



# Opinions of citizen scientists on open access to UK butterfly and moth occurrence data

Richard Fox<sup>1,2,3</sup> · Nigel A. D. Bourn<sup>1</sup> · Emily B. Dennis<sup>1,4</sup> · Richard T. Heafield<sup>1</sup> · Ilya M. D. Maclean<sup>3</sup> · Robert J. Wilson<sup>2,5</sup>

Received: 13 December 2018 / Revised: 28 June 2019 / Accepted: 24 July 2019 / Published online: 29 July 2019  
© The Author(s) 2019

## Abstract

Citizen science plays an increasingly important role in biodiversity research and conservation, enabling large volumes of data to be gathered across extensive spatial scales in a cost-effective manner. Open access increases the utility of such data, informing land-use decisions that may affect species persistence, enhancing transparency and encouraging proliferation of research applications. However, open access provision of recent, fine-scale spatial information on the locations of species may also prompt legitimate concerns among contributors regarding possible unintended negative conservation impacts, violations of privacy and commercial exploitation of volunteer-gathered data. Here we canvas the attitudes towards open access of contributors (104 regional co-ordinators and 510 recorders) of species occurrence records to two of the largest citizen science biodiversity recording schemes, the UK's Butterflies for the New Millennium project and National Moth Recording Scheme. We find that while the majority of participants expressed support for open access in principle, most were more cautious in practice, preferring to limit the spatial resolution of records, particularly of threatened species, and restrict commercial reuse of data. In addition, citizen scientists' opinions differed between UK countries, taxonomic groups and the level of involvement volunteers had in the schemes. In order to maintain successful and democratic citizen science schemes, organisers, funders and data users must understand and respect participants' expectations and aspirations regarding open data while seeking to optimise data use for scientific and societal benefits.

**Keywords** Open data · Citizen science · Volunteer attitudes · Species occurrence · Biodiversity conservation · Lepidoptera

---

Communicated by Peter Bridgewater.

---

This article belongs to the Topical Collection: Biodiversity appreciation and engagement.

---

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s10531-019-01824-6>) contains supplementary material, which is available to authorized users.

---

✉ Richard Fox  
rf382@exeter.ac.uk

Extended author information available on the last page of the article

## Introduction

There is growing expectation and demand for open access to data in many areas of public life including science. In addition to the accepted scientific requirements of transparency and reproducibility, and the responsibility of public funding, this demand has been driven by the development of “big data” technologies enabling the storage and analysis of huge quantities of information (Arzberger et al. 2004; Farley et al. 2018). Scientists are increasingly willing to share data publicly (Tenopir et al. 2015), enabling other researchers to utilise and build upon freely-available archived data, resulting in benefits for society. An open access culture has developed in some scientific fields, notably genetics and genomics (Noor et al. 2006), although even here ethical concerns remain (McGuire et al. 2011; McEwen et al. 2013; Choudhury et al. 2014).

Ecologists, however, have been relatively slow to embrace open data, despite its potential to address many urgent, global, environmental pressures (Hampton et al. 2013; Poisot et al. 2013; Kenall et al. 2014; Soranno et al. 2015). Progress towards a more open approach in ecology is hindered by technological and cultural barriers, but solutions and incentives have emerged, alongside new obligations for public data archiving from funding organisations and scientific journals (Reichman et al. 2011; Michener 2015; Nosek et al. 2015; Culina et al. 2018a). Nevertheless, concerns remain about open access to ecological data, and while the views of scientists and organisations have been reported (Moles et al. 2013; Mills et al. 2015; Pearce-Higgins et al. 2018; Tulloch et al. 2018), the opinions of citizen scientists themselves have been overlooked.

Ecological data gathered through citizen science projects are increasingly useful, particularly for biodiversity monitoring and conservation (Chandler et al. 2017; Sullivan et al. 2017; Soroye et al. 2018). Unrestricted access to and reuse of citizen science ecological data maximises the societal and scientific returns on the efforts of volunteers; for example disclosure of locations of threatened species can encourage informed decision making about land-use changes that might impact biodiversity, improve species’ trend assessments, facilitate applied scientific research and help engage landowners, funders, politicians and the public in conservation (Tulloch et al. 2018). However, in the context of open access, citizen science data differ fundamentally from those collected in professional scientific research because the data are contributed by volunteers, who have their own views on data accessibility. It is widely expected that citizen science ecological data will be open access (Groom et al. 2017; Robinson et al. 2018), perhaps because it is supposed that people who contribute willingly and without material reward to citizen science projects would assume, or even insist, that their data are freely shared and publicly accessible. This assumption may not be justified, in part because the large number of citizen scientists are bound to encompass a diversity of views but also, specifically, because some participants have been engaged in gathering ecological data under different data exchange principles long before the advent of the “big data” era and the contemporary pressure for open access. Indeed, while the term citizen science was coined in the mid-1990s and the field has burgeoned since then (Silvertown 2009; Pocock et al. 2017), there is a long tradition of amateur naturalists gathering ecological, and particularly biogeographical, information (Miller-Rushing et al. 2012; Pocock et al. 2015; Strasser et al. 2019). In this tradition, the individual’s motivation to observe and study nature may have little to do with science or biodiversity conservation, leading to mismatches and tensions between the expectations of the scientific establishment and these participants in projects that are nowadays labelled as ‘citizen science’ (Ellis and Waterton 2004).

Thus, while some citizen science projects have an explicitly open data ethos (e.g. eBird, Sullivan et al. 2014), others do not (Groom et al. 2017). This may simply be because projects and their participants are continuing the historical legacy of mindsets, relationships and practices formed long before the advent of modern citizen science (Strasser et al. 2019) and do not conform to its expectations around open access. Alternatively, access to data may be restricted deliberately due to legitimate concerns from project organisers (Pearce-Higgins et al. 2018; Tulloch et al. 2018). One such concern is that unintended negative consequences of open access, for example harm to threatened species, could lead citizen scientists to cease participation, undermining project viability. It is important, therefore, that organisers, funders and users of citizen science are mindful of the views of participants regarding open access. While the motivations of citizen scientists taking part in biodiversity projects have been surveyed (Evans et al. 2005; Hobbs and White 2012; Wright et al. 2015; Domroese and Johnson 2017), their attitudes towards the onward use of the data that they contribute, and on the specific issue of open access to data, have rarely been considered (Ganzevoort et al. 2017).

These issues are of interest and importance to governmental and non-governmental organisations involved in conservation and research. For example, the charity Butterfly Conservation runs long-term citizen science schemes focussed on butterflies and moths (Lepidoptera) in the United Kingdom (UK). The schemes rely upon collaboration between paid staff (organising and promoting the schemes, managing databases, undertaking research and providing feedback to participants) and unpaid volunteers (undertaking species recording, computerisation and verification of records). Tens of thousands of volunteers are involved annually and the schemes have generated datasets that underpin assessments of UK Lepidoptera biodiversity change (e.g. Fox et al. 2014, 2015) and the delivery of species conservation (Ellis et al. 2012), as well as research e.g. into the impacts of environmental drivers such as climate change (e.g. Mason et al. 2015; Martay et al. 2017; Pearce-Higgins et al. 2017). In most cases, the data assembled through these schemes are not currently open access. Yet, given the considerable potential benefits for both biodiversity protection and scientific research of increasing access to these data, as well as the ethical impetus towards greater inclusivity (Soranno et al. 2015), the availability of these datasets should be reviewed and weighed against possible negative repercussions (e.g. impacts on threatened species or habitats, intrusion on participants' privacy or damage to partnerships with private landowners who have allowed access to otherwise closed land).

Therefore, to inform such a review and to provide practical recommendations to designers, organisers and funders of similar citizen science projects, we conducted surveys of volunteer participants in Butterfly Conservation recording schemes to seek a nuanced understanding of their views on open access to butterfly and moth occurrence data. Our study extends the approach of Ganzevoort et al. (2017), the only similar survey that we are aware of, by exploring the influence of spatial resolution, deferred data release and species threat on the attitudes of two different groups of volunteers with differing roles and levels of involvement in citizen science schemes, as well as contrasting the opinions of recorders of different taxa and in different UK countries. Our principal aim was to document the attitudes of these different groups of participants and understand how these may influence transition towards more open models of data accessibility. We did not seek to explore the motivations or values underlying participants' attitudes to open access and acknowledge that, as a result, the findings in this respect are limited. However, in addition to quantifying opinions, we sought to test the following hypotheses: (1) if the main concerns of citizen scientists related to potential damage to butterflies, moths or their habitats, rather than about privacy, confidentiality or intellectual property rights, then they would be more

reluctant to allow open access to records of threatened species compared to widespread ones and (2) that unwillingness to make threatened species records open access would be ameliorated by limiting (blurring) spatial location information and postponing the release of records for long periods (5 or more years).

