JNCC Report
No. 640

United Kingdom Terrestrial Evidence Partnership of Partnerships data products: improving opportunities for re-use

August, T.[1], Siriwardena, G.[2] & Isaac, N.[1]

October 2019

ISSN 0963-8091

For further information please contact:

Joint Nature Conservation Committee
Monkstone House
City Road
Peterborough PE1 1JY
www.jncc.gov.uk

This report should be cited as:

August, T.A., Siriwardena, G. & Isaac, N.J.B. 2019. United Kingdom Terrestrial Evidence Partnership of Partnerships data products: improving opportunities for re-use. *JNCC Report No. 640*, JNCC, Peterborough, ISSN 0963-8091.

EQA:

This report is compliant with the JNCC Evidence Quality Assurance Policy http://jncc.Defra.gov.uk/default.aspx?page=6675.

Affiliation:

[1] NERC Centre for Ecology & Hydrology, Crowmarsh Gifford, Wallingford, OX10 8BB UK
[2] British Trust for Ornithology, The Nunnery, Thetford, Norfolk, IP24 2PU, UK

Summary

The United Kingdom Terrestrial Evidence Partnership of Partnerships (UKTEPoP) schemes produce a large quantity of data, both in terms of raw observations and results from analyses. UKTEPoP schemes produce ‘data products’ that have many commonalities which provides opportunities for analyses across schemes, creating greater value than if each scheme’s data were analysed in isolation. Here we define a data product as an output from an analysis or survey, in a format that can be disseminated. This could be as simple as monthly means or data aggregated to grid squares, up to species’ trend metrics and multispecies indicators. Currently there is a large amount of variation in the way that data products are documented, stored and disseminated between members of UKTEPoP. This variation creates inefficiencies, particularly in the context of cross-scheme analyses. In this report we review the existing data products across the UKTEPoP partnership. We use the FAIR principles of findability, accessibility, interoperability, and reusability to assess the current state of UKTEPoP data products and suggest opportunities for improvements that will lead to an increase in re-use.

Contents

1 Introduction

The United Kingdom Terrestrial Evidence Partnership of Partnerships (UKTEPoP) is a suite of terrestrial surveillance partnership schemes funded or co-funded by JNCC. UKTEPoP spans a suite of structured monitoring schemes collecting abundance data for plants (NPMS), mammals (primarily NBMP), birds (ADS, BBS, GSMP, WeBS) and insects (UKBMS, PoMS – see Glossary for a full list of acronyms). UKTEPoP includes unstructured data from the Rare Breeding Bird Panel (RBBP) and activities of the Biological Records Centre (BRC, via the BReVI partnership) in supporting the activity of more than 80 national recording schemes and societies. For the purposes of this document, BRC is considered as one scheme, not 85. The Terrestrial Surveillance Development & Analysis project (TSDA) has a remit to work across the family of UKTEPoP schemes.

All schemes within UKTEPoP generate large quantities of primary data on the distribution and abundance of terrestrial biodiversity. These data are used for a variety of purposes, but the principal application is to quantify the status and trends of our native species at UK and, where possible, national and regional scales. UKTEPoP data have been influential in describing changes in wildlife populations over recent decades and contribute an enormous quantity of information to species trends (e.g. annual updates of the UK Biodiversity Indicators, English Biodiversity Indicators, Scottish Biodiversity Indicators), international reporting (e.g. EU Habitats Directive, EU Birds Directive, Agreement on the Conservation of African-Eurasian Migratory Waterbirds) and natural capital assessments (e.g. the Office of National Statistics' natural capital accounts).

Historically, data were created, disseminated and used separately by each scheme, reflecting the fact that schemes have largely operated in parallel with one another. In recent years there has been a growing demand for cross-scheme analyses and outputs. For example, the UK Priority Species Indicator for Abundance data (C4a) combines data from at least five UKTEPoP schemes to produce a single headline metric of trends in species' abundance (Eaton *et al.* 2015). The triennial State of Nature Report (Hayhow *et al.* 2016) is broader still, incorporating all the available data from UKTEPoP schemes (including distributional trends from BRC) and data from other sources (e.g. the Countryside Survey and Rothamsted Insect Survey). A large part of the TSDA work programme is predicated on the notion of analyses across schemes and taxonomic groups. Given these developments, there is a clear need to understand how data are produced by schemes in order to effectively deliver cross-scheme analyses and outputs.

A key concept when discussing scheme outputs is the notion of a data product. In this report, we define data products and discuss their advantages in the context of biodiversity indicators and other national-scale analyses. We then review the data products currently produced by the UKTEPoP schemes and describe the challenges presented by the diversity, based on recent experiences working on the Priority Species Indicator. We place UKTEPoP into the global context by describing efforts to construct international data products known as Essential Biodiversity Variables (EBVs). We draw on the FAIR (Findable, Accessible, Interoperable, and Reusable) data principles (Wilkinson *et al.* 2016) to identify opportunities for UKTEPoP to support re-use of its members' data products. Finally, we make some recommendations to achieve greater harmonization and transparency for UKTEPoP data products.

2 Data products

2.1 What are they and why are they useful?

In the context of UKTEPoP, a data product can be defined by two features: 1) an output from an analysis or survey, 2) in a format that can be disseminated, such as a spreadsheet (figure 1). By this definition, data products can include data from a survey formatted for dissemination, a standardised analysis that accounts for the design of the scheme, and intermediate data sets. In this report we will focus on data products that contain the result of an analysis since these are most relevant to the TSDA work program and other cross-scheme use cases. Discussion of raw data as data products can be found elsewhere (e.g. Groom *et al.* 2017; Michener 2015). However, many of the challenges and solutions we discuss apply equally to data products that contain the results of analyses and those that contain data from surveys.

