critical value for analyses determining the causes of change, for which changes in the biota of interest can be either divided into impacted versus unimpacted groups or arranged along gradients with respect to likely driving variables. This is much easier with structured survey designs, as applied to investigations of impacts of historical agri-environment impacts on birds in Wales (see Box 1).

- The Woodland Trust runs successful citizen science projects with trained (upskilled) volunteers collecting valuable, if unstructured and, therefore, probably geographically biased, data on tree health, phenology and the locations of ancient trees (which inform planning issues) (Appendix Y).
- The mySoil smartphone app provides novel capacity for reporting soil condition, but only via the solicitation of unstructured data. Hence, biases in the representativeness of the data are possible and similar data for the UK Soil Observatory from Wales collected by a self-selected sample of farmers show opposite trends to a well-structured, professional survey, because samples have been collected disproportionately when problems with soils were perceived. See Box 2.
- Another new initiative is the use of volunteer effort to ground-truth Earth observation data, which is being scoped by JNCC, with a view to monitoring of environmental change. A pilot project led by Environment Systems has trialled such an approach in Warwickshire (see Box 3).
- In Wales, survey coverage has historically been sparser than in England, chiefly reflecting the low density of human observers in the upland areas that make up much of the country. This is important to note in assessments of the potential of citizen science approaches based on experience elsewhere.
- Low-intensity survey data are valuable when used to assess large-scale effects/trends, but less useful at small scales, where more intensive monitoring, and, probably, professional surveyors, are required.

BOX 1. Case study: application of Breeding Bird Survey data to evaluating Tir Gofal

The BTO/JNCC/RSPB Breeding Bird Survey (BBS) is a volunteer survey conducted annually in a random sample of 1km squares across Wales using standardized methods. As part of GMEP, species' counts over time were analysed (following Baker et al. 2012, J. Appl. Ecol.) to measure effects of Tir Gofal (TG) management on bird population changes. Options providing Grassland habitat, Arable winter seed, Arable invertebrates, Woodland creation & stock exclusion, Heathland, Scrub management and hedgerow management were considered. The citizen science and management data were combined with Land Cover Map background habitat information (from Earth observation) to remove habitat biases from the analyses. Positive associations with TG options were much more common than negative ones, particularly for woodland and hedgerow management, followed by arable seed provision and scrub management. The evidence suggests, therefore, that this management under TG has had positive net effects on Welsh bird populations, but that the other option types have not been so effective.

Weaknesses with this study include the inability to assess rarer species and options because of small sample sizes, so the results may not reflect high conservation priorities. The balance of effects across species for several option types suggests that TG has been broadly beneficial; for other options, either small sample size effects (e.g. heathland) or failure to address limiting factors (e.g. arable invertebrate options) probably underlie the limited effects.

BOX 2. Case study: A crowd-sourced database of soil condition data.

Soils data in the form of unstructured records, are collected through both the mySoil iphone and android apps (http://www.bgs.ac.uk/mySoil/#ad-image-ad-image-0) and also through the UK Soil Observatory Online (www.UKSO.org). The digital apps and portal are able to collect any written and photographic data; guides are provided for basic texture, pH and photos. The existing soil data tools were designed to raise awareness of soils but have great potential for added value data collection. The next upgrade of mySoil will include Welsh language support, whilst the next version of the UKSO will include crowd-sourcing and verification of landcover map data.

A survey of users shows, 40% are gardeners, 30% are farmers and 30% are in research. The team are currently trying to understand how users apply these tools. Respondents say that mySoil increases knowledge about soils and increases the quality of work they do. The following comment about mySoil shows the utility of these platforms for small business, “I find this really useful in my role as an agricultural crops advisor, it gives me a good idea of the predominant soil type in any particular location”.

BOX 3. Combining citizen science and Earth Observation: opportunistic ground-truthing of habitat maps in England.

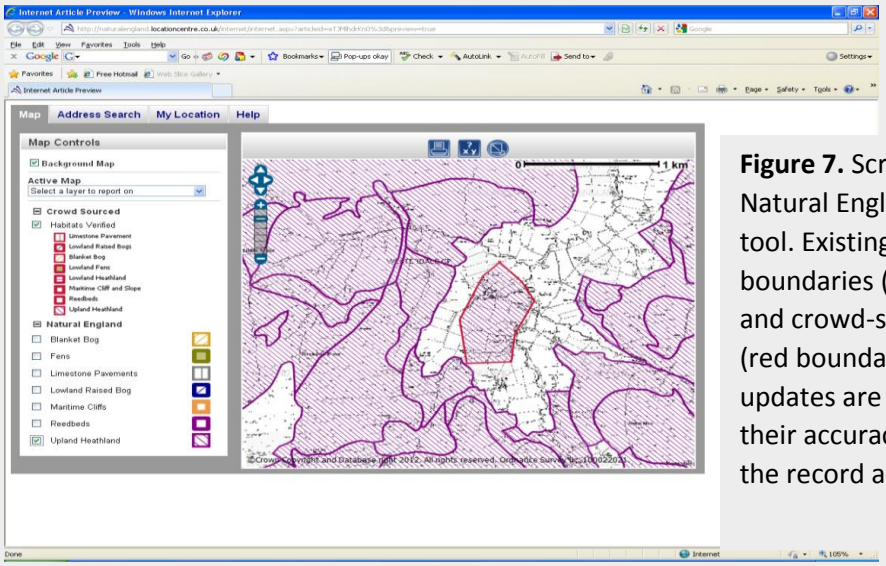


Figure 7. Screenshots from the Natural England pilot interactive tool. Existing Priority Habitat boundaries (a) (purple boundary) and crowd-sourced updates (a) (red boundary). Submitted updates are then moderated for their accuracy and plausibility and the record accepted or rejected.

Maps showing the extent of Priority Habitats are available from the NBN and MAGIC portals. These provide access to finely resolved data but the resulting coverage is based on compilations of survey datasets of varying age and reliability. Natural England have piloted an on-line tool to allow updating of the habitat map by professionals or volunteers.

A similar project has been carried out by Environment Systems and Warwickshire County Council where a remotely sensed habitat map of the county can be updated by volunteers on the ground (Medcalf et al 2015).

BOX 4 : A Citizen Science campaign to measure water quality in Oxfordshire.

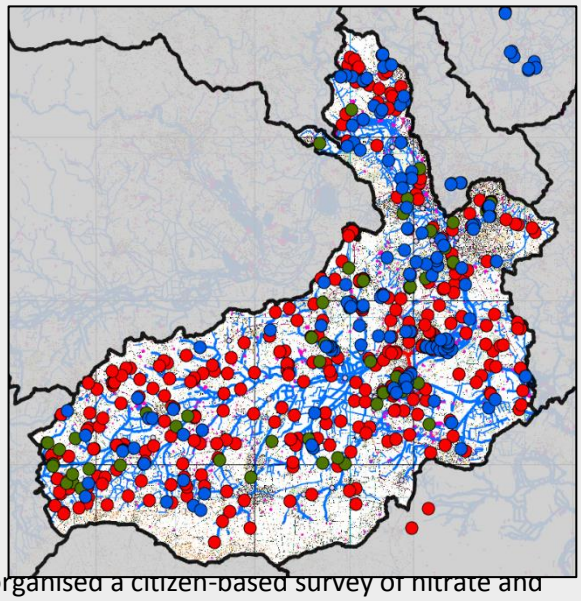


Figure 1. Water quality in the River Ock, Oxfordshire catchment. Undertaken using citizen science methods, this is the first survey of all waterbody types across a whole river catchment.

