

# An operational workflow for producing periodic estimates of species occupancy at national scales

Robin J. Boyd\* , Thomas A. August, Robert Cooke, Mark Logie, Francesca Mancini, Gary D. Powney, David B. Roy, Katharine Turvey and Nick J. B. Isaac

*UK Centre for Ecology and Hydrology, MacLean Building, Benson Lane, Crowmarsh Gifford, Wallingford, OX10 8BB, UK*

## ABSTRACT

Policy makers require high-level summaries of biodiversity change. However, deriving such summaries from raw biodiversity data is a complex process involving several intermediary stages. In this paper, we describe an operational workflow for generating annual estimates of species occupancy at national scales from raw species occurrence data, which can be used to construct a range of policy-relevant biodiversity indicators. We describe the workflow in detail: from data acquisition, data assessment and data manipulation, through modelling, model evaluation, application and dissemination. At each stage, we draw on our experience developing and applying the workflow for almost a decade to outline the challenges that analysts might face. These challenges span many areas of ecology, taxonomy, data science, computing and statistics. In our case, the principal output of the workflow is annual estimates of occupancy, with measures of uncertainty, for over 5000 species in each of several defined ‘regions’ (e.g. countries, protected areas, etc.) of the UK from 1970 to 2019. This data product corresponds closely to the notion of a species distribution Essential Biodiversity Variable (EBV). Throughout the paper, we highlight methodologies that might not be applicable outside of the UK and suggest alternatives. We also highlight areas where the workflow can be improved; in particular, methods are needed to mitigate and communicate the risk of bias arising from the lack of representativeness that is typical of biodiversity data. Finally, we revisit the ‘ideal’ and ‘minimal’ criteria for species distribution EBVs laid out in previous contributions and pose some outstanding questions that should be addressed as a matter of priority. Going forward, we hope that this paper acts as a template for research groups around the world seeking to develop similar data products.

*Key words:* biodiversity, essential biodiversity variable, citizen science, occupancy model, species distributions.

## CONTENTS

I. Introduction	1493
II. Species–space–time cubes	1494
III. The workflow	1495
(1) Raw data acquisition	1495
(2) Risk of bias assessment	1497
(3) Data manipulation	1498
(4) Modelling	1499
(a) State sub-model	1499
(b) Detection sub-model	1499
(c) Model fitting	1500
(5) Model evaluation	1500
(6) Populating the species–space–time cube	1501

\* Author for correspondence (E-mail: [robboy@ceh.ac.uk](mailto:robboy@ceh.ac.uk)).

(7) Applications	1501
(a) Species trends	1501
(b) Multispecies indicators	1501
(c) Comparing trends	1501
(d) Functional diversity	1502
(8) Dissemination	1502
IV. Implementing the workflow	1503
V. Computer infrastructure	1503
VI. Metadata	1503
VII. Generalisability of the workflow to other taxa and geographic locations	1504
VIII. ‘Ideal’ versus ‘minimal’ requirements for species distribution EBVs	1504
IX. Outstanding questions	1505
(1) Can we statistically correct for a wider range of biases in the species occurrence data?	1505
(2) Is the occupancy-detection model appropriate where repeat visits are few?	1505
(3) How do we evaluate model adequacy?	1505
(4) What are the optimal species inclusion criteria and are they generalisable?	1505
(5) Is the one-size-fits-all approach appropriate?	1505
X. Conclusions	1505
XI. Acknowledgements	1505
XII. References	1506
XIII. Supporting information	1508