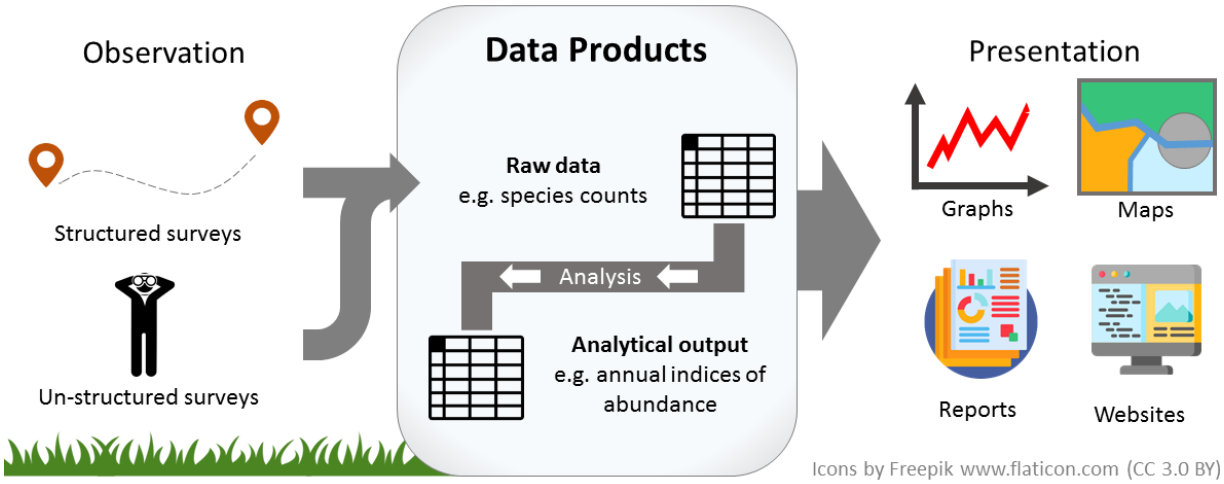

Figure 1. Observers in the field generate raw data, which when stored in a shareable form such as a spreadsheet, are considered a data product. Downstream datasets created through analysis of this data are also data products. Data products can be presented in a variety of formats including graphs and maps.

A good example of a data product would be a table containing annual indices of abundance for each species at the national scale. This is the standard type of input data for biodiversity indicators and the State of Nature Report (Hayhow *et al.* 2016). The data table from the BBS is superficially similar to those from the UKBMS and the NBMP: all contain one number (an index of abundance) for each species in each year. Yet the analytical procedures that produced these data tables differ among schemes, reflecting differences in how the schemes are designed, and the protocols they employ.

Data products of this sort have several types of benefit, both for the individual schemes and the wider community. First, data products provide an effective interface between generic analytical tasks (e.g. collating a biodiversity indicator) and those requiring specialist knowledge of scheme design or species' ecology (e.g. modelling individual species trends). Second, data products create efficiencies, because they can be used for multiple applications. For example, the analytical pipeline to create species' indices of abundance from UKBMS data takes several days to run on a high-performance computing facility, so it's not feasible to go back to the raw data for each and every cross-scheme analysis. Third, data products are visible advertisements for scheme activity and provide a measurable way to track their impact (e.g. through data citation).

Data products are regularly presented as graphs of trends, maps of distributions, densities or movements, and written interpretations of patterns or analytical results. While these presentations are created from data products they fall outside our definition, so are not considered data products themselves. Maps and graphs of temporal or spatial analyses can provide valuable, accessible feedback to surveyors and interpretation of data products for end-users. Examples are the BTO/JNCC/RSPB/WWT Wetland Bird Survey (WeBS) online, the BTO's BirdTrends report (trends and summaries from abundance and demographic data) and online maps from the NBN.

2.2 Use cases for UKTEPoP data products

UKTEPoP data are used for a variety of purposes, and clearly one type of data product will not be sufficient for all of these needs. In this report we focus on one particular use case: the collation of UK biodiversity indicators. For this, the minimum requirement for a data product is a high-level estimate of status (e.g. an index of abundance) for each species in each year. An indicator for any of the nations within the UK would require data products specific to that country.

These national-scale data products would also be suitable for a broad cross-taxon analysis e.g. to test the hypothesis that habitat specialists have declined relative to generalists. However, there are use-cases where information would be required at finer temporal and spatial resolution, and a national scale, biodiversity indicator data product would not be appropriate. For example, assessing the impact of land-use or protected areas might benefit from data products summarised by land-use (or landcover) categories or protected area status, or at the scale of the individual study site.

While the focus here is on the biodiversity indicators use case, many of the challenges and solutions are applicable across a range of different types of data products.

3 Current scheme data products

Recent experience has highlighted two major issues with existing UKTEPoP data products. First, there is considerable variation in the way data products are created, presented and disseminated. This variation creates substantial inefficiencies for cross-scheme analyses. Secondly, much of the variation among scheme products is not adequately documented. These issues highlight a need to catalogue this variation in data products and to explore opportunities for harmonization across UKTEPoP data products.

In this section we describe the national-scale data products for UK biodiversity, with a focus on UKTEPoP schemes. We also describe a number of data products that are not strictly produced by UKTEPoP schemes, but which are closely aligned with the UKTEPoP datasets, or are used in combination with UKTEPoP data products to create new data products (e.g. indicators that cross taxonomic groups).

In reality, few schemes produce data products in a form that is readily accessible and useable by others. Most schemes do produce some kind of trend summary in an annual report, but the data product used to create the various graphs and text summaries are not commonly openly available online.

3.1 UK Butterfly Monitoring Scheme (UKBMS)

UKBMS publishes a suite of data products each year. The two most relevant to this report are the annual measures of abundance (known as collated indices) for each species (Botham *et al.* 2019b) and the long-term trend estimates for each species (Botham *et al.*

2019a). Both are shared via the NERC Environmental Information Data Centre (EIDC)[1] as .csv files with accompanying documentation that details the metadata for the shared data as well as the methods used to create the data.

3.2 Biological Recording, Verification and Interpretation (BReVI) partnership

The Biological Records Centre (BRC) has just published, for the first time, a data product covering 5,293 species of invertebrates, bryophytes and lichens (Outhwaite *et al.* 2019). This data product is openly available online on the EIDC[2]. The dataset includes annual estimates of occupancy (a measure of species geographic range) for each year since 1970, and another dataset of long-term trend estimates. Alongside this data, shared as a series of .csv files, is documentation of the modelling process and metadata describing how the data product was created.

3.3 National Bat Monitoring Programme (NBMP)

NBMP publishes long term trend estimates in its annual report (Bat Conservation Trust 2019). The report includes a series of figures in which the indices of abundance are plotted for each year. These include the index values for each year as well as a smoothed trend with confidence intervals, for each survey method and by country. These presentations are based on a data product generated by the BCT, but which is not shared with the report, or available online.