In April 2010, Freshwater Habitats Trust organised a citizen-based survey of nitrate and phosphate levels on 570 sites (ponds, lakes, streams, rivers, ditches, fens) in the catchment of the R. Ock, Oxfordshire as part of the Clean Water for Wildlife project. This was slightly more than 1 waterbody / km^2 in this 470 km^2 catchment. Most sites are not currently monitored.

Rapid water test kits were successfully able to separate 'clean' water (i.e. those at 'High' status under WFD) from more polluted waters.

Nearly a third of sites were 'clean', predominantly ponds and lakes, with some streams and ditches. Most running waters experienced substantial nitrate or phosphate pollution.

Table 1. **Current assessment of technology.** Cell colours denote the current state of citizen science for monitoring in Wales with respect to the examples listed: Green – happening, possible, doable, achievable; Amber – likely, probable, achievable after a bit more progress; Red – not likely, hard to achieve, challenging; Grey – unknown/unknowable. These categories have been used to be consistent with the other "technologies" considered in this report. For citizen science and biodiversity monitoring, there are several constraints on how information can effectively be fitted into the categories available. First, there are many data owners and the authors have not been able, in the time available, to consult with them all over plans and data quality so as to be able to make reliable, supportable judgements about possible inference now and into the future. Second, an important issue in assessing quality or possible inference is what the monitoring target actually is. For many of these datasets (and from the point of view of the data owners), "success" might be the collection of data, nationally representative sampling, the capacity to detect national trends or the capacity to detect effects of environmental/management change. Moreover, rather different forms of data are required to answer questions involving variation in abundance or presence (distribution). ? And what data are needed – presence or abundance? Third, unstructured biodiversity recording may be useful for selected locations (reserves, gardens, villages, farms, parishes or whatever), but poor/unrepresentative at larger scales, so the definition of the exact scale involved for "local" monitoring could be critical. Fourth, the use of "investigative" implies a study structure with controls as well (usually) as application at the local scale, so it creates rather a narrow category and it has been disregarded in populating the "local" column. Finally, the definition for amber is rather positive and overlaps with that for green. For the data sources considered here, a category for "uncertain, may work with further development of volunteer networks or statistical processing but further trialling is required" would have been useful.

Example	Local (site or grid square)	National surveillance (Wales)	2-5 years	>5 years	Comments
Plants	NPMS	NPMS			Some local inference may be supportable for limited/biased locations; national representativeness will depend upon uptake and taxonomic resolution.
Plants	BSBI recording (TPP)	BSBI recording (Atlases and Local Change)			Recording is underway for Atlas 2020. The Threatened Plants Project (TPP) and Local Change (LC) surveys could be repeated and extended in Wales.
Birds	Schemes designed for national inference	BBS used for national population reporting and tests of management effects; BirdTrack can be	As current – significant increases in volunteer effort are unlikely	As current – significant increases in volunteer effort are unlikely	Some local inference may be supportable for limited/biased locations

		developed further for rarer species			
Butterflies	WCBS designed for national inference, BMS suitable locally for target habitats	WCBS and BMS are used to produce national indices	As current – significant increases in volunteer effort are unlikely	As current – significant increases in volunteer effort are unlikely	Some local inference may be supportable for limited/biased locations
Bees, Wasps & Aculeates		BWARS			Annual or longer term trends probably achievable for more common taxa using modern occupancy modelling with recorder effort correction but representativeness depends upon taxonomic coverage, rarity and the influence of spatial and temporal biases.
Other invertebrates	Unstructured NBN data only – may be suitable for selected locations	Unstructured NBN data only – some national monitoring may be possible using new statistical approaches	As current		
Bats	Schemes designed for national or regional inference, but maternity roost and hibernation site surveys inform at the site level where they are conducted	National Bat Monitoring Programme was designed for the UK, but the data support statistically reliable trends for Wales for the species monitored	Scope for further development and standardization of monitoring methods and to tap into a new volunteer base	As 2-5 years	Extent of additional potential volunteer effort unknown
Other mammals	Possibly some useful records for certain species in NBN	Some species covered by BBS; extent of additional potential			

		volunteer effort unknown			
Aquatic vertebrates	Only presence/ absence tractable, limited value at local scales	New FHT volunteer surveys in development, with species-level eDNA. Power and sample biases uncertain.	As current; scope for structured sampling being investigated. Volunteer interest uncertain.		
Water quality	New FHT volunteer surveys in development. Power and sample biases uncertain.	New FHT volunteer surveys in development. Power and sample biases uncertain.	As current; scope for structured sampling being investigated. Volunteer interest uncertain.		
Habitat mapping	e.g. Case study from Warwickshire when combined with EO	Ground-truthing of CEH Landcover has just been added to mySoil/UKSO, uptake unknown; JNCC are scoping broader potential	Ground-truthing of EO habitat data is being trialled and may be effective but scope and biases unknown		
Soil	Being trialled in mySoil and other apps e.g. the Crap- app. NRM have published their data but bias identified so would need work and only does farmers (so no forests or coast etc)	Could explore possibility for sending in samples from selected squares but untested to date	Unknown	Unknown	
Health and disease: animal pesticide effects	Predatory bird scheme and collection of otter carcasses for rodenticides, etc. provides unstructured	Predatory bird scheme records are probably biased with respect to human population	As current	As current	

	data with unreliable local record density	density but are believed to be reliable			
Pollution recording	EA solicits reactive data on incidents; CEH, in collaboration with a range of partners, have produced an on-line app linking lichen morphotypes to nitrogen deposition levels[2]	Various apps on pollution e.g CEH have various. Not sure if anyone has tried to make a map out of it or interpreted results?			

[2] http://www.apis.ac.uk/nitrogen-lichen-field-manual

f. Data Informatics

- Important ongoing surveys in Wales can be tapped into to inform about changes and impacts/management effects on natural resources. There may be significant added value to be gained from applying Bayesian models to unstructured data, but this requires further development.
- Citizen science data are best-suited to large-scale applications in which local-scale site turnover is not important and detailed local inference is not required, subject to survey uptake being suitably representative (see Figure 1 for an illustration of current citizen science coverage in Wales). Professional surveys are required otherwise. Structured volunteer surveys are the best citizen science option, where practical, followed by unstructured data collation with as much secondary recording of recording effort as possible.

Figure 1. Distribution citizen science survey effort in Wales: (a) randomly selected 1km BBS squares taken up by volunteers, with the depth of grey colour indicating the number of years of survey coverage to date (up to 21); (b) numbers of unstructured BirdTrack observations, across all species, submitted up to 2016, summarized by 10km square.

- Data harvesting and the solicitation of unstructured records clearly enable the collation of large quantities of data. While these data could generate valuable monitoring evidence, this is not necessarily the case, because geographical and other biases are likely to limit reliable inference significantly. Data analysis can account for some issues here, but is not a panacea. Intrinsic data quality, typically involving the degree of structure underlying data collection, will be central in determining the evidence value of available data and should be evaluated in critically in determining the extent to which a given data source meets evidence needs.
- Further engagement-focused initiatives are likely to arise because soliciting data is an effective way of increasing interest and educating the public. However, the value of the data collected from these processes for monitoring and delivering evidence is limited and it is important that this is recognized at policy level.
- Exploiting unstructured records requires that the quantity and quality of data are maximised. This means working in partnership with Wales LERC, whose data holdings may often exceed the

numbers available via the NBN, and with centres of analytical expertise such as CEH and BTO who can help deal with spatial and temporal variation in recording effort.