## Methods

### Focal citizen science projects

The opinions towards open access of contributors to two UK-wide citizen science projects organised by Butterfly Conservation, Butterflies for the New Millennium (BNM; Asher et al. 2001) and the National Moth Recording Scheme (NMRS; Fox et al. 2011), were ascertained by questionnaires. The BNM was launched in 1995 and has, to date, collated 12.7 million butterfly species occurrence records covering the period 1690–2017. The NMRS commenced in 2006, initially focusing on macro-moth occurrence records (although it has now been extended to include all moth species), and has compiled 25 million macro-moth records for the period 1746–2016. These projects are among the largest citizen science biodiversity monitoring schemes globally, but the majority of BNM and NMRS records are not currently open access.

The flow of species occurrence records through the BNM and NMRS projects is organised in the same way. Observations made by citizen scientist recorders are sent to regional co-ordinators (also known as County Recorders), who are expert volunteers with the responsibility to collate and verify sightings for their area and maintain a local dataset of records. Copies of these local datasets are then pooled annually and, following further checks, added to the BNM or NMRS databases. At the time of this study, the BNM project included 65 regional co-ordinators and the NMRS 94. A few individuals fulfilled both roles for their area. The total numbers of citizen scientist recorders participating in the BNM and NMRS annually is unknown, because of inconsistencies in the way that individual recorder identities are logged across the schemes. However, given that each scheme currently collates c.1 million new records per annum, it is likely that there are tens of thousands of contributors at present. Some recorders take part in one but not the other scheme, whereas others contribute sightings to both.

The BNM and NMRS schemes collate opportunistic sightings of species from any location in the UK and on any date. Although there are minimum information standards for valid sightings, there are no sampling protocols—participants can record where, when and for as long as they wish. This traditional model of natural history recording (Pocock et al. 2015), separates the schemes on the one hand from systematic monitoring programmes with rigorous sampling protocols undertaken by experienced amateur or professional naturalists (e.g. the UK Butterfly Monitoring Scheme, Brereton et al. 2011; North American Breeding Bird Survey, Sauer et al. 2013) and, on the other, from modern citizen science projects that often aim to engage people with no previous involvement (e.g. Big Butterfly Count, Dennis et al. 2017; Great Pollinator Project, Domroese and Johnson 2017). Thus, while all BNM and NMRS participants are volunteers, their natural history expertise and recording behaviour vary greatly, as has been found in other biodiversity surveillance projects (Boakes et al. 2016; Everett and Geoghegan 2016).

Although it is difficult to categorize BNM and NMRS recorders on the basis of levels of engagement or expertise, different volunteer roles within the schemes provide a

clear dichotomy; individual regional co-ordinators are essential to the functioning of the schemes in a way that individual recorders are not, as without a regional co-ordinator in place no new records for that area will be provided to the scheme. While the opinions of both groups are important, the integral role of regional co-ordinators in the operation of the schemes necessitates an understanding of their attitudes to data sharing of the records in their custodianship as part of any prospective shift toward open access to the BNM and NMRS data. In addition, as curators of local datasets of species occurrence records, regional co-ordinators are likely to be familiar with the pros and cons of open access and, as expert naturalists, their views will be shaped by the traditions of data exchange within amateur natural history (Ellis and Waterton 2005; Ellis et al. 2005).

## Questionnaires

Separate questionnaires were designed to elucidate the views of regional co-ordinators and recorders and surveys were undertaken in May and June 2017. A longer questionnaire was used for regional co-ordinators so that we could gain a detailed understanding of the views of this key group of volunteers, while a much shorter, ‘light touch’ and entirely anonymous questionnaire was developed for recorders to maximise participation in the study.

### Regional co-ordinator questionnaire

The questionnaire for regional co-ordinators (Online Resource 1) aimed to ascertain the current level of support for and against open access and to gauge how such attitudes vary between volunteers in schemes for different taxa, in different countries and in response to perceived risk of negative impacts. Even when data are made publicly accessible, potential risks to species, habitats, sites or citizens can be moderated by restricting the information that is made available, by delaying the release of data and by legally restricting the uses to which data can be put. Thus, general support for open access was assessed by responses on a ten-point numerical scale (from 1 = serious reservations to 10 = strongly in favour), but subsequent questions asked participants to consider the appropriate spatial resolution of open records (i.e. how much records are blurred to conceal the precise location of species occurrence, with options of full capture resolution or blurring to  $1 \times 1$  km square,  $2 \times 2$  km square or  $10 \times 10$  km square), whether there should be a time lag before records are made public (with options of no lag, 5 year, 10 year or 20 year lags) and on the type of Creative Commons license that should be applied to open access UK butterfly and moth data. Developed as an alternative to traditional ‘all rights reserved’ copyright, Creative Commons licenses enable the copyright holder to choose which rights to reserve and which to waive, and have been widely adopted in many fields of human endeavour, including biodiversity monitoring (Hagedorn et al. 2011; Groom et al. 2017). Regional co-ordinators were asked for their opinion on the most appropriate of three Creative Commons licenses for UK butterfly and moth occurrence data; Zero (CC0), which has no restrictions on reuse, Attribution (CC-BY), which requires users to acknowledge the author/source, and Attribution-NonCommercial (CC-BY-NC), which requires acknowledgement and restricts reuse to non-commercial applications.

In addition to controlling data availability and use, the rarity or threat levels of taxa are likely to influence the perception of risk stemming from open access. The questionnaire sought to quantify this by asking respondents to consider the appropriate spatial resolution for open access records separately for widespread and threatened species. ‘Widespread’

and ‘threatened’ were not defined, so respondents used their own interpretation. In addition, regional co-ordinators were asked whether there were taxa or specific populations of taxa in their area that would require a more restrictive approach than the various open access options already discussed.

In total, the regional co-ordinator questionnaire included six questions with multiple-choice or scaled answers. Respondents were asked to provide their name and the geographical area for which they fulfil the role of regional co-ordinator. Questions were not obligatory and not all respondents completed all questions.

The questionnaire was sent by email attachment as a Microsoft Word document with a covering letter (Online Resource 2) to all UK regional co-ordinators in the BNM and NMRS networks on 10 May 2017. Regional co-ordinators were given until the end of May 2017 to respond, although responses received by 7 June 2017 were included in the analysis.

### Recorder questionnaire

A simpler questionnaire (Online Resource 3) was designed to canvas recorders’ views on open access and how recording behaviour might change in response to it. Just four multiple-choice questions were asked; two to segment respondents by UK country and taxonomic interest (recording butterflies, moths or both) and two relating to open access. Recorders were asked for their preferred open access spatial resolution for their own records via the BNM and NMRS schemes. Three options were provided: all records open at full capture resolution (i.e. the same level of spatial resolution as submitted by the recorder), widespread species at full resolution but scarce/threatened species at a summary (i.e. blurred) resolution, and all records at summary resolution. Secondly, to quantify the impacts (positive or negative) of moving to open access, recorders were asked about their likely behaviour towards the schemes if all records were made fully accessible. Four options were available; withhold future records from the schemes, blur the resolution of future contributed records, continue to participate as before, and increase support for the schemes by submitting more records.

All four questions were obligatory and the survey was anonymous. The questionnaire was an online survey designed using DotMailer ([www.dotmailer.com](http://www.dotmailer.com)). In late May 2017, the online questionnaire was promoted to recorders by the UK regional co-ordinators. It remained live for just over 2 weeks with data being extracted on 13 June 2017.

### Analysis

For each questionnaire, analysis was carried out on the aggregated responses but also separately after categorizing respondents by geographic or taxonomic interest, to test for differences between citizen scientists in different constituent countries of the UK (England, Scotland, Wales; Northern Ireland could not be tested separately due to a low sample size of responses to both questionnaires) and between recorders of butterflies, moths and both taxa. In addition, for the regional co-ordinator questionnaire data, we divided respondents into promoters, neutrals (passives) and detractors on the basis of their general support (on a 10-point scale) for open access to butterfly and moth records, using a slightly modified Net Promoter Score (NPS) methodology (Reichheld 2003; Keiningham et al. 2007). We classified those who scored 9 or 10 as promoters of open access, those who scored 5–8 as neutral and those scoring 1–4 as open access detractors. In standard NPS classification, scores as high as 6 are designated as detractors, but

we increased the neutral segment to reflect better the range of views of our respondents. Categorising in this way enabled us to compare the opinions of regional co-ordinators with different levels of overall support for the principle of open access to the questions about specific details of record resolution, temporal delays in data release and species threat status.