3.4 Breeding Bird Survey (BBS) and Waterways Breeding Bird Survey (WBBS)

BBS produces species-specific annual indices of relative abundance (with confidence intervals) for common and widespread breeding birds, covering all habitat types. The WBBS produces similar trends for the common and widespread breeding birds, but along waterways specifically. Both of which are presented in graphical form on the individual survey webpages and on the BirdTrends website[3]. These schemes also monitor certain mammal species and data products analogous to those provided for birds are produced wherever possible[4]. BBS and WBBS results are available as graphs of smoothed and unsmoothed trends for all species for which trends are produced at UK, country and regional scales since the surveys began (BBS from 1994 and WBBS from 1998). The data used to plot the graphs are available for download as .csv files with one click from the graphs themselves. Trends are produced wherever supported by the data available. Metadata about these data products and details of survey and analytical methods are also published in the same web platform and are clearly accessible from the BirdTrends data pages.

Further data products are available from the same site in respect of static and customizable tables of population changes. Again, methods are provided via clear links. The static tables present population changes over five, ten or the maximum number of years, as well as lists of species for which alerts of identified population declines over given periods are made (with rationale and methods presented in the Methods text) and for species undergoing population increases. For customizable tables, changes can be extracted for all or a single species for five, ten, 25 or the maximum number of years, with BBS/WBBS results presented

[1] https://catalogue.ceh.ac.uk/documents/571a676f-6c32-489b-b7ec-18dcc617a9f1
[2] https://doi.org/10.5285/0ec7e549-57d4-4e2d-b2d3-2199e1578d84
[3] https://www.bto.org/our-science/publications/birdtrends/2018
[4] https://www.bto.org/our-science/projects/bbs/latest-results/mammal-monitoring

alongside those from other schemes (see below). Tables are presented on the website and can be copied manually for use elsewhere.

Outside of the BirdTrends webpages, supplementary information and data products for BBS and WBBS alone are available on the BBS webpages, consisting of downloadable Microsoft Excel files containing UK trends, national trends and trends for nine English regions[5]. The official statistics are published in an annual BBS Report with summary information and articles covering topics such as population changes as well as fieldwork methods and uses of data. This report covers the BBS bird and mammal monitoring, and the WBBS.

3.5 Other bird data/schemes

The Common Birds Census (CBC) pre-dated the BBS (with overlap 1994-2000) and now functions to provide long-term, joint trend information back to 1966 or 1967 for a subset of BBS species that were sufficiently common in the English lowland farmland and woodland habitats on which CBC was focused to support the calculation of indices. Data products representing the joint trends are presented and are available for download alongside the analogous, shorter-term BBS products from the BirdTrends website[6]. Similar joint trends with WBBS are also produced using the Waterways Bird Survey (WBS), which was the scheme's predecessor from 1975 to 2000, and analogous data products are available for species associated with linear waterways.

Avian Demographic Schemes (ADS) encompass sampling informing about variation in bird breeding success and annual survival in time and space. Among these, the Constant Effort (ringing) Scheme (CES) also provides annual abundance data via standardized catching effort in scrub and reed bed habitats. Similar data products (tables and data underlying trend graphs) are available from the BirdTrends website[4] for this dataset, again with methods and metadata provided via clear links.

Demographic data from ADS are summarized as tables and trends showing UK-level changes since 1967 (data derived from the Nest Record Scheme, NRS), 1984 (data from CES) or species-specific start dates post-2000 (data from Retrapping Adults for Survival, RAS). RAS informs annual survival estimates in twelve species that are not well-covered by general ringing but are sampled by multiple, local, species-specific projects. Annual estimates derived from NRS data of laying date, clutch size, brood size, nest failure rate (i.e. daily probability) at each of the egg and chick stages, and a composite measure of productivity (fledglings per breeding attempt), are reported on the BirdTrends website[7], along with fitted linear or curvilinear trends, as supported by the data. Modelled start and end values for the time series are tabulated and the data underlying trend graphs are downloadable as .csv files, as with BBS. Likewise, trends in productivity shown by the juvenile:adult count ratio from CES are presented in the same location, with the same downloadable products, except that changes are tabulated for five, ten, 25 or the maximum number of years. Annual survival probabilities from CES and RAS are presented graphically, with downloadable data. Metadata and methods information are available for all ADS analyses via clear links on the site.

The Wetland Bird Survey (WeBS) Core Counts scheme prioritises coverage of waterbodies holding substantial waterbird populations, such as estuaries, large lakes, gravel pit complexes and reservoirs, particularly those designated for waterbodies as SSSIs and SPAs, but also includes an unstructured set of smaller inland water bodies, rivers and open coast stretches with variable time series of historical monitoring data. Methods are published

[5] https://www.bto.org/our-science/projects/bbs/latest-results/population-trends
[6] https://www.bto.org/our-science/publications/birdtrends/2018
[7] https://www.bto.org/our-science/publications/birdtrends/2018

and data are presented via an online application[8]. WeBS Report Online presents tables of site-specific annual maxima counts for each waterbird species, and some subspecies, that can be copied from the screen under an Open Government Licence but are not downloadable as separate files. Annual site-specific monthly maximum count data are also available, with tabular presentations that are entirely customizable by species and site and can be filtered to species group, habitat or county. Annual and monthly (averaged across years) indices for the UK and each constituent country are shown graphically as unsmoothed indices and smoothed trends within WeBS Report Online and are also available as a spreadsheet download from the WeBS pages of the BTO website (separately from the WeBS Report Online application). These are modelled from the complete WeBS dataset, and do not support re-generation from a subset such as inland sites alone. Further data products presented in the online report are population changes per site and for groups of sites such as Special Protection Areas, evaluated according to categories for 'alerts' of given net declines over periods of five, 10 or 25 years. WeBS Alerts are produced periodically, with the next release due in autumn 2019. Site totals are given for every WeBS site, calculated as the sum of the annual maxima for each species. The WeBS Low Tide Count scheme aims to survey estuaries at Low Tide every six years. Results are presented within the WeBS Report Online as interactive maps of estuaries with dot density for each species per count sector and a table of maximum and average counts and bird density for the whole site. Results from the Non-estuarine Waterbird Survey, which takes place every nine years to improve coverage of a habitat poorly covered by WeBS Core Counts, are also available as tables of birds counted and county population estimates, and maps of birds counted within the WeBS Report Online[6]. Again, tables are customizable by county, site and species as appropriate, but downloads are not supported.

The Goose and Swan Monitoring Programme (GSMP) has a species focus, considering 13 native goose and migratory swan populations that are found in the UK in winter. Results are presented online[9] as annual count totals. Breeding success data are presented in the form of overall percentages of young and mean brood size, based on observations in winter flocks. Specific site or regional splits are presented as is most informative for each species, rather than being standardized, so graphical and tabular products vary, and numerical data are not provided in an easily downloadable format.