- Data ownership is a sensitive issue with citizen science data. First, more structured schemes require investment to support design, volunteer management and data analysis, leading to organizations having significant intellectual property. Second, many individual volunteers commit great time and effort to data collection and, sometimes, monitoring design. In both cases, the (part-)ownership of these data by the organizations or people concerned needs to be respected, recognized and rewarded adequately to maintain effort and data quality in the long term. How to manage this in the long term and across all ultimate uses of data collected via citizen science schemes is a significant challenge.

II. Unique Selling Point

The application of volunteer/amateur recording effort to monitoring priorities can be a cost-effective alternative to large-scale professional surveys, potentially with added value in community engagement.

III. Future potential

Public interest underlies citizen science effort and is secure as a resource in the long term. Funding support for many structured schemes is also probably fairly secure because the data products are valued at policy level. Although this is subject to the maintenance of government and agency support, the ongoing importance of citizen science to UK-wide environmental surveillance is recognized and promoted by the UK Environmental Observation Framework[3]. Increasingly sophisticated remote-sensed data are likely to add value to citizen science by providing more and better habitat context information, while online recording platforms provide increasingly sophisticated data capture and checking facilities. While public interest and regional biases in activity are always likely to limit monitoring in practice, more sophisticated statistical models are likely to improve the utility of unstructured data.

Threats to continued volunteer effort include loss of skills as older naturalists are not replaced because younger generations may lack the level of engagement required to conduct surveys for some taxa. However this may be more of a threat to trying to increase coverage rather than to maintenance of existing levels and may be offset by increasing numbers of retirees joining volunteer surveyor communities, although it is possible that over-reliance on demographic change and

[3] The UKEOF aims to develop a holistic picture of the observation needs of the UK, to share knowledge and information, to understand the use of observation data and tools for knowledge transfer, to enable funding mechanisms for long-term observations and to build a strong community to share data and expertise.
The Citizen Science working group provides a forum for member organisations to share good practice and discuss future needs and plans, helping partners make best use of different monitoring approaches.
Resources: http://www.ukeof.org.uk/resources/citizen-science-resources .
Two project reports have also recently been published on “Understanding Motivations for Participating in CS”; and on “CS and Environmental Monitoring: Opportunities, Costs and Benefits”. These can be found on the main page of the EOF web site.

recruitment does not constitute a sustainable strategy for citizen science. Rising transport costs are likely to be a growing problem for volunteers to contribute high levels of survey effort.

A cost-effective approach to more representative or complete coverage than volunteer effort allows may be to augment existing volunteer effort with professional effort, for example in remote areas. However, effective survey designs for volunteers typically require low field effort per survey so as to be more tractable in people's spare time. Such protocols, without modification, are unlikely to deliver cost-effective use of professional surveyors' time. Paying travel expenses for surveyors may be a further route to increase coverage, assuming that costs are limiting for them, but this could cause organizational problems within surveys. There may also be potential to develop novel data collection systems making use of volunteer effort, supported by new technologies, from online systems to laboratory DNA analyses, for example in the freshwater environment. However, levels of volunteer interest are always likely to be limiting and cannot be taken for granted. Costs of developing new schemes may also be significant. All developments of new monitoring should be piloted to ensure that the data collected can provide appropriate evidence and revised or abandoned if this is not the case.

A general issue with citizen science is clarity of aims and objectives. There is broad scope to develop new data collection protocols and creative approaches for interaction with and encouragement of the public. Schemes can be focused on monitoring, engagement and/or education, so there is potential to achieve multiple objectives at once. However, the optimal scheme designs for these objectives may be very different, varying in factors such as observer skills required, duration, replication, minimum sample size and geographical locations. Particularly from the perspective of national monitoring requirements, it is critical that new scheme designs and the exploitation of existing data match data quality to potential inference appropriately, such that evidence needs can be met reliably.

a. Next steps for development as a monitoring tool

- Citizen science approaches are well-suited, within limits and with careful use, to a range of surveillance, monitoring and evaluation applications with respect to natural resources (including monitoring towards national or regional environmental targets), but they are not suitable for local-scale regulatory applications (e.g. compliance of farms to statutory regulations). Structured surveys such as BBS, NBMP, WCBS and NPMS are valuable for future monitoring and can contribute to the evaluation of management impacts at large scales. With further methodological development, the same may be true for some unstructured datasets. However, if detection of management effects means making existing simple methods more complex and difficult to implement this will probably result in reduced engagement from volunteers.
- Future development of citizen science for monitoring can take four directions: new surveys, exploitation of further unstructured sources, more analyses of existing data and integration of citizen science and professional effort.
- It is possible that additional structured surveys could succeed, but observer interest will be a strong restriction. The fate of new pollinator, plant and Earth Observation ground-truthing survey initiatives will be instructive.
- Freshwaters represent a significant monitoring gap, in Wales as elsewhere. The Freshwater Habitats Trust has identified significant potential for developing semi-structured monitoring of

freshwaters in Wales. These involve ongoing development of the PondNet and Clean Water for Wildlife programmes so that citizen surveys, especially for large scale water pollution monitoring and biodiversity monitoring using eDNA, can now provide data which are not available with other approaches. In particular, citizen science approaches are seen as a valuable, cost-effective approach to covering the large numbers of small water bodies and low-order streams that have high environmental importance. Future work will focus on developing the sample-collection-and-testing approach further for water quality and biodiversity applications, considering the key methodological, statistical and practical application questions. As with other data sources lacking a formal sampling structure, work is needed to reveal the evidence value of the data likely to be collected, i.e. the representativeness of the sampling and the sensitivity to reveal (a) changes over time and (b) effects of management or environmental change. See Appendix Y to the Main Report for further details.

- There may be additional value in the harvesting of unstructured records to contribute to future recording in Wales, both by actively soliciting increased biological recording effort and by carrying out more central collation and processing. Unstructured data should be exploited to its fullest potential where structured surveys would not be feasible due to low uptake, such as may limit the potential of the new National Plant Monitoring Scheme in Wales.
- The most cost-effective approach to future monitoring with representative or complete coverage may be to combine volunteer recording with professional effort, for example in remote areas.
- Effective survey designs for volunteers typically require low field effort per survey so as to be more tractable in people's spare time. Such protocols, without modification, are unlikely to maximize the data collected during professional surveyors' time, so the latter should certainly only be applied to structured surveys. More sophisticated combinations of effort are likely to be more cost-effective, for example using tiered sampling approaches, with volunteer data informing about gross patterns and professional supplementation providing complementary detail, as used in the combination of BBS and professional bird surveys in GMEP.
- Paying travel expenses for surveyors may be a further route to increase coverage, assuming that costs are limiting for them, but this could cause organizational problems within surveys.

Table 1: Types of citizen science scheme and their attributes, with relevance to Wales[4].