Each comparison was analysed initially using Pearson Chi squared and linear-by-linear association tests (Agresti 2002), accounting for the presence of ordinal variables. Where significant associations were found, cumulative link models with logit link were fitted, then Tukey-adjusted pairwise differences were investigated via least-squares means (LSM). All analyses were undertaken in R version 3.5.1 (R Core Team 2018) using the packages *ordinal* (Christensen 2018), *coin* (Hothorn et al. 2008) and *emmeans* (Lenth 2018). Goodness of fit for the cumulative link models was checked using likelihood ratio tests (*nominal\_test* and *scale\_test* in the *ordinal* package), in particular to assess whether the proportional odds assumption was satisfied. In some cases this assumption was not met, suggesting that the cumulative link model may not be appropriate, and in these instances pairwise differences among the explanatory variables were either assessed using the Cochran–Armitage test (with p values adjusted to account for false discovery rate) or only considered on the basis of summary statistics and figures.

## Ethics statement

Butterfly Conservation conforms strictly to appropriate legislation and codes of conduct relating to personal data and both questionnaires were designed and implemented in this context. For the regional co-ordinator questionnaire, full informed consent was obtained from all participants for the use of anonymised, aggregated responses in this research paper. Participants consented to the secure storage of data and access to the data by Butterfly Conservation employees involved in its analysis, and to publication of the arising results, for a period of 5 years, after which the data will be destroyed. Regional co-ordinator responses were anonymized prior to analysis. The online recorder questionnaire was completely anonymous and no personal data were collected. Participation in the questionnaires was voluntary and respondents were informed that the purpose was to gather views relating to open access to UK butterfly and moth occurrence data to aid the ongoing management and development of recording scheme data by Butterfly Conservation and other citizen science organisers.

## Results

### Regional co-ordinators

#### Survey coverage

Completed questionnaires were received from 104 regional co-ordinators representing response rates of 69% for the BNM and 68% for the NMRS networks. Responses were received from all four UK countries: 60 England, 2 Northern Ireland, 28 Scotland, 14 Wales.



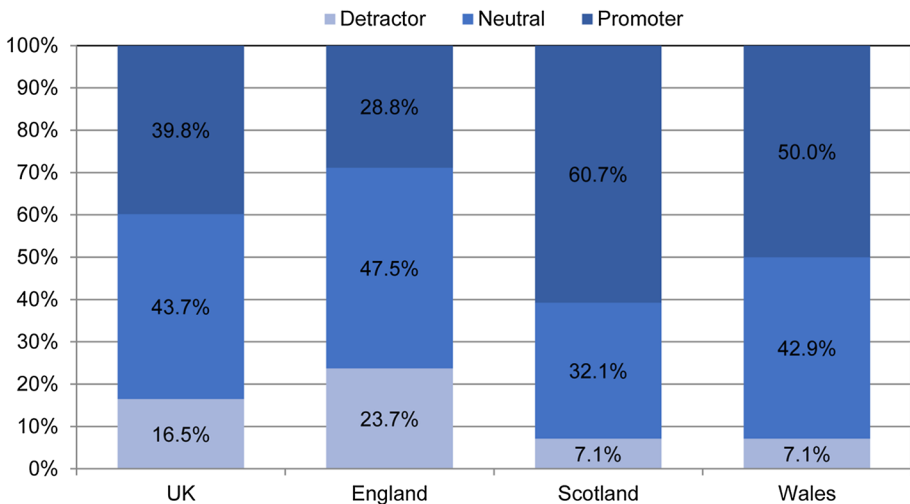
## Support for open access

Using our modified NPS scale, 39.8% of 103 regional co-ordinators who responded to this question were classified as open access promoters, 43.7% as neutrals and 16.5% as detractors. There was no difference in NPS value between respondents responsible for butterfly records, moth records and those who cover both taxa ( $\chi^2=3.257$ ,  $df=2$ ,  $p$  value=0.196), although regional co-ordinators for butterflies generally appeared to have more moderate NPS values than other co-ordinators, with smaller proportions in both the promoter and detractor classes (Online Resource 4).

Levels of general support for open access (measured with NPS) varied significantly between countries (Fig. 1,  $\chi^2=9.766$ ,  $df=2$ ,  $p=0.008$ ); regional co-ordinators in Scotland were more in favour of open access than their counterparts in England (England—Scotland contrast: LSM estimate=−0.485,  $z$  ratio=−3.252,  $p=0.003$ ). Respondents from Wales had similar NPS scores to those from Scotland, but the difference with England was not statistically significant (England—Wales contrast: LSM estimate=−0.364,  $z$  ratio=−1.852,  $p=0.153$ ).

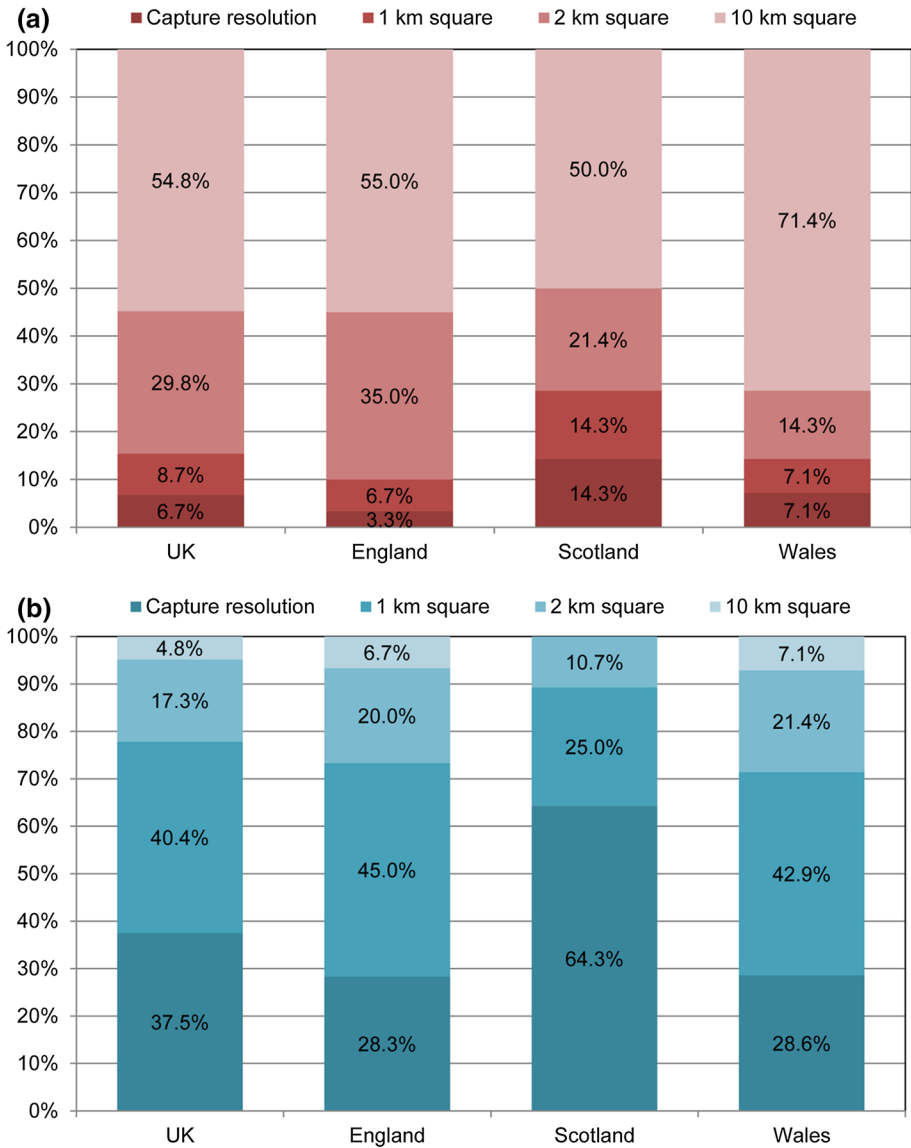
## Spatial resolution of records

For records of threatened species, only 6.7% of the 104 regional co-ordinators were in favour of open access at full capture resolution (Fig. 2a). The majority (54.8%) preferred records of such species to be accessible only at 10 × 10 km square (hereafter ‘10 km square’) scale, the coarsest resolution offered in the questionnaire, with a further 29.8% in favour of 2 × 2 km square (hereafter ‘2 km square’) scale. Attitudes were very different for records of widespread species. For these, 37.5% of regional co-ordinators were in favour of open access at full capture resolution, while a further 40.4% supported open access at 1 × 1 km square (hereafter ‘1 km square’) resolution and 17.3% chose



**Fig. 1** Levels of general support for open access, assessed by modified NPS categories, among regional co-ordinators from UK, England, Scotland and Wales (Northern Ireland not shown separately due to low sample size)





**Fig. 2** Preferred resolution of open access records of **a** threatened species and **b** widespread species among regional co-ordinators from UK, England, Scotland and Wales (Northern Ireland not shown separately due to low sample size)

the 2 km square scale (Fig. 2b). Only 4.8% (5 of 104 respondents) preferred the coarsest resolution option (10 km square) for records of widespread species. These results provide support for our hypotheses, suggesting that fear of ecological damage underlies regional co-ordinators’ concerns about open access (as they were much more restrictive about records of threatened species than widespread ones) and also that these concerns can be ameliorated by blurring the spatial resolution of accessible records.