Outside UKTEPoP and the component schemes' direct antecedents, a range of monitoring activities report on scarce and rare breeding birds in the UK, via a combination of bespoke, funded surveys, voluntary initiatives and unstructured recording, and are collated by the Rare Breeding Birds Panel (RBBP), which is led and funded by JNCC on behalf of country conservation agencies, BTO and RSPB. Data for rarer species are presented as county-level records (i.e. with no further details on location), via an online system for records pre-2010 and via an annual report in *British Birds* magazine for more recent years[10]. 'Trends' (more accurately, ratio changes) for each species are downloadable as a Microsoft Excel file, wherein the most recent available estimate of current breeding population size is provided, along with an estimate of percentage change (currently 1990–2015) and metadata on the estimated level of representation of the national population. Where possible, a 25-year change is presented, with a 15-year one for more recently added species. Most trends are derived from data submitted to the RBBP, but trends from bespoke, periodic surveys are used where they are more robust. The source of the data is noted in the downloadable file and full methods are available in the annual RBBP reports. The Heronries Census has monitored Grey Heron (*Ardea cinerea*) in the UK since 1928. Data on changes over five, 10, 25 or the maximum number of years, for the UK and constituent countries separately, are reported as a data product on the BirdTrends webite[5], along with graphs showing annual

[8] https://app.bto.org/webs-reporting/
[9] https://monitoring.wwt.org.uk/our-work/goose-swan-monitoring-programme/species-accounts/
[10] www.rbbp.org.uk

changes and smoothed trends, the data for which are downloadable as .csv files. Field and analytical methods are available via clear links on the site.

Finally, RSPB and country agencies collaborate in a loose programme known as the Statutory Conservation Agencies/RSPB Rare Breeding Birds Survey (SCARRABBS), sometimes also involving BTO, in which periodic surveys of species that are too rare to be covered by national bird monitoring are organised. Some results of these surveys are summarized by the RBBP and are available as detailed above. Paper summaries are produced in the RSPB-led, annual State of the UK's Birds report[11] (no report was produced in 2018) and individual survey results are commonly published in scientific and semi-popular journals. However, there is no central repository or common format for data products from these surveys, if they exist.

3.6 National Plant Monitoring Scheme (NPMS)

NPMS has been running since 2015 and has yet to produce metrics at the national-scale. Recent activities have focussed on developing methods for creating such metrics using Bayesian methods to combine the occupancy and abundance data collected by the scheme (Pescott *et al.* 2019a). The NPMS openly shares plot-level plant occurrence data (Pescott *et al.* 2019b) on the EIDC[12] in a format that is easy to download as .csv files and is shared under an Open Government Licence. Metadata is downloadable from the same location as the data, and contains a detailed description of the data, as well as the methodology used to collect it.

3.7 Pollinator Monitoring Scheme (PoMS)

PoMS comprises two streams of data collection, flower-insect timed counts (FIT counts) and pan-trap surveys have been conducted since 2017. FIT counts are a simple survey that can be undertaken anywhere in dry weather from April to September. This simple survey collects data on the total number of insects that visit a particular flower over a ten-minute period. The PoMS pan-trap surveys are more systematic, undertaken in a set of 75 1km squares randomly allocated to cropped and non-cropped land. The Pollinator Monitoring and Research Partnership (PMRP), which includes PoMS and existing recording schemes that focus on pollinating insects, have been involved in producing the UK pollinator indicator[13]. In the future, it is envisaged that data products will be produced that integrate multiple data streams (i.e. PoMS surveys and observations from recording schemes such as BWARS) into a single composite estimate of species' status and can be produced for the constituent countries of the UK.

3.8 People's Trust for Endangered Species (PTES)

While the People's Trust for Endangered Species (PTES) is not a UKTEPoP scheme, it produces indices for Dormouse and Hedgehogs which are used in the UK Priority Species Indicator C4a[14]. Like indices produced by UKTEPoP members, the indices for dormouse and hedgehog have yearly estimates of abundance and associated error. These indices are used to support assessments of trends in hedgehogs and dormice over time, for example in the State of Britain's Hedgehogs 2018 (Wilson & Wembridge 2018) and the State of Britain's Dormice 2016 (Wembridge *et al.* 2016). However, these data products are not available online and when accessed are in Microsoft Word and Microsoft Excel files which are not accompanied with detailed metadata.

[11] https://www.rspb.org.uk/our-work/conservation/centre-for-conservation-science/state-of-the-uks-birds/
[12] https://doi.org/10.5285/79604721-049b-42ab-a8d7-a9c8b18a193f
[13] https://jncc.gov.uk/our-work/ukbi-d1c-pollinating-insects/
[14] https://jncc.gov.uk/our-work/ukbi2018-c4a-species-abundance/

3.9 Rothamsted Insect Survey (RIS)

The Rothamsted Insect Survey (RIS) has been running since 1964 collecting data on large moths and other insects. The traps use a light to draw in moths at night and the catch is identified and recorded. Since the light traps are standardised and the effort (number of nights the trap is run) is known the abundance data is well suited for long term trend assessments. RIS is not a UKTEPoP scheme but is shares similarities with UKTEPoP schemes such as the UKBMS, which also has fixed sites visited repeatedly and a protocol to ensure constant effort. RIS is used in the UK Priority Species Indicator C4a[10]. RIS moth trap data has recently been reanalysed using the analytical methods and pipelines developed for UKBMS. The resulting data products will soon be made available on EIDC. These data can also be explored graphically on the RIS website, though the data behind each graph cannot be downloaded directly from this location[15]. The aspiration is to update the data on the EIDC and the presentations on the RIS website annually.

4 FAIR data principles

The FAIR principles offer an established framework to reflect upon the challenges UKTEPoP members face when managing and sharing their data products, and to consider the steps UKTEPoP members could take to increase re-use of their data products. The FAIR principles (Wilkinson *et al*. 2016), were created to ensure scientific data is open to analyse in downstream studies, ensuring the maximum value can be gained from existing data products.

Under the FAIR principles data products should be **F**indable, **A**ccessible, **I**nteroperable, and **R**eusable. The authors stress that these principles should apply to both humans and machines, highlighting that given the number of data products available, researchers often use computers to find and retrieve data from the internet.

To be **findable**, data products should have a unique and persistent identifier, should be richly described, and should be registered or indexed in a searchable resource. **Accessible** data products can be easily retrieved using open protocols, and the metadata should be accessible even if the data is removed at some point in the future. To be **interoperable**, data products should use a common terminology to represent the data, this terminology should be based on vocabularies that are themselves FAIR, and metadata should link to other relevant metadata/data. To ensure **reusability** data products need to be richly described, clearly licensed, meet the quality needed for downstream analyses, and have a detailed provenance. We used this framework to identify areas where UKTEPoP data products do not fulfil these principles and suggest a number of approaches to address this shortfall.