Forms of scheme/data collection	**(Stratified) random surveys (type i)**	**Surveys with observer selected locations (type i)**	**Atlases (type i)**	**Biological records (type ii)**	**List data (type ii)**	**Volunteer run projects (type iii)**	**Sample collection (type v)**
Example data sets	BBS, WCBS, NPMS	BMS, CBC, ObservaTree	Bird Atlas 2007-11, Butterfly Atlas	NBN, Ancient Tree Inventory, Nature's Calendar, Earthworm Watch	BirdTrack, WildWalks	Welsh Chough monitoring	mySoil, Newt eDNA, PondNet, PBMS, Norfolk Bat Survey.
Primary purpose	Tracking temporal change	Tracking temporal change	Mapping (change in) distributions	Personal interest of recorders; some with more direction	Personal interest of recorders	Personal interest of recorders, tracking temporal change	Mostly mapping and surveillance for gross change
Species/other target coverage in practice	Widespread species	Widespread species and some habitat specialists	All species	Rarer or specific species (and phenology)	All species (and phenology)	Specific rare species	Specific targets of sampling methods
Primary spatial unit	Standard survey areas (often 1km squares)	Patches of target habitats	Grid squares (size taxon-dependent)	Simple locations of records	Locations of records	Locations of records or patches of habitats	Specific sampling sites or simple locations of records
Representativeness/bias	Representative of habitats in sampling design	Typically broadly representative of target habitats	Complete coverage (at large scales, at least as an aim)	Biased according to recording effort	Biased according to recording effort	Typically complete for restricted target areas	Probably biased according to recording effort, but variable
Value at national scale	High	Moderate to high, depending on coverage	High	Can be high if coverage is high or unbiased with respect to distributions	Moderate	Typically low unless whole populations are measured in a single local area	Can be high if coverage is high or unbiased
Value at local scale	Low	Low to high, depending on field method	Low	Can be high if there is local standardization in recording	Low	High	Low to high, depending on field method
Stability of funding support	High (subject to government and agency support)	High (subject to government and agency support)	Moderate	High	Moderate	Low	Moderate (often start-up funds only; expensive)
Effort control	Strong	Strong	Strong (large scales) to moderate	Weak	Moderate	Moderate	Strong

[4] See also http://ecsa.citizen-science.net/community/map for Citizen Science capacity-building across Europe.

Forms of scheme/data collection	**(Stratified) random surveys (type i)**	**Surveys with observer selected locations (type i)**	**Atlases (type i)**	**Biological records (type ii)**	**List data (type ii)**	**Volunteer run projects (type iii)**	**Sample collection (type v)**
Quality control/verification	On data entry, automated (moderate)	On data entry, automated (moderate)	Strong	Weak to strong, depending on taxon and location	On data entry, automated (moderate)	Strong	Strong
Counterfactual/zero records	Included automatically	Included within target habitats	Included automatically	Not considered	Included but may be biased	Not included but reliably inferred	Included automatically
Sensitivity to change: value for measuring changes over space and time	High	High	Low (long periods between repeat surveys)	Low to moderate (mostly presence data with variation not controlled)	Low (more often presence data than counts, variation not controlled)	High	Variable/unknown (many methods yet to be evaluated)
Suitability for measuring management or environmental effects	High (for long-term, widespread effects subject to uptake and dispersion)	Moderate (depends upon coincidence of target habitat and drivers of interest)	Low (only via space for time substitution)	Low (high uncontrolled variation and difficult to associate with spatial data)	Low (high uncontrolled variation and difficult to associate with spatial data)	Moderate (limited spatial extent)	Variable/unknown (many methods yet to be evaluated)
Spatial coverage (extent)	Wales (but with habitat biases due to uptake).	Wales (in target habitats)	Wales	Wales (with habitat biases)	Wales (with habitat biases)	Restricted areas	Restricted areas, but growing
Standardization of spatial sampling unit	Fully standardized	Loose only	Fully standardized	None	None	Loose or none	Loose to fully standardized
Sampling design	Randomized (but with observer-selected details)	Observer selected with restrictions	Complete	None	None	Observer selected with restrictions	None to observer selected from a restricted set
Standardization of sampling method	Fully standardized	Fully standardized	Partly standardized	None or full (in more directed schemes)	Some (effort recording)	Partly standardized	Fully standardized
Frequency	Annual	Annual	Periodic (less than decadal)	Flexible dependent upon summarization (including sub-annual)	Flexible dependent upon summarization (including sub-annual)	Annual	Variable (at least annual)
Examples of use for evidence	AES evaluation in England and Wales	Identification of farmland bird decline due to agricultural change	Identification of bird range expansion due to climate change	Records summarized for some State of Nature recording and UK Biodiversity indicators.	Identification of phenological change in bird migration		Identification of ponds with Great Crested Newts to inform development

The potential of citizen science data for monitoring in Wales

FURTHER READING

1. Key datasets derived from volunteer-based schemes that currently contribute to the evidence needs of Natural Resources Wales (courtesy Dr Liz Howe, Head of Species Team, NRW Bangor).

- **Bird data**- all BTO datasets and trends analyses plus red listing and birds of conservation concern.
- **Bat data**- Bat Conservation Trust and all rare bats recording projects.
- **Mammal data**- mammal society datasets.
- **Dormouse**- National Dormouse Monitoring Program run by the PTES (Peoples Trust for Endangered Species) http://surveydata.ptes.org/dormousemonitoring/.
- **Otter**- UK otter survey (https://naturalresources.wales/media/4590/osw-5-english-24-06-2015.pdf)
- **Herpetofauna** – National Amphibian and Reptile Recording Scheme (http://www.narrs.org.uk/) and the rare reptile and amphibian database held by ARC.
- **Butterflies**- the UKBMS and in Wales the Butterfly Conservation, Marsh Fritillary and other rare species surveys.
- **Moths**- National Moth Recording Scheme also run by Butterfly Conservation (http://www.mothscount.org/text/27/national_moth_recording_scheme.html).
- **Other animals** – also rely on evidence and data from a range of other recording schemes and societies.
- **Inverts** – The Wales invertebrate recorder database has over 0.5 million records and will be going onto the NBN soon.
- **Plants** – BSBI and Plantlife recording schemes.
- **Non-vascular plants** - various recording schemes (primarily the British Bryological Society and the British Lichen Society) and the red lists that go with them.

2. Examples of newer citizen science schemes and derived indicators relevant to Wales

Soils data and the mySoil app

Soils data in the form of unstructured records, are collected through both the mySoil iphone and android apps (http://www.bgs.ac.uk/mySoil/#ad-image-ad-image-0) and also through the UK Soil Observatory online (www.UKSO.org); all platforms provide information, with more than 50,000 users and 4000 records crowdsourced from across the globe. The digital apps and portal are able to collect any written and photographic data; guides are provided for basic texture, pH and photos. The existing soil data tools were designed to raise awareness of soils but have great potential for added value data collection. For instance, farmer soil analysis records could be collected, peat depth mappers could record across Wales and games could be developed around data collection; there is

no reason to prevent the tools being used in professional survey. The tools would need some upgrading for these applications, for example mySoil has no offline capability to record information, and this could be added and is important for Wales. The next upgrade of mySoil will include Welsh language support, whilst the next version of the UKSO will include crowd-sourcing and verification of landcover map data.

Atlas of Living Wales

This project is ongoing and involves harnessing the functionality of the recently completed Atlas of Living Australia. The Atlas of Living Wales will be built using an open source biodiversity data infrastructure and is intended to deliver on the fourth Strategic aim of the NBN Strategy, The Atlas of Living Wales will offer the ability to create a Welsh view, bringing together species and habitat data.

To quote from the NBN website "*The Atlas of Living Wales will offer the ability to create a Welsh view, bringing together species and habitat data and offering functionality including the ability to view and upload photographs, search for biological data by predefined areas, by postcode or by polygon search tools, find organisations working in a particular area (geographic or taxonomic) and create alerts for species records. Additionally, the Atlas of Living Wales will provide bilingual functionality, offering users the option to switch between Welsh and English language pages – functionality which has not previously been available via the NBN Gateway.*

This project is part of a work programme to build Atlas infrastructure for England, Wales, Scotland and Northern Ireland and use the same scalable platform to create an Atlas of Living UK. Each of these five atlases would have the same functionality and same basic design and be supported by a single database but offers users a more focussed, local view of the national data holdings."