For threatened species records there was no apparent difference between the responses from regional co-ordinators in different countries ( $\chi^2=3.364$ ,  $df=2$ ,  $p=0.186$ ), but there was a significant difference for widespread species ( $\chi^2=9.513$ ,  $df=2$ ,  $p=0.009$ ); regional co-ordinators in Scotland favoured finer resolution of open access records of widespread species than those in England (Scotland—England contrast: LSM estimate =  $-0.585$ ,  $z$  ratio =  $-3.493$ ,  $p=0.001$ ) (Fig. 2b). There was also a tendency for regional co-ordinators in Scotland to favour finer resolution access than those in Wales (Scotland—Wales contrast: LSM estimate =  $-0.604$ ,  $z$  ratio =  $-2.298$ ,  $p=0.056$ ). For example, in Scotland, 64.3% supported capture resolution access for widespread species compared to 28.3% in England and 28.6% in Wales.

There was a significant negative relationship between NPS category and preferred spatial resolution for both threatened (linear-by-linear association test  $Z=-3.794$ ,  $p=0.0001$ ) and widespread species ( $Z=-5.197$ ,  $p\leq 0.0001$ ), with detractors favouring the coarsest resolutions. For records of widespread species, detractors favoured a coarser resolution than both neutrals (detractors—neutrals contrast: LSM estimate =  $0.995$ ,  $z$  ratio =  $4.358$ ,  $p\leq 0.0001$ ) and promoters (detractors—promoters contrast: LSM estimate =  $1.396$ ,  $z$  ratio =  $6.207$ ,  $p\leq 0.0001$ ), and neutrals favoured a coarser resolution than promoters (neutrals—promoters contrast: LSM estimate =  $0.401$ ,  $z$  ratio =  $2.690$ ,  $p=0.020$ ).

Based on the goodness-of-fit tests, the cumulative link model was not reliable for pairwise contrasts between NPS categories and preferred spatial resolution for records of threatened species, but the responses suggest that detractors favoured coarser resolutions than neutrals who, in turn, favoured coarser resolutions than promoters (Online Resource 5). None of the regional co-ordinators who classified as detractors or neutrals and only 17.1% of promoters were in favour of capture resolution open access for threatened species records. Even at the 2 km square scale, only 17.6% detractors and 44.4% of neutrals were supportive, compared to the majority (58.5%) of promoters who were in favour of open access to records of threatened species at this resolution or even finer. In contrast, all of the regional co-ordinators classified as promoters or neutrals were in favour of open access to widespread species records at 2 km square resolution, along with 70.6% of detractors. However, even for records of widespread species there was only limited support for full resolution open access, with 61.0% promoters, 28.9% neutrals and just 5.9% detractors (corresponding to one respondent) in favour.

The preferred spatial resolution of open access records of threatened species differed between regional co-ordinators covering butterflies, moths or both taxa ( $\chi^2=9.376$ ,  $df=2$ ,  $p=0.009$ ) but there was no apparent difference for widespread species ( $\chi^2=0.852$ ,  $df=2$ ,  $p=0.653$ ). Regional co-ordinators for butterflies preferred finer resolution open access for threatened species records than their moth counterparts (butterfly co-ordinators—moth co-ordinators contrast: LSM estimate =  $-0.627$ ,  $z$  ratio =  $-3.441$ ,  $p=0.002$ ) or for those covering both taxa (butterfly co-ordinators—co-ordinators of both taxa contrast: LSM estimate =  $-0.676$ ,  $z$  ratio =  $-3.101$ ,  $p=0.006$ ). Only 28.9% of regional co-ordinators for butterflies considered that the coarsest resolution (10 km square) was required for open access records of threatened species, while 68.1% of regional co-ordinators for moths felt this was the appropriate resolution, as did 73.7% of co-ordinators responsible for both taxa.

## Time lags

Of the 100 regional co-ordinators that responded to the question about time lags, 74 favoured no delay to records being made open access, 21 supported a 5-year lag, 1 a

10-year lag and 4 a 20-year lag. NPS was significantly related to time lag (linear-by-linear association test  $Z = -5.351$ ,  $p \leq 0.0001$ ), with higher NPS correlated with shorter time lags. We were unable to undertake pairwise comparisons between NPS categories and time lags as the models did not satisfy goodness-of-fit tests. However, the significant relationship supports our hypothesis that concerns about open access can be lessened by deferring the release of records, at least among those regional co-ordinators who are generally more concerned about open access.

There was no apparent difference in the responses on time lags between regional co-ordinators covering different taxa ( $\chi^2 = 2.371$ ,  $df = 2$ ,  $p = 0.306$ ), but there was between countries ( $\chi^2 = 8.495$ ,  $df = 2$ ,  $p = 0.014$ ); only 11% (3 of 28 respondents) of regional co-ordinators in Scotland and 8% (1 of 13 respondents) in Wales advocated any time lag at all, and all of these were at the 5-year level, while 39% of 57 respondents in England were in favour of a delay in the release of records, including 9% who supported at least a 10-year delay. The difference in opinion on time lags was statistically significant (at the 5% level) between regional co-ordinators in England and Scotland (Cochran–Armitage test  $Z = 2.403$ ,  $p = 0.049$ ), but not between England and Wales (Cochran–Armitage test  $Z = 1.780$ ,  $p = 0.113$ ).

### **Additional restrictions for species or colonies**

70.1% of 97 regional co-ordinators who answered this question stated that no additional restrictions on open access were required for species and/or sites in their area beyond those provided by constraints on spatial resolution and time lags.

### **Creative commons licences**

Of the 103 regional co-ordinators who answered this question, 79.6% favoured the Attribution-NonCommercial licence (CC-BY-NC), the most restrictive of the three Creative Commons licence options offered on the questionnaire. Only 3.9% of respondents selected the most open licence option (CC0).

Opinions about Creative Commons licences differed between countries ( $\chi^2 = 8.105$ ,  $df = 2$ ,  $p = 0.017$ ). 46.2% of regional co-ordinators in Wales favoured the more open licences (CC0 and CC-BY), compared to 21.4% in Scotland and just 15.0% in England, but none of the pairwise comparisons were statistically significant (at the 5% level) using cumulative link models. There was no difference in views on Creative Commons licences between regional co-ordinators responsible for different taxa ( $\chi^2 = 0.659$ ,  $df = 2$ ,  $p = 0.719$ ).

### **Recorders**

#### **Survey coverage**

A total of 510 people completed the online questionnaire aimed at contributors of occurrence records to the BNM and NMRS. 25.3% identified as butterfly recorders, 25.5% as moth recorders and 49.2% stated that they recorded both groups. 367 (72.0%) respondents record mainly in England, 80 (15.7%) in Scotland, 58 (11.4%) in Wales and 5 (1.0%) in Northern Ireland.

## Spatial resolution of open access to own records

32.7% of respondents preferred full open access, opting for public access to all their records at capture resolution. A further 50.8% indicated that they were happy for their records of widespread species (but not those of scarce/threatened species) to be available at full capture resolution. Thus, for widespread species, 83.5% of respondents supported open access at capture resolution. In contrast, 16.5% of citizen scientists opposed capture resolution open access to any of their records (i.e. the spatial resolution of all records should be blurred to obscure precise locations), along with the 50.8% of respondents who thought that their records of scarce/threatened species should be blurred. Thus, 67.3% of respondents were against open access at capture resolution for some of their records. There were no significant differences between the views of recorders of different taxa ( $\chi^2=2.022$ ,  $df=2$ ,  $p=0.364$ ) or in the different countries ( $\chi^2=2.324$ ,  $df=2$ ,  $p=0.313$ ). The overall pattern, with a majority of recorders preferring to have their records of scarce/threatened species blurred but those of widespread species available at capture resolution provides further support for our two hypotheses; concern about ecological harm resulting from open access appears to be widespread among recorders and can be reduced by blurring the spatial resolution of records that are made publicly accessible.

## Future support for open access recording schemes

The majority of respondents (76.7%) indicated that their participation in the projects would be affected positively (4.5% would provide more records) or unaffected (72.2%) if all records were made open access in full detail. In contrast, the results suggest that the participation in the recording schemes of 23.3% of respondents would be detrimentally impacted, either due to them reducing the precision of the records they submit (21.2%) or withholding records entirely (2.2%). There were no significant differences in responses between countries ( $\chi^2=1.267$ ,  $df=2$ ,  $p=0.531$ ) or between recorders of different taxa ( $\chi^2=2.393$ ,  $df=2$ ,  $p=0.302$ ).