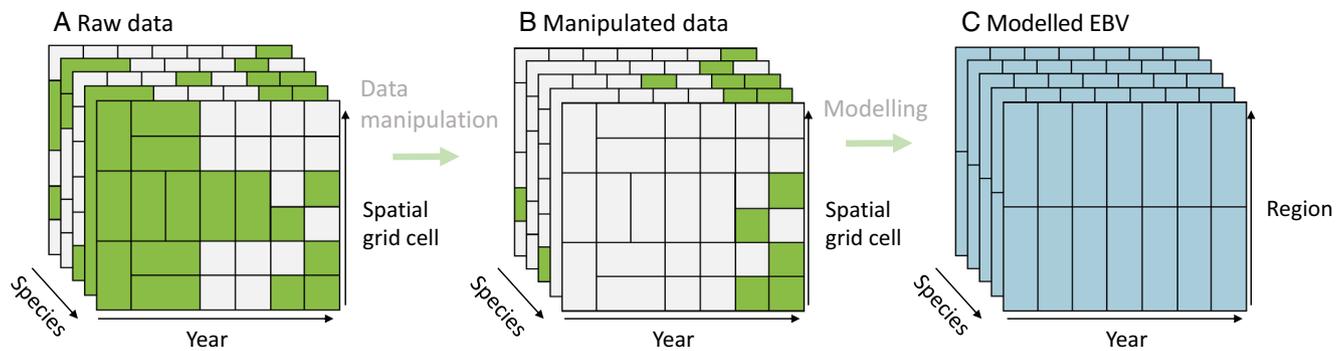
## I. INTRODUCTION

Information on the status of biodiversity and trends thereof is needed to monitor progress towards biodiversity targets and evaluate the effectiveness of conservation action. The rudiments of this information are primary (raw) data, but policy makers require high-level summaries such as indicators. The route from raw data to biodiversity indicator is not straightforward because the data typically derive from disparate sources and are heterogeneous in terms of sampling protocol, extent and resolution (grain size). To bridge this gap, the Group on Earth Observations Biodiversity Observation Network (GEO BON) conceptualised Essential Biodiversity Variables (EBVs; Pereira *et al.*, 2013) as intermediary products that synthesise the available information in a common spatial, temporal and taxonomic framework. Several categories of EBV have been characterised to summarise the major dimensions of biodiversity and biodiversity change: genetic composition, Species populations (abundance or distribution), Species traits, Community composition, Ecosystem structure and Ecosystem function (Pereira *et al.*, 2013). Together, these EBVs form a key component of a global information infrastructure for biodiversity (Peterson & Soberón, 2018). For example, EBV-type data products underpin multinational biodiversity syntheses, such as the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Service’s (IPBES) Global Assessment, the Global Biodiversity Outlook and the Biodiversity Indicators Partnership dashboard and are increasingly being used at national and local levels (Vihervaara *et al.*, 2017).

Species population EBVs characterise species’ populations along the axes of taxonomy, space and time (Jetz *et al.*, 2019; Kissling *et al.*, 2018). One way to view species population EBVs is as three-dimensional grids in which each cell denotes the status of some species’ population in some spatio-temporal

unit – the species–space–time cube (Fig. 1; Jetz *et al.*, 2019; Kissling *et al.*, 2018; Schmeller *et al.*, 2017). Within each cell, population status may be quantified using one of two state variables: abundance, i.e. an index of the number of individuals present; or occurrence, i.e. whether at least one individual is present (or the probability thereof). The choice of state variable determines the specific category of EBV; i.e. the species abundance or species distribution EBV, respectively. Abundance is often the preferred measure of species’ population status, but data on species’ abundances are expensive and complicated to collect. Hence, for most taxa, places, and time periods – and therefore most cells in the species–space–time cube – occurrence is the only feasible measure of species’ populations.

Populating the species–space–time cube with information on species’ occurrences requires data and models. Structured monitoring data are the gold standard but are not available for most taxa in most parts of the world. Instead, analysts rely on unstructured, presence-only data of the types held in biological collections or collected through citizen science initiatives. These data are available for more cells in the species–space–time cube than structured data; however, they are not available for all cells, and those cells with data may be subject to measurement error (e.g. false absences; Franklin, 1999). Hence, modelling is required to infer information on species’ occurrences in cells with no data, and to correct for measurement error in those cells for which data are available. Several types of model might be considered: correlative habitat suitability models (Amini Tehrani, Naimi & Jaboyedoff, 2021); deductive habitat suitability models, which are based on expert advice about habitat associations (e.g. <https://mol.org/indicators/habitat>; Jetz, McPherson & Guralnick, 2012); or models with a temporal component that estimate changes in species’ distributions (e.g. Outhwaite *et al.*, 2020).