New functionality is also likely to include alerts to expert recorders allowing them to moderate newly uploaded records. This should favour an increasing quantity of useable records but not at the expense of quality. It is also worth noting that this facility already exists through iRecord (supported by national recording schemes and societies and available at http://www.brc.ac.uk/irecord/) and BirdTrack (http://www.bto.org/volunteer-surveys/birdtrack/taking-part/birdtrack-apps).

UKBMS indicator for Wales

BRC with Butterfly Conservation are due to produce an indicator for Wales that is consistent with the UK JNCC C6 indicator of butterfly trends (http://jncc.defra.gov.uk/page-4236) but based on improved Bayesian occupancy modelling. This will be updated annually under the current UKBMS contract which runs for another year till end of financial year 16/17.

Most of the work on the indicator has been done, but the information (sample sizes for species, trends, indicator plots) needs consolidating into a form agreed by NRW. Once finalised the indicator will be placed in the public domain.

Plantlife section 42 species monitoring activities

Plantlife have devoted considerable effort toward accumulating high resolution records for section 42 higher plants, grassland fungi and lichens in Wales. However they do not currently run structured monitoring of these taxa but may be able to fund future activities pending the outcome of funding bids (Cath Shellswell pers.comm.).

Developments in bird monitoring in Wales

Multiple structured and semi-structured volunteer schemes led by the BTO contribute to annual or periodic bird abundance monitoring in Wales, and some then provide data for Wales-specific indicators (Appendix 1). In addition, professional monitoring and periodic, targeted volunteer surveys record various rare and priority species under the SCARABBS programme (e.g. raptors and twite), or led by volunteers (e.g. chough); the SCARABBS surveys are led by NGOs. Various schemes also monitor bird demography, but these are not strictly relevant here. The BTO/JNCC/RSPB Breeding Bird Survey (BBS) is the principal annual scheme for monitoring terrestrial breeding birds, and a recent drive to increase survey engagement via peer-to-peer mentoring has seen a 35% increase in coverage. The Waterways Breeding Bird Survey (WBBS) is a sister scheme for linear waterways that has specific relevance for Wales because of the importance of the rivers for specialists such as dipper and grey wagtail; it is currently supported by BTO and reported along with the main BBS. Wintering waterbirds are covered annually on estuaries and a sample of freshwaters by the BTO/RSPB/JNCC Wetland Bird Survey (WeBS); complete coverage is achieved of coastal sites, whereas it is more patchy inland and increased uptake would be valuable. Non-Estuarine Waterbird Surveys are conducted approximately every nine years to record wintering birds away from estuarine sites and are particularly important for the rocky shoreline in Wales. Bird Atlas 2007-11 was the latest in a series of periodic (c. 20-yearly) Britain and Ireland projects measuring distribution and relative abundance of all species in winter and summer, which included complete coverage of Wales at the 10km square level. In addition to these general schemes, the BTO organizes specific, periodic monitoring of particular target groups, with notable examples for Wales including the Wales Chat Survey from 2012 (for whinchat, stonechat and wheatear) and the Peregrine Survey from 2014. Finally, the BirdTrack system (partners include the Welsh Ornithological Society) is an online recording portal to capture casual bird records, including recording of complete lists, which provide a measure of effort and thus an element of structure to the data. Methods for the analysis of these data are still in development, but they have the potential to fill information gaps for scarce and localized species all year round, as well as providing information about the timing of migration.

National Plant Monitoring Scheme

For more widespread plants and CSM indicators linked to semi-natural habitats the nascent NPMS scheme may have a prominent role to play in future monitoring. The great advantage of the scheme is its low cost. It is managed by Plantlife, BSBI and CEH and among volunteer-based schemes uniquely addresses the challenge of annually monitoring common plants within fixed vegetation quadrats that can be explicitly grouped by habitat type. The scheme has been running for one year across the UK. Uptake in terms of number of quadrats recorded in 2015 in Wales is indicated in Table A1 alongside quadrat numbers for the two professionally funded vegetation monitoring schemes that have covered Wales; GMEP and Countryside Survey.

The NPMS scheme clearly has the potential to be an important contributor to future monitoring. Currently uptake is however relatively low in Wales and Plantlife are involved in ongoing attempts to increase participation.

A number of issues pertain to the use of NPMS in addressing possible questions about monitoring change in vegetation and common species and identifying the drivers of those changes in the future.

Issues:
1. Differences in plot sizes between schemes (NPMS versus CS and GMEP versus NRW datasets). The requirement here is to measure diversity and other variables of interest in such a way that they are corrected for differences in area censused. This only applies if there is a real need to amalgamate datasets but in some cases this may be the case.

2. How many plots are actually required? Could it be that despite low current uptake of NPMS it may in fact provide enough quadrats to answer relevant questions? This depends on the question; attribution of changes over time to multiple driver gradients requires adequate randomised, replicated and crossed samples along each hypothesised gradient. Modern Bayesian modelling methods can readily deal with missing data but the critical point is that Bayesian imputation does not guarantee lack of bias in inference. It therefore does not correct for biased sampling across the domain of interest. In essence there is a limit on the extent to which sampling variation can be compensated by sophisticated modelling. If the question is about identifying trends over time then it is still the case that biased sampling will produce a trend estimate representative of some areas but not of others.

3. Roughly a third of the NPMS plots in 2015 are recorded at wildflower level and another third at CSM indicator level. Joint analysis of NPMS and other quadrat datasets could be carried out by reducing the taxonomic coverage of all datasets to an equivalent level; for example only selecting wildflowers or CSM vascular plants from CS, GMEP and NRW quadrats. Work is underway to determine the cost versus benefit of this approach with respect to the use of NPMS plots in England as a counterfactual for the current HLS re-survey.

4. It may be important that bryophyte (moss & liverwort) cover is not recorded in NPMS plots. Total bryophyte could presumably be easily added to the guidance for NPMS in Wales. In the western oceanic seaboard of Britain and in upland habitats, bryophytes provide important ecosystem functions including moderating run-off, N fixation, substrate protection, C storage and habitat for other species of animals plants. They are therefore likely to contribute to ecosystem resilience.

5. It would be useful to explore the effect of any bias in NPMS locations toward freely accessible land for which land-owner permission did not need to be sought. Again, information as to whether volunteers sought permission or not could presumably be recorded in future years and retrospectively gathered for 2015 plots.

6. By design the NPMS preferentially targets 1km squares rich in semi-natural habitats. This is because its purpose is to measure change in the abundance of species typical of these more threatened habitats across the UK. Square selection was achieved using an objective weighting of all UK 1km squares by land-cover diversity. Therefore since all 1km squares in Wales have a weighting the coverage of NPMS plus GMEP and the extent to which they are severally and jointly representative of Wales could be readily quantified.

7. NPMS targets semi-natural habitats. Improved land and conifer plantation are therefore deliberately avoided by NPMS yet these habitats attract a range of Glastir interventions and so NPMS may not be optimal in these habitats. Conversely NPMS plots ought to help detection of impacts in semi-natural habitats. Further consultation is required to determine how far NPMS could be adapted to help with detecting Glastir impacts. At a recent workshop discussing future monitoring of HLS and Countryside Stewardship options in England it was thought that asking NPMS volunteers to additionally stratify by in or out of option land would foist prohibitively complex protocols on them and risk drop-out.