5 Challenges and solutions

We use the FAIR principles to identify challenges to findability, accessibility, interoperability and reusability in existing UKTEPoP data products. In each case we identified solutions that could address these challenges. Given the large number of data products produced by UKTEPoP members we focussed on the example of the Priority Species Indicator update 2017 that drew together annual indices of abundance to create a cross-taxonomic indicator. Through this case study, we highlight challenges and solutions that are applicable to a wide range of data products produced by UKTEPoP members and similar schemes.

[15] https://insectsurvey.com/trends

5.1 Findability

The FAIR principle of findability presses upon the creators of data products the importance of making it easy for potential users to find the data they need. Our review of existing data products (section 4) shows that most of the data products created by schemes are already presented in some form. Typically, graphs of the data have been made public, either online or in paper reports, or both. In this way, the data products underpinning these graphs are somewhat findable, but with plenty of room for improvement.

Finding UKTEPoP data products requires prospective users to search different organisational websites and reports. This can be a time-consuming process when undertaking an analysis requiring data products from multiple schemes. Where these data products are available for download, they are rarely associated with a persistent unique identifier (i.e. a URL or Digital Object Identifier (DOI)). Data products that are not associated with a persistent unique identifier can become hard to find in the future, for example if the website cited for a data product change. In order for a user to know that a data product fulfils their needs, the data product must be adequately described. We found that most shared data products are described (in terms of data format and collection/analysis methodology), although the format of this description varies from one data product to the next. When taken together, the findability of data products across schemes does not meet the standard outlined in the FAIR principles, which creates inefficiencies to analyses that seek to use these data products. There is also no standardisation across data products which makes machine-lead data searches impossible at present.

During the update of the Priority Species Indicator in 2017, none of the UKTEPoP annual indices of abundance data products were available on data repositories. Most data products were findable through descriptions in reports and on websites but email correspondence with the relevant schemes was required in most cases in order to access a detailed description of the data products.

To reduce the time taken for users to locate data products, data products can be added to a data repository that uses indexing and search services. Data repositories, such as the EIDC, have been used as standard practice for UKBMS (Botham *et al.* 2019) for several years already (see 3.1), and has now been adopted by the BRC/BReVI (Outhwaite *et al.* 2019) and the Rothamsted Insect Survey. Having these datasets in one location means users can find all the data they need in one place, reducing the time spent searching organisational websites. An extensive list of available data repositories, including the EIDC, can be found at https://www.re3data.org/. By placing data products in data repositories, the data are usually professionally archived and given a DOI. The DOI allows the data product to be cited, giving the original authors credit. This persistent link ensures findability continues into the future, ensuring that the original data is accessible to those who might want to critically assess the analysis, or re-run the analysis using a different methodology. Regardless of where a data product is stored it is important that the data is accompanied by a description that allows users to assess whether the data product is appropriate for their needs. This is usually enforced by data repositories, but where it is not done users will need to invest more time to investigate the data and its appropriateness and could lead to spurious use of data.

5.2 Accessibility

A data set, once found, must be accessible to users in order for it to be used. There are a number of potential barriers to accessibility; can the data product be downloaded, is the format of the data accessible, is the metadata accessible? Amongst the data products produced by UKTEPoP members there is a significant amount of variation in their accessibility. Some data products, such as the NPMS (Pescott *et al.* 2019) data are easily

accessible via the EIDC, while others required email correspondence with scheme organisers to access.

Data products that meet the FAIR definition of accessible should be easy to retrieve via an open and free to use tool (e.g. the Mozilla web browser). At the time of the Priority Species Indicator update in 2017, the majority of UKTEPoP scheme's annual indices of abundance were not accessible via a web browser, though it is clear from our review in section 4 that schemes are increasingly sharing these data products via the web. The raw data used by schemes to undertake analyses and produce new data products such as species indices are not typically freely available. However, downstream data products, such as annual indices of abundance, are typically presented publicly, in the form of graphs in reports and on websites (figure 1). Improving accessibility to UKTEPOP annual indices of abundance is therefore often a case of making data available that could already be accessed by reading figures from graphs that are freely available online.

In the case of the Priority Species Indicator update in 2017, a significant amount of time was spent collating data from UKTEPoP partners over email. Since this collated data is also not accessible, this task would need to be repeated with each update of the Priority Species Indicator. Data products that were only accessible via email correspondence tended to have limited or no metadata associated with them. Consequently, when retrieving data products for the Priority Species Indicator update in 2017 we frequently needed to return to the data providers to gain additional metadata that was not available with the data product (for examples see 5.3). Incomplete metadata also led to the need to re-analyse data sets once assumptions about the data products' structure were found to be incorrect in the light of metadata that had initially been unavailable. Where UKTEPoP schemes made their data products publically accessible this was always via a web interface, allowing free access. Any move away from systems that allow free access, such as using proprietary systems for sharing data products, should be strongly discouraged.

When approached via email for the Priority Species Indicator, UKTEPoP data providers provided access to their data and provided invaluable expertise about the data they curate. However, this requires a significant amount of time from both data providers and data analysts. Publishing data products in a data repository with associated metadata overcomes this accessibility problem by making the data product easily accessible, in perpetuity, via a web browser with little or no additional input required from the data provider. This also makes it easier for the user of a data product to cite the data that they have used, in turn making the data accessible to anyone reading their research. However, publishing data products requires an initial investment from the data provider to format and upload data and accompanying metadata which may or may not be paid off by the reduced time spent handling data requests.

After the data product has been downloaded there can still be barriers to accessibility due to the format of the data. There are many formats in which data products can be stored, some of these are more accessible than others. Using the correct file format when sharing data ensures that data are accessible now and in the future. It is important to use file formats that are human readable, machine readable, and are not controlled by a commercial organisation (i.e. non-proprietary). Datasets used in the Priority Species Indicator analyses were sent in a variety of formats, with approximately 50% using a proprietary format (i.e. Microsoft Word or Excel). These formats cannot reliably be read into statistical programming workflows such as R and require commercial software to read. The other half of data products came in comma-separated files (.csv) which are human readable and machine readable using free and open software, and therefore more accessible. These same principles apply to supporting metadata, which should be made available alongside the data products in open file formats.