**Fig. 1.** Progression of the species–space–time cube through various stages in the workflow. Grey cells indicate a lack of information, green cells indicate that data is available and blue cells indicate that information on species’ occupancy has been inferred through statistical modelling. Cube A represents the raw data, which is the Essential Biodiversity Variable (EBV)-usable data set (Kissling *et al.*, 2018). Note that data are available for many cells, but that the cells vary in size, denoting variable spatial and temporal resolutions. Cube B represents the EBV-ready data set (*sensu* Kissling *et al.*, 2018), which is obtained after the data manipulation stage. At this step, spatially and temporally imprecise data have been removed, which is reflected by a common cell size, but also by the fact that fewer cells are populated. Cube C represents the derived and modelled EBV data set (*sensu* Kissling *et al.*, 2018). We use occupancy-detection models to infer information on species’ occupancy in every sampled cell in B, then calculate the proportion of those cells that are occupied in each ‘region’ (e.g. country within the UK).

The majority of the first wave of species distribution EBVs were derived using correlative or deductive habitat suitability models, whose primary purpose is to *map* rather than *monitor* species’ distributions. For example, Amini Tehrani *et al.* (2021) constructed an EBV for 14 species of bird in Switzerland using an ensemble of habitat suitability models. At the time of writing, there are two species distribution EBVs available on the GEO BON EBV data portal (<https://portal.geobon.org/home>), both of which were derived using habitat suitability models. The *Map of Life* web platform (Jetz *et al.*, 2012) holds deductive habitat suitability models for over 6000 vertebrate species. Whilst it is possible to extrapolate habitat suitability to new time periods, thereby populating the temporal component of the species–space–time cube, this requires strong assumptions about the stationarity of species–environment relationships (Damgaard, 2019).

Moving beyond the choice of data and model, there is a growing literature on the multitude of steps required to create, evaluate and disseminate species distribution EBVs and derivatives such as biodiversity indicators. Kissling *et al.* (2018) and Jetz *et al.* (2019) proposed high-level workflows for developing species population EBVs (distribution and abundance). Hardisty *et al.* (2019) produced the ‘Bari Manifesto’ comprising 10 principles for producing interoperable EBVs of all categories. Rapacciuolo, Young & Johnson (2021) proposed four general steps for mitigating the unstructured nature of community-contributed (or citizen science) data and using them to create indicators. These contributions provide a conceptual framework for constructing species distribution and other EBVs. However, as noted by Fernández *et al.* (2020, p. 491), ‘At present, fully operational workflows that facilitate the automated and widespread production of EBVs are missing’.

Here, building on the generic frameworks cited above, we describe an operational workflow for producing periodic

estimates of species occupancy at national scales. By operational, we mean that the workflow is implemented, and produces outputs (e.g. national biodiversity indicators), on a regular basis. Measuring change over our time is our main focus so the workflow is built on a temporally explicit occupancy detection model (Royle *et al.*, 2007).

We describe each of eight steps in the workflow sequentially. For each step, we review the implementation challenges, describe the methods employed in our specific realisation of the workflow in the UK and provide links to relevant R packages and tools that can be used to replicate or adapt them. Having described each step in the workflow, we then outline some considerations for implementing the workflow on a regular basis to update the EBV and biodiversity indicators. Acknowledging that the UK is among the best-sampled countries globally, we describe options that might be suitable internationally. Finally, we revisit the ‘ideal’ requirements for species distribution EBVs laid out by Kissling *et al.* (2018) in the light of our experiences.