## Discussion

We have shown that while there are high levels of support in principle for open access among UK citizen scientists that contribute, collate and verify Lepidoptera occurrence data, they do not endorse full capture resolution open access nor unrestricted use of such data. Among the two groups of citizen scientists surveyed, only 6.7% of regional co-ordinators and 32.7% of recorders stated that records of all butterfly and moth species (widespread and threatened) should be open access at capture resolution, and 79.6% of regional co-ordinators felt that data reuse should be limited to non-commercial purposes. These findings are broadly similar to those in the only other study of citizen scientists' opinions that we are aware of; Ganzevoort et al. (2017) surveyed the demographics, motivations and views on data ownership and sharing of nearly 2200 volunteer biodiversity recorders in the Netherlands. They found that only 12.3% of biodiversity recorders in the Netherlands supported unconditional reuse of their data, while 36.7% were opposed to commercial use of their data.

Current limitations to access and reuse of citizen science data are often attributed to the scientists or organisations running citizen science projects, who may face a range of technological, economic and cultural barriers and disincentives to data sharing (Reichman et al. 2011; Schmidt et al. 2016; Groom et al. 2017; Pearce-Higgins et al. 2018). However, our UK results and those from the Netherlands suggest that some limitation is in accordance with the wishes and expectations of citizen science participants.

### Citizen scientist support for open access

Despite data quality concerns (Kosmala et al. 2016; Aceves-Bueno et al. 2017; Specht and Lewandowski 2018), citizen science has great potential to address pressing matters in biodiversity monitoring, conservation and research (Theobald et al. 2015; Chandler et al. 2017; Pocock et al. 2018). Open access to citizen science data would maximise this potential through increased reuse and the application of new ‘big data’ techniques and cross-disciplinary studies (Culina et al. 2018b; Farley et al. 2018; Ma et al. 2018; Tulloch et al. 2018), as well as yielding benefits of increased transparency and public trust in science (Soranno et al. 2015).

Surveys of citizen scientists’ motivations suggest support for these goals, with factors such as contributing to biodiversity conservation and science ranking highly (Hobbs and White 2012; Wright et al. 2015; West and Pateman 2016; Ganzevoort et al. 2017; Lewandowski and Oberhauser 2017). In keeping with this, our surveys of attitudes among UK citizen scientists suggest general support for open access, albeit with some concern about threatened species. 39.8% of UK regional co-ordinators were classified as promoters of open access on the basis of NPS, with a further 43.7% as neutrals, and support was stronger in some UK countries (60.7% promoters in Scotland and 50.0% promoters in Wales). Among the much larger group of recorders, 32.7% felt that all their records should be open access at capture resolution and 76.7% indicated that they would maintain or enhance their participation if the data were to be made completely open. Considering just records of widespread species, 37.5% of regional co-ordinators and 83.5% of recorders were in favour of open access at capture resolution, with the proportion of regional co-ordinators in favour rising to 77.9% if records were restricted to 1 km square resolution. In their survey of Dutch biodiversity recorders, Ganzevoort et al. (2017) also found evidence of general support for open access; 76.1% of citizen scientists regarded the data they contributed as a public good or as belonging to the organisation running the recording scheme i.e. they did not consider the data to be their personal property.

### Concerns and alleviating factors

Set against this general desire for data to be available and utilised are clear signals from our results and from other studies of concern regarding inappropriate use (Pearce-Higgins et al. 2018). As we did not ask participants about the motivations underlying their opinions on open access, discussion of their concerns must be speculative. It is well established that many citizen scientists want their records to contribute towards biodiversity conservation (e.g. Hobbs and White 2012; Lewandowski and Oberhauser 2017) but may be concerned that open access to data will undermine this goal. Threats to species (e.g. collectors targeting rare species, deliberate habitat destruction by land-owners to avoid conservation responsibility/land use restrictions, accidental damage to sites by naturalists wanting to see scarce species) are real (Tulloch et al. 2018),

but the levels of perceived risk are subjective and individualistic. Such concerns may also engender support for licences that prohibit commercial reuse; citizen scientists appear to support uses of their data that are likely to benefit biodiversity conservation, but not those that are thought to cause harm (Ellis and Waterton 2005; Ganzevoort et al. 2017). The perceived commodification of volunteer-gathered records, which runs counter to the traditional culture of data exchange within natural history, and a lack of transparency and feedback about the onward uses of the data may also contribute to restrictive attitudes towards licensing (Ellis and Waterton 2005). Other concerns may exist around privacy and the potential malicious use of personal information (e.g. names and locations of recorders) derived from species occurrence data (Bowser et al. 2014).

We extended the previous study by Ganzevoort et al. (2017) to gain a more nuanced understanding of these concerns and explored how citizen scientists' attitudes to open access were moderated by variation in spatial and temporal factors. We hypothesised that if concerns about open access related to potential damage to individual organisms, populations and habitats, then citizen scientists would be more restrictive with records of threatened species than widespread ones. Additionally, we posited that restricting the spatial resolution of publicly accessible data or delaying the release of data may both be expected to reduce the perceived risk. Other commonly raised fears around the personal privacy of the recorders themselves and of private land where charismatic species are present (which may be subject to trespass if the precise locations are made public) might also be ameliorated by such restrictions.

We found strong evidence to support both our hypotheses. There was a clear effect of spatial scale on attitudes to open access for UK Lepidoptera records (but not for the use of deferred release of data i.e. time lags). 37.5% of regional co-ordinators were in favour of capture resolution open access for records of widespread species and this rose cumulatively as the spatial scale was coarsened, such that 77.9% were in favour at 1 km square resolution and 95.2% in favour at 2 km square resolution. The impact of spatial resolution on open access opinions was even more pronounced when considering records of threatened species; regional co-ordinators were more cautious, with only 6.7% in favour at capture resolution, rising cumulatively to 15.4% at 1 km square and 45.2% at 2 km square resolution. Similar patterns were found when regional co-ordinators were grouped by general levels of support for open access (NPS categories) and each analysed separately.

The survey of recorders also suggested that spatial scale was an important factor in citizen scientists' attitudes towards open data. Generally, recorders were more supportive than regional co-ordinators of open access at capture resolution. Nevertheless, two-thirds (67.3%) of recorders felt that some (i.e. threatened species) or all of their records should be blurred to a coarser resolution than capture level for open access.

Therefore, although we did not attempt to determine the rationale underlying the opinions of citizen scientists, these results support both our hypotheses. The greater unwillingness to release records of threatened species at full capture resolution compared to records of widespread species suggests that the main concerns of citizen scientists relate to potential negative ecological impacts, rather than unease about privacy, confidentiality or intellectual property rights. Second, for the majority of contributors these concerns can be alleviated by blurring spatial location information. Interestingly, most respondents did not support deferral of the open release of records in addition to spatial restrictions, although 26.0% were in favour of a delay of at least 5 years.

## Differences between roles, countries and taxa

The differing nature of the roles of regional co-ordinators and recorders and the fact that they were asked different questions makes it inappropriate to undertake a direct statistical comparison of their views. In addition, it is probable that some regional co-ordinators also completed the recorder questionnaire and so the two samples may not be independent. The findings on spatial resolution suggest, however, that the regional co-ordinators were more restrictive, on average, than recorders in their attitudes to open access. Further work is required to elucidate the causes of the seemingly greater risk aversion among regional co-ordinators, as our questionnaires did not examine the reasons underlying stated opinions. They may stem from complex combinations of ecological (e.g. increased awareness of possible threats to species), legal (e.g. concerns about acts of trespass and personal data under the General Data Protection Regulation), personal (e.g. greater time investment in the data), ethical (e.g. a sense of responsibility as custodians of records contributed by other citizen scientists) and cultural (e.g. traditional principles of data exchange in natural history) considerations. The latter may be particularly important given that regional co-ordinators are amateur expert naturalists, whereas recorders are a much more diverse group ranging from committed amateur naturalists to complete beginners (e.g. see Everett and Geoghegan 2016). Irrespective of the causes, however, if restrictions on open access to recording scheme data, informed by the views of regional co-ordinators, are contrary to the wishes of most citizen scientist participants, this may risk demotivation, loss of support and, ultimately, reduced levels of species recording.

Significant differences were found between the opinions of regional co-ordinators in England and Scotland. Regional co-ordinators in Scotland had higher NPS values than their counterparts in England, indicating greater support in general for the principles of open access to Lepidoptera occurrence records. This predisposition was reflected in attitudes to more specific options, whereby regional co-ordinators in Scotland favoured finer spatial scale resolution of open access records for widespread species and shorter time lags before records are released than their colleagues in England.