8. Options for further exploring the contribution of NPMS to monitoring in Wales could include adoption of existing GMEP squares, or at least some of the GMEP quadrats within squares, by NPMS volunteers. The emphasis would presumably be on ‘interesting’ and ‘accessible’ squares near to volunteers’ homes.

9. Analysis of NPMS plots and GMEP plots in accidentally coincident squares could also shed light on differences in species and habitat coverage by the two methods.

Table A1. Total numbers of fixed vegetation quadrats in Wales currently available for analysis from three monitoring programs, the volunteer-based National Plant Monitoring Scheme, which went live in 2015, and the professionally funded Glastir Monitoring and Evaluation Program and Countryside Survey of Great Britain. Note that neither of these two latter schemes have secure funding for any future re-recording. Quadrats are grouped by the habitat or feature they sample. Note that finer divisions of plots to section 42 habitat is possible. NPMS quadrats include those from all three levels of recorder effort; wildflower, indicator and inventory (see http://www.npms.org.uk for more information). Numbers of NPMS plots were correct at 18th March 2016 (courtesy Oli Pescott, CEH Wallingford).

Table A1.

Quadrat types and broad habitats from CS/GMEP	**NPMS habitat types**	**NPMS**	**GMEP**			**CS**		
		2015	**2013**	**2014**	**2015**	**1990**	**1998**	**2007**
A plots on cultivated field margins	Arable margins	2	7	14	8	0	11	17
Bog	Bog and wet heath	9	104	90	63	9	41	58
Broadleaved woodland + linear H and D plots	Broadleaved woodland, hedges and scrub	58	106+ 388	143+ 515	83+ 362	50+ 52	74+ 300	159+ 608
All coastal broad habitats	Coast	21	17	11	16	18	43	44
Streamside plots	Freshwater	12	183	226	174	209	257	458
Heath	Heathland	13	54	54	66	18	56	101
Neutral grassland	Lowland grassland	55	125	135	107	53	88	152
Fen, Marsh & Swamp	Marsh and fen	14	93	121	64	41	74	96
Inland rock	Rock outcrops, cliffs and scree	8	5	5	5	1	9	17
Acid grassland	Upland grassland	20	86	128	136	60	138	209

Woodland Trust led or partnered citizen science projects.

Project	Background	What does the project do?	How is the data used?	Partners	Who are the key WT contacts?
Nature's Calendar www.naturescalendar.org.uk	The longest running citizen science project at WT, started 2000	Thousands of untasked volunteers across the UK collect information about the timing of natural events where they live eg date of first swallow of spring, first tinting leaf of autumn. A sister project called Track a Tree is run by our PhD student based at University of Edinburgh	The project has a huge database (modern and historic records) and is used by academics and government to show how natural timings are changing as a result of climate change.	WT is lead partner. Supported by Centre for Ecology and Hydrology.	**Kate Lewthwaite** is project manager. **Judith Garforth** is project administrator
Ancient Tree Inventory http://www.ancient-tree-hunt.org.uk/	Began as a five year, WT-led, HLF funded project in 2006 as the 'Ancient Tree Hunt'. Name changed to reflect the value of the data now held.	Untasked volunteers seek and record ancient , veteran and notable trees; an estimated half of all these trees in the UK are now on the project database. Tasked volunteer verifiers visit and check each tree eg that species correctly identified	Data used strategically to aid in conservation decisions such as the designation of Sites of Special Scientific Interest and in deciding planning applications.	Project partners of ATI include specialist charities the Tree Register of the British Isles (TROBI) and the Ancient Tree Forum.	**Jill Butler** is project manager**. David Alderman** and **Kylie Knight** provide additional support.
Observatree	4 year project, began in 2013,	Recruited and trained a network of 200 tasked	Data used to help track the impact of	Forest Research is lead partner.	**Kate Lewthwaite** leads WT activity.

www.observatree.org.uk	funded by European funder LIFE+ WT leads on volunteer management and project communications	volunteers to add capacity to tree health scientists by helping to sift pest and disease records and carry out site visits. Promotes the use of FC online tool 'Tree Alert' to encourage reporting of pests and diseases of concern.	pests and diseases in the UK and to support more general scientific research	Other partners are National Trust and FERA. DEFRA, Natural Resources Wales and APHA are supporting partners.	**Helen Jones** is volunteer officer, **Anna O'Connor** is comms officer. **Judith Garforth** provides additional support.
UK National Tree Seed Project www.kew.org/ukntsp (NB This is not a monitoring project)	Five year project launched by Millennium Seed Bank Kew in 2013	WT recruits, trains and manages skilled tasked volunteers called "seed collection champions". The collected tree seed is stored deep frozen by Kew where possible, species that cannot survive this (eg oak) are grown on straight away by Kew or FC.	The project aims to gather a genetically comprehensive collection of important UK tree seeds to aid research and conservation efforts.	Kew is the lead partner. Other partners include Forestry Commission and National Trust.	**Kay Haw** and Kylie Knight lead WT activity.

Some definitions

Tasked volunteers- Specific number of people recruited via application to the WT volunteer team. People have a task outline (a bit like a job description), training for the role and a named WT task manager. They record their volunteering hours and receive out of pocket expenses.

Untasked volunteers – A more typical model for citizen science where people volunteer as and when they wish, no limit to the number of people that can help or the time spent. No formal training, volunteer manager or task outline. Do not claim expenses since carry out tasks as part of their normal day to day living.

Future citizen science development in freshwaters: Comments from Freshwater Habitats Trust

Jeremy Biggs, 3 June 2016

We provide some brief comments and suggestions on the main themes discussed at the monitoring workshop as they relate to freshwater ecosystems.

Summary

1. Citizen science: A citizen science approach to freshwaters offers several opportunities in Wales to complement, and extend cost-effectively, current freshwater monitoring work. Much of the infrastructure has been established thorough current and on-going FHT work developing the PondNet and Clean Water for Wildlife programmes so that citizen surveys, especially for large scale water pollution monitoring and biodiversity monitoring using eDNA, can now provide data which are not available with other approaches.

2. Freshwater monitoring: A key requirement of freshwater monitoring in Wales, as elsewhere, is the effective incorporation of small waterbodies (headwater streams, ponds, small lakes, ditches) into monitoring programmes. Although increasingly recognised as important, small waters generally remain outside of current regulatory monitoring programmes. A major hurdle to effective monitoring of small waters is cost, and citizen science programmes can provide a way round this substantial problem.

3. Possible ways forward for citizen-science based freshwater monitoring programmes: Work during 2016 has provided a proof of concept of the value of rapid water quality test kits for large scale evaluation of water quality across whole catchments (including both large and small waters). Similarly, Great Crested Newt eDNA work has also clearly demonstrated the value of this technique for volunteer monitoring of protected freshwater species.

We suggest that in 2017, using the freshwater citizen science monitoring network established in Wales by Freshwater Habitats Trust with HLF support, there is a good opportunity to explore further the potential of this approach for monitoring freshwater biodiversity and pollution problems. We recommend a larger practical trial to address key methodological, statistical and practical application questions, in three or four key catchments, as part of work to assess the potential for a longer term national citizen based freshwater monitoring network in Wales.

Background

Freshwaters in Wales include ponds, lakes, streams, rivers and ditches. As in most parts of the world, it is likely that in terms of numbers and length, small waters (ponds and small lakes; zero to second order streams; ditches) greatly outnumber the larger waters (lakes, rivers), although larger waters of course occupy a larger area.