## II. SPECIES–SPACE–TIME CUBES

The workflow described in this paper produces a series of data sets, each of which can be represented as a species–space–time cube (Fig. 1). The first, the ‘EBV-usable’ data set (Kissling *et al.*, 2018), is populated with the available species occurrences records. The spatial, temporal and taxonomic resolutions and extents of these data, and therefore of the EBV-usable data set, are not fixed. The second species–space–time cube, the ‘EBV-ready’ data set, is more standardised. At this stage, the data have been harmonised in a common spatial, temporal and taxonomic framework. The final data cube, the ‘derived and modelled’ EBV data,

contains estimates of species occurrence (or summaries thereof) derived from a statistical model. The resolutions of this final data cube may or may not be the same as the previous one. For clarity, we list the spatial, temporal and taxonomic resolutions and extents of the three species–space–time cubes that we produce for the UK in Table 1.

### III. THE WORKFLOW

For simplicity, we present the workflow as a linear process with eight steps (Fig. 2; Table 2). In practice, some steps are iterative and some components could reasonably be placed in more than one step. We highlight iterative components of the workflow throughout this paper and indicate some of these in Fig. 2 using black arrows.

#### (1) Raw data acquisition

The first task when constructing a species distribution EBV is to obtain reliable data on species' occurrences for as many cells in the species–space–time cube as possible. Many data types might be considered: preserved specimens from museums and herbaria (Jönsson, Broad & Umner, 2021), observational data documenting sightings of some taxon (Sullivan *et al.*, 2014) and more modern forms of monitoring such as passive (e.g. acoustic) sensors and environmental DNA (eDNA) (August *et al.*, 2015), amongst others. These data types have different properties, which has important implications for how they are treated downstream.

Data sources vary in terms of their reliability. For example, records from preserved specimens are generally reliable in terms of taxonomic identity but lack precise information on where and when they were collected. On the other hand, community-contributed data (e.g. from eBird) often come with precise information on where and when they were