The causes of these differences are not known and require further research. However, we speculate that two factors may contribute to these contrasting attitudes. First, long-term abundance trends of butterflies and moths differ geographically within the UK. The abundance of 337 species of widespread moths has decreased significantly in southern Britain (most of England and all of Wales) but not in northern Britain (Scotland plus part of northern England) (Conrad et al. 2006). Similarly, the abundance of wider countryside butterflies has decreased significantly in England but not in Scotland (Fox et al. 2015). Thus, regional co-ordinators in England, where greater declines have occurred, might be more sensitive to potential adverse effects on butterflies and moths arising from open access to data and this results in more restrictive attitudes than regional co-ordinators in Scotland.

Second, there are substantial differences between Scotland and the rest of the UK in the legal framework relating to public access to land. The Land Reform (Scotland) Act 2003 confers a public ‘right to roam’ over almost all land in Scotland, while similar rights (under the Countryside and Rights of Way Act 2000) cover only c.8–12% of the total land area of England and Wales (Lovett 2012). The situation is even more restrictive in Northern Ireland. Regional co-ordinators in Scotland may have reduced concerns, therefore, compared to their counterparts in other UK countries, about either



exposing acts of trespass by recorders or inadvertently encouraging others to trespass on private land (thereby undermining relationships between recorders and landowners) as a result of records being made open access.

Interestingly, the online survey of recorders found no significant differences between UK countries. This suggests that the differing attitudes of regional co-ordinators in England and Scotland relates to their roles as custodians of local datasets.

In contrast to the clear country-level differences, attitudes of regional co-ordinators varied very little depending on the taxon (butterflies, moths or both) for which they have responsibility. The only significant result in our analysis was that regional co-ordinators for butterflies favoured finer spatial resolution open access for records of threatened species than regional co-ordinators who cover moths or both taxa. Possible reasons for this might include that there are more UK populations of the most threatened butterflies than the most threatened moths, that sites for threatened butterflies are often well known already or that extra visitors to sites of threatened butterflies are likely to be less intrusive for landowners than those wanting to see threatened moths if the latter are nocturnal. There were no significant differences between the opinions of recorders based on taxon of interest.

### Wider applicability

The wider applicability of our findings depends on the representativeness of our sampling. With 69% and 68% response rates among regional co-ordinators, we can have high confidence that our results are representative of this key group of UK Lepidoptera-recording volunteers. However, we do not know how many people participate annually in the BNM and NMRS recording schemes, so we cannot measure the response rate for our online questionnaire aimed at recorders. While 510 responses is reasonable for statistical analysis, it likely represents only a small proportion of the total number of citizen science contributors to these projects. In addition, the sample is likely to be biased, as the online survey was not distributed randomly or systematically but promoted to recorders by the regional co-ordinators. This clearly limits our ability to generalise from the findings.

Another limitation stems from variation between participants. Analyses of this variation have classified citizen scientists by expertise in species identification and by temporal and spatial patterns of participation in particular projects (Ponciano and Brasileiro 2014; Boakes et al. 2016; Everett and Geoghegan 2016; Johnston et al. 2018). Boakes et al. (2016), for example, categorised citizen scientists undertaking biodiversity recording as ‘dabbler’, ‘steady’ or ‘enthusiast’ depending on their temporal participation, while Everett and Geoghegan (2016) utilised a continuum of engagement, on the basis of past involvement in natural history. While all citizen scientists can contribute useful data, their motivations and strength of commitment to particular projects vary considerably between individuals and also over time for individuals. It is likely that attitudes towards open access to citizen science data would also vary between individuals and over time, and might covary with other metrics describing the engagement behaviour of citizen scientists. By definition, given their role and responsibilities to the BNM and NMRS projects, the regional co-ordinators that took part in our study are highly motivated, committed and knowledgeable volunteers, many of whom have a passion for biodiversity conservation. Their views on open access are of fundamental importance for the ongoing development of the BNM and NMRS projects, but cannot reasonably be generalised to the thousands of citizens who participate to a greater or lesser extent in the schemes. Similarly, as the recorders who responded to our online questionnaire were not selected at random, it

is likely that these may also be a biased sample, with views on open access that might differ from less active or more recent participants.

Even within our sampled audience of citizen scientists, we found evidence of differences in attitude towards open access between countries. Whatever the causes, this variation within the UK suggests that there will also be differences between the UK and other countries. This limits the applicability of our results but stresses the importance of seeking the opinions of and establishing dialogue with citizen scientists on this issue, rather than making assumptions.

### Practical recommendations for citizen science

A key factor in the creation of a citizen science project is the development of a comprehensive yet clear data policy (James 2011). This needs to take into account not only the requirements of the project itself, and its aspirations for future data sharing and scientific publication, but also any legal requirements for open access imposed by funding organisations. For example, in the UK Butterfly Monitoring Scheme, a systematic monitoring programme run by Butterfly Conservation and partner organisations, it is a condition of long-standing financial support from government departments and agencies that data are made freely available under an Open Government Licence. Schemes such as those addressed in this study, which are not bound by funder requirements regarding open data, provide an opportunity to plan data access in the light of contributors' attitudes. A data policy must, of course, also comply with relevant legislation relating to the protection of personal data, such as the European Union's General Data Protection Regulation. The use of widely recognised licences, such as Creative Commons licences, is recommended to ensure clarity for both participants and prospective data users, as well as compatibility with other projects and data repositories (e.g. the Atlas of Living Australia, [www.ala.org.au](http://www.ala.org.au)) locally and globally. Most importantly, we recommend that any data policy developed for a citizen science project should be actively disseminated to potential contributors to ensure that they are aware of the uses to which their data will be put and are therefore able to make an informed choice prior to participation.

Despite its limitations, our study provides useful information on the development of open access data policies that is of wider relevance to biodiversity citizen science projects. In particular, the heterogeneity of views present in these relatively small samples shows that organisers would be well-advised to consult with potential participants on matters of data access in advance as part of project development. Similarly, funding organisations, statutory agencies and policy makers may wish to reflect on the diversity of views revealed by our questionnaires, and previous studies (e.g. Ellis et al. 2005), in their drive for open citizen science data. Our results suggest that the cultural context is likely to be extremely important in influencing attitudes to open access among citizen scientists; not only are these likely to differ substantially between nationalities, but also between participants with different roles in projects and levels of past engagement with natural history and citizen science.

### Conclusions

In order to maximise the scientific and societal benefits of citizen science, the views and motivations of participants must be considered. Our study shows that, contrary to common assumptions, UK citizen scientists taking part in butterfly and moth recording have diverse, in some cases polarised, views on open access and there was substantial variation between different countries and between volunteers with different roles. Overall, many participants

are supportive, in principle, of open access to the data they gather, but are mindful of possible negative ecological impacts that may result. Our results suggest that the majority of participants favour increasing access to these data, and that the concerns of many could be ameliorated by limiting the spatial resolution of open records, particularly of threatened species, and licencing reuse for non-commercial purposes. Globally, citizen science schemes have great potential to help address the enormous challenges facing biodiversity, but to do so effectively, must be responsive to the changing attitudes and new opportunities afforded by open data.

**Acknowledgements** We are grateful to all of the volunteer regional co-ordinators and recorders who contribute to the BNM and NMRS projects and, in particular, to those who participated in the surveys of opinion reported here. We would like to thank Dr Katie Cruickshanks and two anonymous reviewers for helpful comments on previous drafts of this paper. The BNM and NMRS are run by Butterfly Conservation and receive financial support from Natural England.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Aceves-Bueno E, Adeleye AS, Feraud M, Huang Y, Tao M, Yang Y, Anderson SE (2017) The accuracy of citizen science data: a quantitative review. *Bull Ecol Soc Am* 98:278–290. <https://doi.org/10.1002/bes2.1336>
- Agresti A (2002) *Categorical data analysis*, 2nd edn. Wiley, Hoboken
- Arzberger P, Schroeder P, Beaulieu A, Bowker G, Casey K, Laaksonen L, Moorman D, Uhlir P, Wouters P (2004) An international framework to promote access to data. *Science* 303:1777–1778. <https://doi.org/10.1126/science.1095958>
- Asher J, Warren M, Fox R, Harding P, Jeffcoate G, Jeffcoate S (2001) *The millennium atlas of butterflies in Britain and Ireland*. Oxford University Press, Oxford
- Boakes EH, Gliozzo G, Seymour V, Harvey M, Smith C, Roy DB, Haklay M (2016) Patterns of contribution to citizen science biodiversity projects increase understanding of volunteers' recording behaviour. *Sci Rep* 6:33051. <https://doi.org/10.1038/srep33051>
- Bowser A, Wiggins A, Shanley L, Preece J, Henderson S (2014) Sharing data while protecting privacy in citizen science. *Interactions* 21:70–73. <https://doi.org/10.1145/2540032>
- Brereton TM, Roy DB, Middlebrook I, Botham MS, Warren MS (2011) The development of butterfly indicators in the United Kingdom and assessments in 2010. *J Insect Conserv* 15:139–151. <https://doi.org/10.1007/s10841-010-9333-z>
- Chandler M, See L, Copas K, Bonde AMZ, Claramunt López B, Danielsen F, Kristoffer Legind J, Masinde S, Miller-Rushing AJ, Newman G, Rosemartin A, Turak E (2017) Contribution of citizen science towards international biodiversity monitoring. *Biol Conserv* 213:280–294. <https://doi.org/10.1016/j.biocon.2016.09.004>
- Choudhury S, Fishman JR, McGowan ML, Juengst ET (2014) Big data, open science and the brain: lessons learned from genomics. *Front Hum Neurosci* 8:239. <https://doi.org/10.3389/fnhum.2014.00239>
- Christensen, RHB (2018) Ordinal—regression models for ordinal data. R package version 2018.8-25. <http://www.cran.r-project.org/package=ordinal/>. Accessed 19 Oct 2018
- Conrad KF, Warren M, Fox R, Parsons M, Woiwod IP (2006) Rapid declines of common, widespread British moths provide evidence of an insect biodiversity crisis. *Biol Conserv* 132:279–291. <https://doi.org/10.1016/j.biocon.2006.04.020>