At present, most formal monitoring of freshwaters, in terms of hydrological, chemical and ecological quality, is focussed on larger waters. Although Wales is notable for having taken a lead in work on ponds and small lakes, worldwide there is generally little monitoring of smaller waters, whether still or flowing, despite increasing recognition that small waters are important both in their own right, and through their critical influence on larger waters.

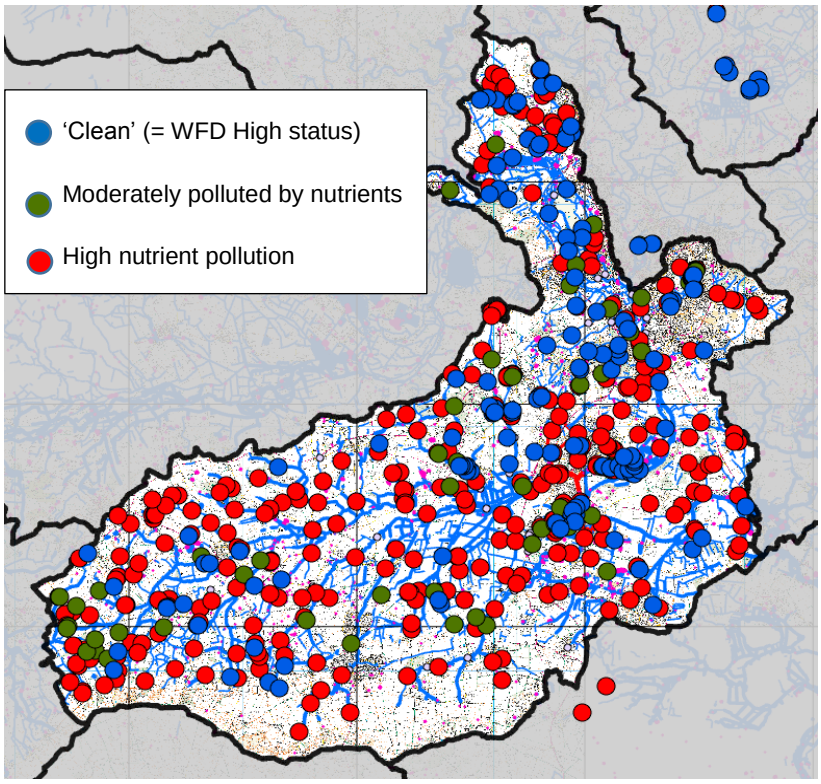
Freshwater Habitats Trust's monitoring interests

Freshwater Habitats Trust's primary interest in monitoring is to encourage, and undertake, effective monitoring of freshwater biodiversity. This also includes aspects of ecosystem service delivery, particularly water quality.

There are three areas of monitoring work in which Freshwater Habitats Trust is currently involved which could contribute to understanding of the water environment in Wales:

1. The new national, volunteer-based, pond monitoring network, PondNet, which has been established with the support of Defra, Natural England and the Heritage Lottery Fund and is currently being rolled-out to cover all of Wales and England. The programme is based around a nationally stratified sample of 1 km squares and is initially focused on assessing the quality of all ponds nationally, of Priority Ponds (a subset of all ponds) and of c.30 freshwater priority plants and animals, including the Great Crested Newt. This programme has a national database capable of managing both species and habitat data (including water quality) called WaterNet which, as well as dealing with ponds and small lakes, is designed to be extended to manage datasets from all types of freshwater, still and flowing. The investment in this programme to date is about £500,000, and has created an infrastructure that can be used by both professional and non-professional workers. A new bespoke website for WaterNet will be launched later in June.

Figure 1. Water quality in the River Ock, Oxfordshire catchment. Undertaken using citizen science methods, this is the first survey of all waterbody types across a whole river catchment.



2. A detailed technical manual for the use of rapid test kits will be published at the end of June. The use of a new generation of rapid nutrient test kits for nitrate and phosphate which can quickly and cost effectively provide an over-view of diffuse pollution at catchment and landscape scale. This could provide datasets which have not previously been available for an integrated form of water management planning covering all types of freshwaters and wetlands. The kits are usable by both

professionals and volunteers (and programmes in which both groups work together are probably going to prove most effective). An example dataset from a catchment (the River Ock, which includes Oxford), hosted as part of the Defra Catchment-based Approach by FHT, is shown in Figure 1. In April 2016, Freshwater Habitats Trust organised a citizen-based survey of nitrate and phosphate levels on 570 sites (ponds, lakes, streams, rivers, ditches, fens) in the catchment of the R. Ock, Oxfordshire as part of the Clean Water for Wildlife project. This was slightly more than 1 waterbody / km2 in this 470 km2 catchment. Most sites are not currently monitored. Rapid test kits were successfully able to separate 'clean' water (i.e. those at 'High' status under WFD) from more polluted waters. Nearly 1/3rd of sites were 'clean', predominantly ponds and lakes, with some streams and ditches. Most running waters experienced substantial nitrate or phosphate pollution. The data are now contributing to a range of practical projects. We believe this is the first example of a whole catchment, all waterbody type, analysis of water quality.

3. The exploitation of new eDNA techniques to collect datasets describing the status of waterbodies or species that are not covered by traditional monitoring programmes (e.g. most small waters, many freshwater species of conservation concern). Although there is currently considerable interest in using eDNA to replicate 'traditional' approaches (e.g. invertebrate surveys for WFD), there is also considerable potential to do things with eDNA which currently cannot be done practically by traditional methods e.g. fish surveys in lakes, large scale Great Crested Newt presence/absence surveys, large-scale surveys of fish in headwater systems. As well as having developed the Great Crested Newt eDNA programme, FHT is exploring opportunities for further single and multi-species work for monitoring freshwater biota.

Next steps

We believe that a citizen-based national or regional monitoring programme in Wales to assess the status of a representative sample of all waterbodies, focusing particularly on smaller, largely unmonitored, waters is technically and practically feasible. Such a monitoring programme could help provide a better understanding of the status of (a) water quality, particularly nutrient pollution, in a much wider cross section of freshwaters than is currently possible, (b) selected species, using eDNA techniques, cost-effectively filling gaps in existing monitoring approaches.

Examples of the kinds of practical issues such an approach could help tackle are:

1. Providing water quality data from sites of importance for freshwaters biodiversity, particularly small waters. Such work could include screening of headwater streams to identify High status sites which should be subject to 'No deterioration' objectives; monitoring lakes which are not in the existing SSSIs/SAC programme; monitoring Priority Ponds and monitoring SSSI ditch networks which are currently more or less unmonitored. As eDNA techniques develop it is likely that, in addition to water quality data, further single or multi species tests could be used by non-specialists to monitor individual freshwater species of conservation concern for which there is currently little regular monitoring. We believe there may also be benefits in assessing the potential of eDNA to detect water plants (e.g. charophytes, which are taxonomically challenging for most freshwater botanists), with the first studies of eDNA detection of water plants suggesting this may be possible.

2. Evaluating the success of measures to improve water quality such as agri-environment schemes to reduce local point source pollutions or diffuse pollution. The test kits would again be used to focus on smaller systems, rapidly screening large numbers of sites which may currently have only limited, or no, monitoring, with follow-up using standard regulatory approaches where kits provide the first

evidence of impacts (either positive or negative). This also opens up the possibility of landowners being able to see for themselves the extent of pollution, and the effects of agri-environment schemes, which has the potential to both empower land managers and encourage co-operation. Although test kits are not as accurate as laboratory analysis, they can distinguish between clean and polluted habitats, and can be used at large numbers of sites to provide a scale of survey which it is hard to fund using laboratory analysis.