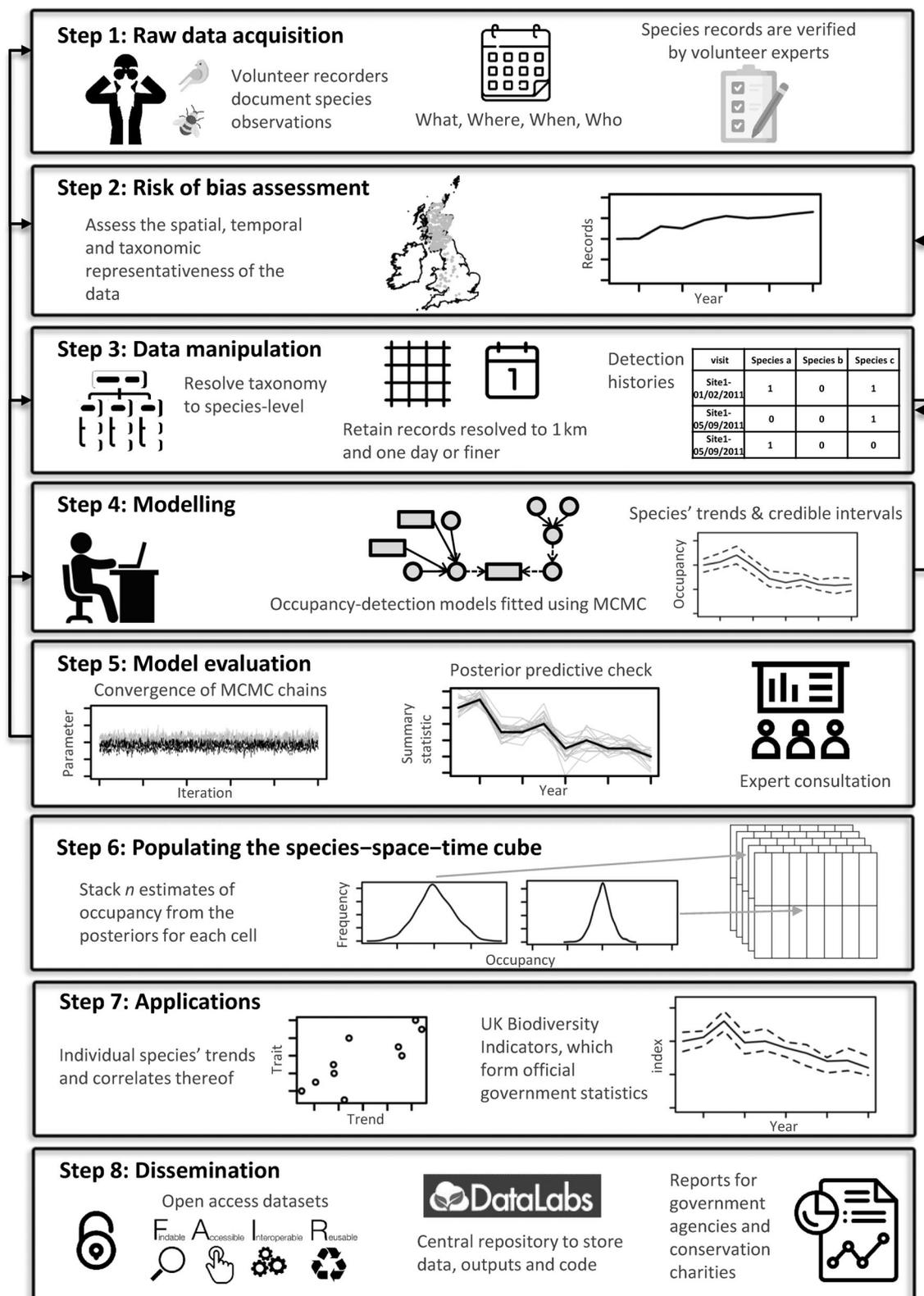
collected, but are more likely to contain misidentifications or to be identified at a coarse taxon level (e.g. species aggregates or genera). Many data providers have procedures to identify dubious records: the Global Biodiversity Information Facility (GBIF) – a global data aggregator – flags records with various spatial, temporal and taxonomic issues; eBird (Sullivan *et al.*, 2014) flags ‘unusual’ records which are then reviewed by regional experts; and iNaturalist designates only those records which have been photographed and accepted by the community as ‘research grade’. Software has also been developed to identify dubious records in species occurrence data sets (Zizka *et al.*, 2019). Dubious and coarse records may be removed downstream (see Section III.3).

In addition to reliability, data sets will vary in terms of supporting metadata. Many repositories, including GBIF and the National Biodiversity Network (NBN) Atlas, have adopted the DarwinCore metadata standard (<https://dwc.tdwg.org/>) for museum specimens and occurrence records. DarwinCore provides a standardised way to store and share ancillary information about species occurrence records. Extensions, such as the Event Core (Wieczorek *et al.*, 2014) and Humbolt Core (Guralnick, Walls & Jetz, 2018), have been developed to capture the sampling protocol, survey effort and other relevant information.

In our realisation of the workflow, we use observational species occurrence data. These data comprise information on the four ‘Ws’ of biological recording: What was seen, Where, When and by Whom (Isaac & Pocock, 2015). Whilst providing the same information (the four ‘Ws’), the data derive from disparate sources such as structured surveys, atlas projects and mass participation projects aiming to engage audiences with a range of expertise. Hence, they comprise a mix of ‘opportunistic’ records, checklists and inventories as well as structured monitoring with a defined protocol and repeated sampling of the same location among years (Pocock *et al.*, 2015).

Table 1. Spatial, temporal and taxonomic extents of the three species–space–time cubes that we produce using the workflow in the UK. In the column on the right, text in italics indicates that the resolution/extent of the derived and modelled EBV meets the minimal criteria as defined by Kissling *et al.* (2018), and bold text indicates that it meets the ‘ideal’ requirement. Kissling *et al.* (2018) proposed additional criteria, but we do not include them in this table because they are not relevant to the extents and resolutions of the EBVs. See Section VIII for more on these criteria.