- Culina A, Baglioni M, Crowther TW, Visser ME, Woutersen-Windhouwer S, Manghi P (2018a) Navigating the unfolding open data landscape in ecology and evolution. *Nat Ecol Evol* 2:420–426. <https://doi.org/10.1038/s41559-017-0458-2>
- Culina A, Crowther TW, Ramakers JJC, Gienapp P, Visser ME (2018b) How to do meta-analysis of open datasets. *Nat Ecol Evol* 2:1053–1056. <https://doi.org/10.1038/s41559-018-0579-2>
- Dennis EB, Morgan BJT, Brereton TM, Roy DB, Fox R (2017) Using citizen science butterfly counts to predict species population trends. *Conserv Biol* 31:1350–1361. <https://doi.org/10.1111/cobi.12956>
- Domroese MC, Johnson EA (2017) Why watch bees? Motivations of citizen science volunteers in the Great Pollinator Project. *Biol Conserv* 208:40–47. <https://doi.org/10.1016/j.biocon.2016.08.020>
- Ellis R, Waterton C (2004) Environmental citizenship in the making: the participation of volunteer naturalists in UK biological recording and biodiversity policy. *Sci Public Policy* 31:95–105. <https://doi.org/10.3152/147154304781780055>
- Ellis R, Waterton C (2005) Caught between the cartographic and the ethnographic imagination: the whereabouts of amateurs, professionals, and nature in knowing biodiversity. *Environ Plan D* 23:673–693. <https://doi.org/10.1068/d353t>
- Ellis R, Grove-White R, Vogel J, Waterton C (2005) Nature: who knows? English Nature, the Natural History Museum and Lancaster University, Lancaster
- Ellis S, Bourn NAD, Bulman CR (2012) Landscape-scale conservation for butterflies and moths: lessons from the UK. *Butterfly Conservation*, Wareham
- Evans C, Abrams E, Reitsma R, Roux K, Salmonsén L, Marra PP (2005) The Neighborhood Nestwatch Program: participant outcomes of a citizen-science ecological research project. *Conserv Biol* 19:589–594. <https://doi.org/10.1111/j.1523-1739.2005.00s01.x>
- Everett G, Geoghegan H (2016) Initiating and continuing participation in citizen science for natural history. *BMC Ecol* 16:15–22. <https://doi.org/10.1186/s12898-016-0062-3>
- Farley SS, Dawson A, Goring SJ, Williams JW (2018) Situating ecology as a big-data science: current advances, challenges, and solutions. *Bioscience* 68:563–576. <https://doi.org/10.1093/biosci/biy068>
- Fox R, Randle Z, Hill L, Anders S, Wiffen L, Parsons MS (2011) Moths Count: recording moths for conservation in the UK. *J Insect Conserv* 15:55–68. <https://doi.org/10.1007/s10841-010-9309-z>
- Fox R, Oliver TH, Harrower C, Parsons MS, Thomas CD, Roy DB (2014) Long-term changes to the frequency of occurrence of British moths are consistent with opposing and synergistic effects of climate and land use changes. *J Appl Ecol* 51:949–957. <https://doi.org/10.1111/1365-2664.12256>
- Fox R, Brereton TM, Asher J, August TA, Botham MS, Bourn NAD, Cruickshanks KL, Bulman CR, Ellis S, Harrower CA, Middlebrook I, Noble DG, Powney GD, Randle Z, Warren MS, Roy DB (2015) The state of the UK's butterflies 2015. *Butterfly Conservation and the Centre for Ecology & Hydrology*, Wareham
- Ganzevoort W, van den Born RJG, Halfman W, Turnhout S (2017) Sharing biodiversity data: citizen scientists' concerns and motivations. *Biodivers Conserv* 26:2821–2837. <https://doi.org/10.1007/s10531-017-1391-z>
- Groom Q, Weatherdon L, Geijzendorffer IR (2017) Is citizen science an open science in the case of biodiversity observations? *J Appl Ecol* 54:612–617. <https://doi.org/10.1111/1365-2664.12767>
- Hagedorn G, Mietchen D, Morris RA, Agosti D, Penev L, Berendsohn WG, Hobern D (2011) Creative Commons licenses and the non-commercial condition: implications for the re-use of biodiversity information. *ZooKeys* 150:127–149
- Hampton SE, Strasser CA, Tewksbury JJ, Gram WK, Budden AE, Batcheller AL, Duke CS, Porter JH (2013) Big data and the future of ecology. *Front Ecol Environ* 11:156–162. <https://doi.org/10.1890/120103>
- Hobbs SJ, White PCL (2012) Motivations and barriers in relation to community participation in biodiversity recording. *J Nat Conserv* 20:364–373. <https://doi.org/10.1016/j.jnc.2012.08.002>
- Hothorn T, Hornik K, van de Wiel MA, Zeileis A (2008) Implementing a class of permutation tests: the coin package. *J Stat Softw* 60:257–263. <https://doi.org/10.18637/jss.v028.i08>
- James T (2011) Running a biological recording scheme or survey. National Biodiversity Network, Nottingham
- Johnston A, Fink D, Hochachka WM, Kelling S (2018) Estimates of observer expertise improve species distributions from citizen science data. *Methods Ecol Evol* 9:88–97. <https://doi.org/10.1111/2041-210X.12838>
- Keiningham TL, Cooil B, Andreassen TW, Aksoy L (2007) A longitudinal examination of net promoter and firm revenue growth. *J Mark* 71:39–51. <https://doi.org/10.1509/jmkg.71.3.39>
- Kenall A, Harold S, Foote C (2014) An open future for ecological and evolutionary data? *BMC Evol Biol* 14:66. <https://doi.org/10.1186/1471-2148-14-66>