Metadata can be presented in verbose reports as is the case for NPMS data[16]. In these cases, it is best to use open file formats (e.g. PDF), rather than proprietary ones (e.g. Microsoft Word). However, metadata that is presented in a report style is typically not machine readable. Machine readable metadata requires a machine-readable file format (e.g. HTML and not PDF), but also common standards for the presentation and definition of metadata (explored further in section 5.3). An extensive discussion of digital file formats, their openness and appropriateness for different data types, curated by the Library of Congress can be found at https://www.loc.gov/preservation/digital/formats/.

5.3 Interoperability

One of the stated ambitions of UKTEPoP is to facilitate "collaboration and knowledge exchange with the anticipated benefits of enhancing and developing efficient and effective joint working"[17]. To realise this vision data products created by UKTEPoP members need to be interoperable. Interoperable data products use a common language so that data and metadata can be consistently interpreted by users and data products can be easily combined. This language should itself be based on a vocabulary that is FAIR. Where appropriate, metadata should link to other relevant data products. Linked datasets make it easier to find additional related data products which can be combined for new analyses or presentations.

Currently the majority of UKTEPoP data products have a low level of interoperability predominantly due to variation in methodologies used to create them that are not adequately documented in metadata. In the case of the Priority Species Indicator update in 2017 a significant investment of time was required to make the data products interoperable. In this case, we were interested in two metrics: species-year indices of abundance, and estimation errors associated with these. Even across just these two metrics there was a large amount of variation in the definitions used which precluded integration of these data products without significant amounts of data manipulation.

The lack of interoperability between UKTEPoP data products used in the Priority Species Indicator update in 2017 is a result of considerable variation in the methods each data provider uses to create its species level index values. Consequently, when using species indices from different schemes the meanings of terms such as 'annual index' are not consistent. For the Priority Species Indicator, most metadata provided with the data products was not sufficient to know the methods used to create the species indices. Consequently, additional clarification was needed from schemes to ensure the interpretation of the data was correct. Some of the key, undocumented, variation between datasets include scale, smoothing, indexing, and representation of errors. We discuss each in turn.

The scale on which the indices were reported varied. The majority of schemes reported on the measurement scale, meaning that the index values were proportional to species' abundance. However, UKBMS reported the species indices on a log scale, so a doubling of species' abundance would not result in a doubling of the index value. Whilst reporting indices on the measurement scale is intuitive, it's not necessarily the obvious choice for schemes such as BBS and NBMP in which the index values are derived from a statistical model on the log scale. Furthermore, the presence of multiple log scales (e.g. $\log_{10}$, $\log_e$) can add to the ambiguity.

In some schemes (BBS, NBMP) the species indices had been smoothed. Smoothing and trend line (e.g. by applying a GAM or spline) can help to reduce the impact of inter-annual variation on the species yearly values but combining taxonomic groups with a mixture of

[16] https://doi.org/10.5285/79604721-049b-42ab-a8d7-a9c8b18a193f
[17] http://archive.jncc.gov.uk/default.aspx?page=7490

smoothed and unsmoothed indices is suboptimal, because the data from different taxa are not directly comparable. It may also be desirable to smooth the indicator after all taxonomic groups have been combined (Soldaat *et al.* 2017), which would result in the smoothing of some trends that had already been smoothed which is undesirable.

Species indices were indexed in the majority of schemes but how this was done varied. Indexing is a process by which all species are set to have the same value in a given year, making the trends in species easier to compare when the absolute abundance of the species might vary considerably. In a third of cases indexing was not used, in another third the first year of the time series (across all species) is used, and in the remaining third species are either indexed to a year other than the first, or the index year is species specific. This situation is further complicated as the value given to species in the index year varies across schemes (one, two, or 100). Indexing results in a loss of information, since indexed data have no units. It's trivial to convert un-indexed into indexed data but impossible to do the reverse, so data products that are not indexed will in general be more interoperable and useable.

Errors on species' yearly values were given in the majority of scheme's data. A quantification of the error on species yearly values is important in order that the uncertainty at the species level can be represented in the final indicator (Soldaat *et al.* 2017). Some schemes did not report errors because the data come from raw counts. In one case, standard errors were reported on a different scale to the species-year estimates, which is a potential source for confusion. Note that species indices are subject to many sources of uncertainty including measurement error (including imperfect detection, i.e. the measured count differs from the true count), sampling error (the study sites don't represent the whole country), model error (shortcomings in how the statistical model describes reality), and statistical error (on the parameters, given the data and the model). Describing which forms of uncertainty are represented in the errors is a further challenge.

This case study of interoperability is likely to reflect a common picture across data products created by UKTEPoP members. Data products are typically created as an output of an analysis that is specifically tailored to the biases and structure of the raw data. Consequently, the outputs of schemes are rarely the same. Overcoming this challenge requires the adoption of a common language for describing these data products, and where possible adopting common standards. Ascribing a standard to data products of a certain type (e.g. yearly indices of species occurrence/abundance), across all data providers, would make them interoperable, since all data providers would use the same definitions and data structure. However, 1) there are no existing standards for biodiversity data products that are the results of analyses (e.g. model outputs), and 2) implementing such a change would require a significant amount of work for those creating and curating data products.

Data standards for raw biodiversity observations data are relatively well established. Biodiversity observations are another form of data product, though not the focus of this report. The Darwin Event core has recently been adopted as a standard for biodiversity observations and is widely used across many systems, making these data interoperable (Wieczorek *et al.* 2012). To date, a standard for biodiversity data products that are the result of analyses of raw data has yet to emerge, but this is likely to change soon. GEO BON[18] has adopted Essential Biodiversity Variables (EBVs) as a key tool for collating, handling and disseminating biodiversity data products (Kissling *et al.* 2015, 2018). EBVs are biological concepts, such as species distribution and abundance (Schmeller *et al.* 2017), which can be described by statistical models. Thus, EBVs are data products that sit between raw observations and biodiversity indicators. The GEO BON working group on Data is now

[18] https://geobon.org/

actively considering the standards for data products to be badged as EBVs, and these could be used to aid interoperability between data products created by UKTEPoP members.

5.4 Reusability

The principle of reusability ensures that data products have sufficient metadata and are of sufficient quality that they can be used again in the future. The data product must be richly described so that future users are able to understand the meaning and structure of the data and make use of the data correctly. The data product must be clearly licensed so that both humans and machines know how the data can be used. The data must be of sufficient quality to be usable in downstream analyses and the data must have a detailed provenance, allowing users to trace back to the analyses that created the data product as well as the input data products to those analyses.