3. Find clean water locations, encouraging stakeholders to more highly value these sites and ensure that small point and diffuse sources potentially affecting these areas are prioritised for remediation. At present, much of the focus of monitoring is on improving the bad rather than protecting the good. We believe that there is much to be gained by helping people focus on, and celebrating, what is already good, looking after that well, and trying to build out from it. There is also a wealth of biological evidence that shows that this is more likely to work, at least for biodiversity, and will be an essential part of improving the degraded. Thus it is clear that in many cases recovery of freshwater biodiversity depends to a large extent on recolonisation from 'good' locations.

We currently hope to continue, and extend, the freshwater monitoring programme involving citizen science in Wales which has been established in the People, Pond and Water project. Practically, our main requirement is to support FHTs Wales Officer who is co-ordinating citizen monitoring at present. We would also recommend further exploration of the pilot work undertaken with rapid test kits and eDNA during 2016 and 2017, to evaluate optimum designs for rapid test kits surveys. For example, although we have run a quite detailed programme of testing comparing the kits with lab data we still have a range of questions about the variability of the kits and their statistical power to detect change. With eDNA we would like to test the single species approach further in its ability to detect individual protected species, given the success of the great crested newt approach. Similarly, can citizen scientists collect fish or amphibian multi-species eDNA samples? We would also like to explore whether other rapid water quality test kits are useful (e.g. heavy metals, aluminium) and to further develop links between citizen survey data and practical actions to improve the environment.

Specific comments on freshwater monitoring options

1. WG could specify a vision for how freshwater monitoring activities might support a Natural Resource Management Programme including the assessment of ecosystem resilience and ecosystem service delivery, and articulate the economic, social and environmental benefits of basing management decisions on sound evidence. Through consultation, this vision could be translated into an agenda for collective action involving all stakeholders.
We suggest that through involving a citizen science element it would be possible to incorporate a wider range of both small and large waterbodies into the monitoring network, providing an excellent but practicable representation of the freshwaters of Wales.

2. NRW in partnership with Phase 2 of Future Options should undertake a comprehensive review of all freshwater monitoring activities in Wales with the goal of identifying opportunities for greater co-operation and co-ordination. Building on earlier work by the UK Environmental Observation Framework (UKEOF), the review could seek to identify information gaps, areas of duplication and overlap, and opportunities to harmonise methods and standards. Meta-data for each monitoring programme could be consolidated and made publically available to facilitate future co-ordination.
We agree with this and would include in this analysis the strengths and limitations of citizen generated datasets.

3. NRW in partnership with Phase 2 of Future Options could explore the core NRW freshwater monitoring networks to see how they can be supplemented and complemented by data and information from other sources. Working with other stakeholders, consideration could be given to the pros and cons of using models to integrate disparate data sources, and how separate lines of evidence could be combined to build a coherent, unified assessment of the state of natural resources.
We agree with this recommendation and would comment only that it should ensure effective incorporation of the wide range of new knowledge on the importance of small waters.

4. Proposed reductions to NRW's statutory monitoring networks could be subject to an impact assessment to understand the associated increase in risk. The implications could be communicated to interested parties so that they can adapt their own data gathering and reporting activities accordingly. A series of statistical and modelling approaches could be used to develop the most efficient and cost-effective approaches including a cost-benefit analysis.
The potential to use citizen networks as a 'backstop' where statutory networks must be curtailed should be assessed. It is important not to oversell the value of citizen data, but there may be situations where, as well as providing something that cannot be generated using 'traditional' statutory networks, citizen datasets may help maintain a watching brief, with less sensitive techniques, on waterbodies which would otherwise go completely unmonitored.

5. NRW in partnership with Future Options could explore the possible benefits to Wales of pooling data with environmental regulators in England, Scotland and Wales and co-operating on the development of future tools and models, including the advantages and disadvantages of modelled data. Lessons learned and new technologies being exploited by other countries could also be explored.
Freshwater Habitats Trust is happy to pool/share/exchange data. We have a policy of openly sharing all datasets.

6. WG could explore options for supporting the exchange of monitoring data between organisations in a way that encourages multifunctional data use. This could take the form of a consolidated data hub/warehouse or a de-centralised data sharing portal that allows organisations to retain ownership and control of their data.
Freshwater Habitats Trust is happy to encourage the use of our data platforms (especially WaterNet – which is intended for multi-user collaboration) to share datasets.

3. Notes (specific analytical approaches, considerations of evidence quality)

- From an analytical perspective, most structured, designed scheme citizen science survey data are fundamentally just survey data: all standard analytical approaches can be used and it is irrelevant that observers are volunteers. However, scheme-by-scheme assessment of survey site uptake and of the distribution of surveyor ability may indicate that additional controls or post-hoc weighting are required to reduce bias in estimated parameters.
- Required sample sizes and their spatial arrangement will vary with the question being asked. The challenge is to estimate the point below which the number of records and their locations lead to a) unrepresentative answers, because of bias, and b) uninformative answers because of too much uncertainty. There is also likely to be another upper threshold beyond which extra numbers of records start to add less and less statistical power. In a

voluntary scheme these extra records may not incur significant extra cost whereas in a professionally funded field campaign such an excess represents an inefficiency. These issues highlight the value of the low cost per record of citizen science schemes and of the need for careful design of professionally funded schemes where these are necessary because of low observed or expected uptake of voluntary recording.

- Summarization of data from local/point locations can improve standardization/representativeness at larger scales
- Bayesian approaches can consider bias in unstructured data, but in no way represent a panacea because information may still be lacking for some areas or periods. Bayesian imputation allows models to be constructed but does not make up for missing data. If it did then we would not need data!
- Proofs of concept and potential from larger scales or other geographical areas (e.g. UK versus Wales) may not be reliable at the Wales scale because data availability and biases are different for this subset of the full dataset concerned.
- Survey uptake per head of population in Wales for the BBS, for example, is the highest in the UK, which illustrates that simple observer density may be the limit to recruitment of volunteer effort in Wales, as opposed to levels of interest. Recruitment may also be negatively affected in some Welsh communities by a perception that survey organization is “English”.

.

NERC SCIENCE OF THE ENVIRONMENT

PEER

INVESTORS IN PEOPLE

Athena SWAN Bronze Award

BANGOR
Centre for Ecology & Hydrology
Environment Centre Wales
Deiniol Road
Bangor
Gwynedd
LL57 2UW
United Kingdom
T: +44 (0)1248 374500
F: +44 (0)1248 362133

EDINBURGH
Centre for Ecology & Hydrology
Bush Estate
Penicuik
Midlothian
EH26 0QB
United Kingdom
T: +44 (0)131 4454343
F: +44 (0)131 4453943

LANCASTER
Centre for Ecology & Hydrology
Lancaster Environment Centre
Library Avenue
Bailrigg
Lancaster
LA1 4AP
United Kingdom
T: +44 (0)1524 595800
F: +44 (0)1524 61536

WALLINGFORD - Headquarters
Centre for Ecology & Hydrology
Maclean Building
Benson Lane
Crowmarsh Gifford
Wallingford
Oxfordshire
OX10 8BB
United Kingdom
T: +44 (0)1491 838800
F: +44 (0)1491 692424

enquiries@ceh.ac.uk
www.ceh.ac.uk