- Kosmala M, Wiggins A, Swanson A, Simmons B (2016) Assessing data quality in citizen science. *Front Ecol Environ* 14:551–560. <https://doi.org/10.1002/fee.1436>
- Lenth R (2018) emmeans: estimated marginal means, aka least-squares means. R package version 1.2.4. <https://CRAN.R-project.org/package=emmeans>. Accessed 19 Oct 2018
- Lewandowski EJ, Oberhauser KS (2017) Butterfly citizen scientists in the United States increase their engagement in conservation. *Biol Conserv* 208:106–112. <https://doi.org/10.1016/j.biocon.2015.07.029>
- Lovett JA (2012) The right to exclude meets the right of responsible access: Scotland's bold experiment in public access legislation. *Probate Prop* 26:52–59
- Ma A, Bohan DA, Canard E, Derocles SAP, Gray C, Lu X, Macfadyen S, Romero GQ, Kratina P (2018) A replicated network approach to 'big data' in ecology. *Adv Ecol Res* 59:225–264. <https://doi.org/10.1016/bs.aecr.2018.04.001>
- Martay B, Brewer MJ, Elston DA, Bell JR, Harrington R, Brereton TM, Barlow KE, Botham MS, Pearce-Higgins JW (2017) Impacts of climate change on national biodiversity population trends. *Ecography* 40:1139–1151. <https://doi.org/10.1111/ecog.02411>
- Mason SC, Palmer G, Fox R, Gillings S, Hill JK, Thomas CD, Oliver TH (2015) Geographical range margins of a wide range of taxonomic groups continue to shift polewards. *Biol J Linn Soc* 115:586–597. <https://doi.org/10.1111/bij.12574>
- McEwen JE, Boyer JT, Sun KY (2013) Evolving approaches to the ethical management of genomic data. *Trends Genet* 29:375–382. <https://doi.org/10.1016/j.tig.2013.02.001>
- McGuire AL, Basford M, Dressler LG, Fullerton SM, Koenig BA, Li R, McCarty CA, Ramos E, Smith ME, Somkin CP, Waudby C, Wolf WA, Clayton EW (2011) Ethical and practical challenges of sharing data from genome-wide association studies: the eMERGE Consortium experience. *Genome Res* 21:1001–1007. <https://doi.org/10.1101/gr.120329.111>
- Michener WK (2015) Ecological data sharing. *Ecol Inform* 29:33–44. <https://doi.org/10.1016/j.ecoinf.2015.06.010>
- Miller-Rushing A, Primack R, Bonney R (2012) The history of public participation in ecological research. *Front Ecol Environ* 10:285–290. <https://doi.org/10.1890/110278>
- Mills JA et al (2015) Archiving primary data: solutions for long-term studies. *Trends Ecol Evol* 30:581–589. <https://doi.org/10.1016/j.tree.2015.07.006>
- Moles AT, Dickie JB, Flores-Moreno H (2013) A response to Poisot et al.: publishing your dataset is not always virtuous. *Ideas Ecol Evol* 6:20–22. <https://doi.org/10.4033/iee.2013.6b.15.f>
- Noor MA, Zimmermann KJ, Teeter KC (2006) Data sharing: how much doesn't get submitted to GenBank? *PLoS Biol* 4:e228. <https://doi.org/10.1371/journal.pbio.0040228>
- Nosek BA et al (2015) Promoting an open research culture. *Science* 348:1422–1425. <https://doi.org/10.1126/science.aab2374>
- Pearce-Higgins JW et al (2017) A national-scale assessment of climate change impacts on species: assessing the balance of risks and opportunities for multiple taxa. *Biol Conserv* 213:124–134. <https://doi.org/10.1016/j.biocon.2017.06.035>
- Pearce-Higgins JW, Baillie SR, Boughey K, Bourn NAD, Foppen RPB, Gillings S, Gregory RD, Hunt T, Jiguet F, Lehtikoinen A, Musgrove AJ, Robinson RA, Roy DB, Siriwardena GM, Walker KJ, Wilson JD (2018) Overcoming the challenges of public data archiving for citizen science biodiversity recording and monitoring schemes. *J Appl Ecol* 55:2544–2551. <https://doi.org/10.1111/1365-2664.13180>
- Pocock MJO, Roy HE, Preston CD, Roy DB (2015) The Biological Records Centre: a pioneer of citizen science. *Biol J Linn Soc* 115:475–493. <https://doi.org/10.1111/bij.12548>
- Pocock MJO, Tweddle JC, Savage J, Robinson LD, Roy HE (2017) The diversity and evolution of ecological and environmental citizen science. *PLoS ONE* 12:e0172579. <https://doi.org/10.1371/journal.pone.0172579>
- Pocock MJO, Chandler M, Bonney R, Thornhill I, Albin A, August T, Bachman S, Brown PMJ, Gasparini Fernandes Cunha D, Grez A, Jackson C, Peters M, Romer Rabarijaon N, Roy HE, Zaviezo T, Danielsen F (2018) A vision for global biodiversity monitoring with citizen science. *Adv Ecol Res* 59:169–223. <https://doi.org/10.1016/bs.aecr.2018.06.003>
- Poisot T, Mounce R, Gravel D (2013) Moving toward a sustainable ecological science: don't let data go to waste! *Ideas Ecol Evol* 6:11–19. <https://doi.org/10.4033/iee.2013.6b.14.f>
- Ponciano L, Brasileiro F (2014) Finding volunteers' engagement profiles in human computation for citizen science projects. *Hum Comput* 1:247–266. <https://doi.org/10.15346/hc.v1i2.12>
- R Core Team (2018) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
- Reichheld FF (2003) The one number you need to grow. *Harv Bus Rev* 81:46–54
- Reichman OJ, Jones MB, Schildhauer MP (2011) Challenges and opportunities of open data in ecology. *Science* 331:703–705. <https://doi.org/10.1126/science.1197962>

- Robinson LD, Cawthray JL, West SE, Bonn A, Ansin J (2018) Ten principles of citizen science. In: Hecker S, Haklay M, Bowser A, Makuch Z, Vogel J, Bonn A (eds) *Citizen science: innovation in open science, society and policy*. UCL Press, London, pp 27–40
- Sauer JR, Link WA, Fallone JE, Pardieck KL, Ziolkowski DJ (2013) The North American Breeding Bird Survey 1966–2011: summary analysis and species accounts. *N Am Fauna* 97:1–32. <https://doi.org/10.3996/nafa.79.0001>
- Schmidt B, Gemeinholzer B, Treloar A (2016) Open data in global environmental research: the Belmont Forum’s open data survey. *PLoS ONE* 11:e0146695. <https://doi.org/10.1371/journal.pone.0146695>
- Silvertown J (2009) A new dawn for citizen science. *Trends Ecol Evol* 24:467–471. <https://doi.org/10.1016/j.tree.2009.03.017>
- Soranno PA, Cheruvilil KS, Elliott KC, Montgomery GM (2015) It’s good to share: why environmental scientists’ ethics are out of date. *Bioscience* 65:69–73. <https://doi.org/10.1093/biosci/biu169>
- Soroye P, Ahmed N, Kerr JT (2018) Opportunistic citizen science data transform understanding of species distributions, phenology, and diversity gradients for global change research. *Glob Change Biol* 24:5281–5291. <https://doi.org/10.1111/gcb.14358>
- Specht H, Lewandowski E (2018) Biased assumptions and oversimplifications in evaluations of citizen science data quality. *Bull Ecol Soc Am* 99:251–256. <https://doi.org/10.1002/bes2.1388>
- Strasser BJ, Baudry J, Mahr D, Sanchez G, Tancoigne E (2019) “Citizen Science”? Rethinking science and public participation. *Sci Technol Stud* 32:52–76. <https://doi.org/10.23987/sts.60425>
- Sullivan BL et al (2014) The eBird enterprise: an integrated approach to development and application of citizen science. *Biol Conserv* 169:31–40. <https://doi.org/10.1016/j.biocon.2013.11.003>
- Sullivan BL, Phillips T, Dayer AA, Wood CL, Farnsworth A, Iliff MJ, Davies IJ, Wiggins A, Fink D, Hochachka WM, Rodewald AD, Rosenberg KV, Bonney R, Kelling S (2017) Using open access observational data for conservation action: a case study for birds. *Biol Conserv* 208:5–14. <https://doi.org/10.1016/j.biocon.2016.04.031>
- Tenopir C, Dalton ED, Allard S, Frame M, Pjesivac I, Birch B, Pollock D, Dorsett K (2015) Changes in data sharing and data reuse practices and perceptions among scientists worldwide. *PLoS ONE* 10:e0134826. <https://doi.org/10.1371/journal.pone.0134826>
- Theobald EJ, Ettinger AK, Burgess HK, DeBey LB, Schmidt NR, Froehlich HE, Wagner C, HilleRisLambers J, Tewksbury J, Harsch MA, Parrish JK (2015) Global change and local solutions: tapping the unrealized potential of citizen science for biodiversity research. *Biol Conserv* 181:236–244. <https://doi.org/10.1016/j.biocon.2014.10.021>
- Tulloch AIT et al (2018) A decision tree for assessing the risks and benefits of publishing biodiversity data. *Nat Ecol Evol* 2:1209–1217. <https://doi.org/10.1038/s41559-018-0608-1>
- West S, Pateman R (2016) Recruiting and retaining participants in citizen science: what can be learned from the volunteering literature? *Citiz Sci Theory Pract* 1:15. <https://doi.org/10.5334/cstp.8>
- Wright DR, Underhill LG, Keene M, Knight AT (2015) Understanding the motivations and satisfactions of volunteers to improve the effectiveness of citizen science programs. *Soc Nat Resour* 28:1013–1029. <https://doi.org/10.1080/08941920.2015.1054976>

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Affiliations

Richard Fox<sup>1,2,3</sup>  · Nigel A. D. Bourn<sup>1</sup>  · Emily B. Dennis<sup>1,4</sup>  · Richard T. Heafield<sup>1</sup>  · Ilya M. D. Maclean<sup>3</sup>  · Robert J. Wilson<sup>2,5</sup> 

<sup>1</sup> Butterfly Conservation, Wareham, Dorset, UK

<sup>2</sup> Biosciences, University of Exeter, Exeter, UK

<sup>3</sup> Environment and Sustainability Institute, University of Exeter, Penryn, UK

<sup>4</sup> School of Mathematics, Statistics and Actuarial Science, University of Kent, Canterbury, UK

<sup>5</sup> National Museum of Natural Sciences (MNCN-CSIC), 28006 Madrid, Spain