Dimension	EBV usable dataset	EBV ready dataset	Derived and modelled EBV
Spatial extent	Sampled locations in the UK	Sampled 1 km grid squares in the UK	<i>UK wide</i>
Spatial resolution	Variable (1 m to e.g. 100 km)	1 km	<i>Regions: typically countries (e.g. Scotland, England) but may be more finely resolved</i>
Temporal extent	All years	Since 1970	<b>Since 1970</b>
Temporal resolution	Variable (depending on the data)	Visits on a particular day nested within years	<b>Annual</b>
Taxonomic extent	Poorly monitored taxon groups (one data set per group)	Poorly monitored taxon groups (one data set per group)	<i>Poorly monitored taxon groups (combined data set for all groups)</i>
Taxonomic resolution	Variable (including subspecies and genera)	Species and a few species aggregates	<i>Species and a few species aggregates</i>



**Fig. 2.** Schematic representation of the workflow as applied in the UK. In this case study, the raw data are biological records provided by taxon-specific schemes and societies, and the downstream products include regional and national indicators. The black arrows indicate that some steps in the workflow are iterative; i.e. in completing a downstream step it might become clear that they need to be revisited. Potentially iterative components of the workflow are highlighted throughout the main text. Icons from Flaticon. MCMC, Markov Chain Monte Carlo methods used to fit the occupancy-detection models.

Table 2. Examples of the methods that we use at each step in our realisation of the workflow and links to the R code used to implement them, as well as alternatives that are available. Text in italics refers to the names of R packages and functions within those packages after the double colon. We do not include prospective methods, which we are yet to implement. All R packages listed under 'Specific methods' are openly available on the UK Biological Records Centre's github page (<https://github.com/BiologicalRecordsCentre/BRC>), with the exception of *occAssess* which is maintained at <https://github.com/robboyd/occAssess>. Each package is extensively documented, unit tested and includes vignettes demonstrating how to use it. EIDC, Environmental Information Data Centre; GBIF, Global Biodiversity Information Facility; MCMC, Markov Chain Monte Carlo methods; NBN, National Biodiversity Network; ROBITT is a tool in ecology for assessing the Risk of Bias in studies of Temporal Trends.

Workflow step	General approach	Specific methods	Alternative methods/code
Raw data acquisition	Sourced from NBN Atlas or from taxon recording schemes	—	Retrieve from GBIF via <i>rgbif</i> (Chamberlain <i>et al.</i> , 2021)
Risk of bias assessment	Bias heuristics and ROBITT	ROBITT (e.g. Boyd & Turvey, 2023) and <i>occAssess</i> (Boyd <i>et al.</i> , 2021)	<i>sampbias</i> (Zizka <i>et al.</i> , 2021) or ignorance maps (Ruete, 2015)
Data manipulation	Data formatting, filtering and harmonisation	<i>sparta::formatOccData</i>	<i>CoordinateCleaner</i> (Zizka <i>et al.</i> , 2019) and <i>taxize</i> (Chamberlain & Szöcs, 2013)
Modelling	Occupancy-detection model	<i>sparta::occDetFunc</i>	Frescalo (Hill, 2012); <i>spOcc</i> (Doser <i>et al.</i> , 2022)
Model evaluation	Convergence of MCMC chains; posterior predictive checks; precision of estimates	<i>sparta</i> (August <i>et al.</i> , 2020b)	<i>spOccupancy::ppcOcc</i> , <i>PresenceAbsence</i> (Freeman & Moisen, 2008) for spatial metrics; Wright <i>et al.</i> (2019) for residual plots
Populating the species–space–time cube	Occupancy per species/year/region, with uncertainty	<i>wrappeR::applySamp</i>	Collate point estimates, measures of uncertainty and model fit for each cell
Applications	Multispecies indicators	<i>BRCindicators::bma</i> (August <i>et al.</i> , 2022)	MSI tool (Soldaat <i>et al.</i> , 2017)
Dissemination	Data products on EIDC; indicators published under Open Government License; DataLabs (Hollaway <i>et al.</i> , 2020) for internal dissemination	<i>wrappeR::summariseMSI</i> (Boyd <i>et al.</i> , 2022b)	Deposit EBVs on GEO BON portal ( <a href="https://portal.geobon.org/home">https://portal.geobon.org/home</a> )