At present the majority of UKTEPoP data products describing annual indices are not richly described. We have covered elements of this in the sections on findability (5.1) and interoperability (5.3). In addition to the types of description required to meet those FAIR principles there are additional descriptive elements particularly important for reusability. These elements describe details that will be important for users to understand if the data is appropriate for their research question. Without access to these descriptions future studies that seek to reuse these data products will be delayed while the required information is found, and there is an increased risk of data being used inappropriately. Descriptions should include information on how the raw data were collected, known biases in the data (e.g. years of intensive species-specific surveys), details on the analyses that have been undertaken to transform the raw data into the data product, and the data structure of the data product (i.e. data fields, their meaning and units). This should include the provenance of the data, for example including persistent unique identifiers (i.e. DOIs) to raw data used in analyses. These descriptions are likely to already exist, in internal and published reports, for many of the UKTEPoP and non-UKTEPoP datasets we have already discussed but may not be formalised or published alongside data products.

Licenses are required in order for users to know what can and what can't be done with a data product. Data products that are shared without a license leave it to the user to make assumptions about permitted uses. In most cases this means that the user must assume the data has a strict license which will reduce the reuse of the data. Licensing a data product does not mean limiting its use, instead it simply describes the allowed use of the data product, which can be an open license allowing unrestricted reuse and sharing. To ensure maximum reusability of data products open licenses should be used. It is the position of JNCC that all biodiversity data generated as a result of JNCC funding should be published openly, using an open license[19]. A full list of licenses that conform to the principles set out in the open definition can be found at http://opendefinition.org/licenses/. Two principal restrictions are permitted in open licenses: 1) 'BY' or attribution requires any user of the data to attribute the data in a derived product (e.g. a publication or report), and 2) 'SA' or share-alike requires users who share the data product to share it under the same license, preventing restrictions to openness being placed on derivatives of the data product. Examples of restrictions in licenses that do not conform to the open definition include: 'ND' or no-derivatives which violates the requirement for re-use in open license and 'NC' or non-commercial which violates the requirement for no discrimination against fields of endeavour. In the case of the non-commercial restriction there is some debate around how this is interpreted legally (Creative Commons 2009; Hagedorn *et al.* 2011). In the narrowest sense it can be taken to mean that the data product cannot be use for profit, for example through resale, which is often the intent of the person applying the license. However, in a broad sense it could also be interpreted to mean that the data product could not be served from a

[19] https://hub.jncc.gov.uk/assets/d6381e39-baa4-4f12-93d7-fa16dd3600b8

website that contains advertising. It is therefore advisable to carefully consider potential adverse consequences of applying this restriction.

6 Recommendations

We have reviewed the existing data products produced by members of UKTEPoP and, using the FAIR principles as a framework, we have explored the challenges to re-use and considered solutions. Here we highlight recommendations which we believe have the greatest chance to increase the re-use of data products while still being achievable. We suggest:

1) Annual publication of data products arising from UKTEPoP schemes.
2) Publication of data products using machine readable open file formats (e.g. .csv).
3) Publication of data products with licenses that conform to the open definition[20] wherever possible (e.g. OGL, CC-0, CC-BY, and CC-BY-SA).
4) Publication of data products with persistent unique identifiers (i.e. DOIs).
5) Metadata should be made available with data products using an open file format, preferably one that is machine readable. This should include definitions of terms, data structure, provenance, analytical methods used to create the data product, and links to related data products (such as the raw data used in an analysis).
6) Publication of data products in searchable, indexed, data repositories of similar data products.

In addition to these recommendations, we also conclude that some solutions to the issue of re-use require additional development before they can be applied:

7) UKTEPoP members should reflect on recommendations 1-6 and agree to minimum expectations for data publication, metadata quality, and openness, for data products arising from UKTEPoP schemes.
8) UKTEPoP members should move towards a common language to describe data products, preferably by adopting EBV standards as they evolve.

7 Glossary of terms and abbreviations

BBS:	Breeding Bird Survey
BRC:	Biological Records Centre
BReVI:	Biological Recording, Verification and Interpretation
BTO:	British Trust for Ornithology
CEH:	Centre for Ecology & Hydrology
Defra:	Department for the Environment, Food and Rural Affairs
DOI:	Digital Object Identifier
EIDC:	Environmental Information Data Centre
GEO BON:	Group on Earth Observations Biodiversity Observation Network
GSMP:	Goose and Swan Monitoring Programme
JNCC:	Joint Nature Conservation Committee
NERC:	Natural Environment Research Council
NBMP:	National Bat Monitoring Programme
NPMS:	National Plant Monitoring Scheme
PoMS:	Pollinator Monitoring Scheme
PMRP:	Pollinator Monitoring and Research Group
RBBP:	Rare Breeding Birds Panel
RSPB:	Royal Society for the Protection of Birds

[20] https://opendefinition.org/od/2.1/en/

SCARRABBS: Statutory Conservation Agencies/RSPB Rare Breeding Birds Survey
SPA: Special Protection Area
SSSI: Site of Special Scientific Interest
TSDA: Terrestrial Surveillance Development & Analysis
UKBMS: UK Butterfly Monitoring Scheme
UKTEPoP: United Kingdom Terrestrial Evidence Partnership of Partnerships
WBBS: Waterways Breeding Bird Survey
WeBS: Wetland Bird Survey

8 References

Bat Conservation Trust. 2019. The National Bat Monitoring Programme Annual Report 2018. Bat Conservation Trust, London. http://www.bats.org.uk/pages/nbmp_annual_report.html.

Botham, M.S., Brereton, T., Harris, S., Harrower, C., Middlebrook, I., Randle, Z. & Roy, D.B. 2019a. United Kingdom Butterfly Monitoring Scheme: species trends 2017. NERC Environmental Information Data Centre. https://doi.org/10.5285/ddcd1fa7-27d8-4082-93bd-a550c98208cf.

Botham, M.S., Brereton, T., Harris, S., Harrower, C., Middlebrook, I., Randle, Z. & Roy, D.B. 2019b. United Kingdom Butterfly Monitoring Scheme: collated indices 2017. NERC Environmental Information Data Centre. https://doi.org/10.5285/ace3c3ef-df89-40b9-ba8b-106997fd6d9c.

Creative Commons. 2009. Defining 'Noncommercial': A study of how the online population understands "Noncommercial Use". https://wiki.creativecommons.org/wiki/Defining_Noncommercial.