In the UK, we are fortunate in that biological recording has a wide taxonomic coverage: there are more than 80 schemes and societies each focussing on the compilation and review of records for a taxonomic group of interest (Baker *et al.*, 2021; Pocock *et al.*, 2015). The primary motivation for the development of the workflow described here was to enable reporting of species groups that are not subject to structured monitoring, so our main focus is on bryophytes, lichens and ~20 invertebrate groups. Through collaboration with the relevant schemes, we have access to >24 million records for >10,000 species in these taxon groups (noting that many species are removed downstream; see Section III.3). The records are verified by taxon experts, who employ automated checks – e.g. the NBN record cleaner used by the NBN Atlas (<https://nbnatlas.org/>) – to screen for spurious records, consider the method of identification (e.g. use of field guides or dichotomous keys), the presence of a supporting image or specimen and a range of additional online supporting information. Given that each scheme is taxon focussed and curates its own data, it makes sense to treat each schemes' data as separate, with distinct properties. Treating the data sets in this way has several advantages, which we describe throughout this paper.

## (2) Risk of bias assessment

Constructing species distribution EBVs and biodiversity indicators is a matter of statistical inference. The analyst intends to infer something about species' distributions – say, a trend in the proportion of sites occupied by some species (occupancy) – using information contained in the species–space–time cube. This task is difficult because data are usually available for a sample of cells only (Meyer *et al.*, 2015). Exactly how difficult depends largely on whether the sample is *representative* of the species–space–time cube as a whole.

Sample representativeness is usually defined in terms of whether the variable of interest (here species occurrence) is similar in sampled to non-sampled population units (cells in the species–space–time cube; e.g. Meng, 2018). Ascertaining whether a data set is representative in this sense is challenging, however, because information is not available for non-sampled cells. Instead, it is common to assess whether the available data are representative of the three dimensions of the cube: taxonomy, space and time.

Tools exist to assess whether a data set is representative of the three dimensions of the species–space–time cube. Oliver *et al.* (2021) developed indicators that track temporal trends

in geographic coverage and sampling effectiveness at national or global levels. Ruete (2015) proposed ‘ignorance maps’, which identify locations that have been poorly sampled. R packages, too, have been developed for bias assessment (Boyd *et al.*, 2021; Zizka, Antonelli & Silvestro, 2021).

Such tools arguably become most informative when used within a qualitative ‘Risk of Bias’ (RoB) assessment. RoB assessments are routine in medicine and other disciplines, but it was only recently that the first RoB tool for biodiversity science, ROBITT (risk of bias in studies of temporal trends in ecology), was developed (Boyd *et al.*, 2022*d*). ROBITT comprises a series of ‘signalling’ questions, mostly about the potential for issues of representativeness to bias estimates of biodiversity time trends. Some ask about the potential for other issues, such as variation in detection probabilities, to cause biases. The remainder ask the user to describe steps that will be taken to mitigate the risk of bias, which generally fall under the Data manipulation (Section III.3) and Modelling stages (Section III.4) below.

Initial ROBITT assessments have revealed that our data sets are not representative (Boyd & Turvey, 2023). Two prevailing issues are that few locations have been sampled consistently over time, and the English and Welsh lowlands are better sampled than elsewhere in the UK. We know that species’ occupancy varies geographically, so these issues probably cause a bias, the extent of which is not clear.