Eaton, M.A., Burns, F., Isaac, N.J.B., Gregory, R.D., August, T.A., Barlow, K.E., Brereton, T. Brooks, D.R., Fulaij, N.A., Haysom, K.A., Nobel, D.G., Outhwaite, C., Powney, G.P., Procter, D. & Williams, J. 2015. The Priority Species Indicator: Measuring the trends in threatened species in the UK. Biodiversity, **16**, 108–19, https://doi.org/10.1080/14888386.2015.1068222.

Groom, Q., Weatherdon, L. & Geijzendorffer, I.R. 2017. Is citizen science an open science in the case of biodiversity observations? Edited by Marc Cadotte. *Journal of Applied Ecology* 54 (2): 612–17. https://doi.org/10.1111/1365-2664.12767.

Hagedorn, G., Mietchen, D., Morris, R.A., Agosti, D., Penev, L., Berendsohn, W.G. & Hobern, D. 2011. Creative Commons Licenses and the non-commercial condition: Implications for the re-use of biodiversity information. ZooKeys **149**, 127–49, https://doi.org/10.3897/zookeys.150.2189.

Hayhow, D.B., Burns, F., Eaton, M.A., Fulaij, N.A.I., August, T.A., Babey, L., Bacon, L., Bingham, C., Boswell, J., Boughey, K.L., Brereton, T., Brookman, E., Brooks, D.R., Bullock, D.J., Burke, O., Collis, M., Corbet, L., Cornish, N., De Massimi, S., Densham, J., Dunn, E., Elliott, S., Gent, T., Godber, J., Hamilton, S., Havery, S., Hawkins, S., Henney, J., Holmes, K., Hutchinson, N., Isaac, N.J.B., Johns, D., Macadam, C.R., Mathews, F., Nicolet, P., Noble, D.G., Outhwaite, C.L., Powney, G.D., Richardson, P., Roy, D.B., Sims, D., Smart, S., Stevenson, K., Stroud, R.A., Walker, K.J., Webb, J.R., Webb, T.J., Wynde, R. & Gregory, R.D. 2016. State of Nature 2016. The State of Nature Partnership.

Kissling, W.D., Ahumada, J.A., Bowser, A., Fernandez, M., Fernández, N., García, E.A., Guralnick, R.P., Isaac, N.J.B., Kelling, S., Los, W., McRae, L., Mihoub, J.-B., Obst, M.,

Santamaria, M., Skidmore, A.K., Williams, K.J., Agosti, D., Amariles, D., Arvanitidis, C., Bastin, L., De Leo, F., Egloff, W., Elith, J., Hobern, D., Martin, D., Pereira, H.M., Pesole, G., Peterseil, J., Saarenmaa, H., Schigel, D., Schmeller, D.S., Segata, N., Turak, E., Uhlir, P.F., Wee, B. & Hardisty, A.R. 2018. Building Essential Biodiversity Variables (EBVs) of species distribution and abundance at a global scale. Biological Reviews, **93**, 600–625, https://doi.org/10.1111/brv.12359.

Kissling, W.D., Hardisty, A., García, E.A., Santamaria, M., De Leo, F., Pesole, G., Freyhof, J., Manset, D., Wissel, S., Konijn, J. & Los, W. 2015. Towards global interoperability for supporting biodiversity research on Essential Biodiversity Variables (EBVs). Biodiversity, **16**, 99–107, https://doi.org/10.1080/14888386.2015.1068709.

Michener, W.K., 2015. Ecological Data Sharing. Ecological Informatics, **29**, 33–44, https://doi.org/10.1016/j.ecoinf.2015.06.010.

Outhwaite, C.L., Powney, G.D., August, T.A., Chandler, R.E., Rorke, S., Pescott, O., Harvey, M., Hepper, D., Hubble, D., Kramer, J., Lee, P., MacAdam, C., Morris, R., Norris, A., Palmer, S., Plant, C., Simkin, J., Stubbs, A., Sutton, P., Telfer, M., Wallace, I. & Isaac, N.J.B, 2019. Annual estimates of occupancy for bryophytes, lichens and invertebrates in the UK (1970-2015). NERC Environmental Information Data Centre. https://doi.org/10.5285/0ec7e549-57d4-4e2d-b2d3-2199e1578d84.

Pescott, O.L., Powney, G.P. & Walker, K.J. 2019. Developing a Bayesian species occupancy/abundance indicator for the UK National Plant Monitoring Scheme. Wallingford, NERC/Centre for Ecology & Hydrology and BSBI. https://doi.org/10.13140/RG.2.2.23795.48161.

Pescott, O.L., Walker, K., Day, J., Harris, F. & Roy, D.B. 2019. National Plant Monitoring Scheme survey data (2015-2018). NERC Environmental Information Data Centre. https://doi.org/10.5285/79604721-049b-42ab-a8d7-a9c8b18a193f.

Schmeller, D.S., Mihoub, J.B., Bowser, A., Arvanitidis, C., Costello, M.J., Fernandez, M., Geller, G.N., Hobern, D., Kissling, W.D., Regan, E., Saarenmaa, H., Turak, E. & Isaac, N.J.B. 2017. An operational definition of Essential Biodiversity Variables. Biodiversity and Conservation, **26**, 2967–2972, https://doi.org/10.1007/s10531-017-1386-9.

Soldaat, L.L., Pannekoek, J., Verweij, R.J.T., van Turnhout, C.A.M. & van Strien, A.J. 2017. A Monte Carlo method to account for sampling error in multi-species indicators. Ecological Indicators, **81**, 340–347, https://doi.org/10.1016/j.ecolind.2017.05.033.

Wembridge, D., Al-Fulaij, N. & Langton, S. 2016. The state of Britain's dormice 2016. People's Trust for Endangered Species. https://ptes.org/wp-content/uploads/2016/09/State-of-Britains-Dormice-2016.pdf.

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

Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.W., Bonino da Silva Santos, L., Bourne, P.E., Bouwman, J., Brookes, A.J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C.T., Finkers, R., Gonzalez-Beltran, A., Gray, A.J.G., Groth, P., Goble, C., Grethe, J.S., Heringa, J., 't Hoen, P.A.C., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S.J., Martone, M.R., Mons, A., Packer, A.L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M.A., Thompson, M., van der Lei,

J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J. & Mons, B. 2016. Comment: The FAIR guiding principles for scientific data management and stewardship. Scientific Data, **3**, 160018, https://doi.org/10.1038/sdata.2016.18.

Wilson, E. & Wembridge, D. 2018. The state of Britain's hedgehogs 2018. British Hedgehog Preservation Society and the People's Trust for Endangered Species. https://www.britishhedgehogs.org.uk/state-britains-hedgehogs-2018/.