Ideally, RoB assessments should inform decisions taken downstream in the workflow, including which modelling framework is most appropriate. As RoB assessments are a new development, however, we implemented them retrospectively to identify where modifications are required (see Section IX).

### (3) Data manipulation

Having assessed the raw data for biases, the next step is to prepare those data for modelling. This data manipulation step includes harmonisation to common spatial, temporal and taxonomic resolutions, cropping the data to the desired extents in those dimensions and other types of (dis)aggregation and filtering. It may be necessary to revisit the data assessment stage if the data are modified appreciably at this stage (e.g. if the data are substantially coarsened or reduced in extent; Fig. 1).

Our choice of resolution is informed by several factors. First, we consider the resolution(s) at which the data were recorded (e.g. in DarwinCore the spatial resolution may be captured in the ‘coordinateUncertaintyInMetres’ field). Second, we consider the trade-off between coverage (the proportion of each dimension in the species–space–time cube for which we have data) and resolution (Rapacciolo *et al.*, 2021). Finally, we consider assumptions related to our modelling framework: for example, the spatial and temporal resolution at which it is reasonable to assume population closure.

At present, we unify the raw data at the species level (with some exceptions due to taxonomic separation

difficulties), 1 km (British Ordnance Survey grid) and day-level resolution. This involves discarding imprecise records and duplicates – both true duplicates, i.e. multiple records of the same observation, and records that become duplicates at the ‘visit’ level. [A visit comprises a unique combination of 1 km grid square (henceforth ‘site’) and date; aggregating the data at the visit level is necessary for our modelling framework.] However, acknowledging that the ecological and data-generation processes differ among taxonomic groups, it may be preferable to move beyond our ‘one size fits all’ approach in the future. One option would be to choose the spatial or temporal resolutions that result in the most even coverage (Jönsson *et al.*, 2021; Pescott *et al.*, 2019) for each group. However, scale effects mean that estimates made at different resolutions are not directly comparable, so workflow design faces a trade-off between generality and specificity.

Having discarded imprecise and duplicate records, we organise the remainder of the data into ‘detection histories’: dataframes indicating whether each species was recorded on each visit. This step has three purposes. The first is to reverse-engineer the survey structure (i.e. visits to some place on some day). The second is to infer non-detections of each species [what Rapacciolo *et al.* (2021, p. 1226) referred to as ‘borrowing strength across taxa’]. A species is considered to have gone undetected where it was not recorded on a visit but another species in the same taxon group was (Phillips *et al.*, 2009). The third is to approximate sampling effort per visit. We use the length of the list of species recorded on each visit, hereafter ‘list length’, as a proxy for sampling effort (Franklin, 1999; Szabo *et al.*, 2011; Van Strien, Van Swaay & Termaat, 2013). Arranging the data as detection histories is possible because we treat the records for each taxonomic group as a combined data set.

For many species, there are simply not enough data to estimate a trend in its distribution. A key question, therefore, is how to select which species should be taken forward to modelling in a way that introduces the fewest additional biases in downstream applications. In the past, we did not model species with fewer than 50 observations from the UK between 1970 and 2019 (Outhwaite *et al.*, 2019*a,b*). More recently, we have adopted thresholds based on the properties of those data sets that produce estimates with acceptable precision (Pocock *et al.*, 2019). Specifically, we set thresholds for the number of observations in the most frequently observed years and the number of sampling events that did not produce an observation of the focal species. An alternative approach would be to retain all species, even those which are likely to have low precision, to be transparent about our lack of knowledge about these species in downstream applications. Understanding the strengths and weaknesses of these choices is a priority as we further develop the workflow. Further research is required to explore whether these ‘rules of thumb’ are transferrable, whether they are applicable to all taxa, or whether alternative selection criteria would be preferable.

In addition to the taxonomic filters described above, we also remove data from poorly sampled portions of the



observation will be reported. Selective reporting arises from the tendency of some observers to record opportunistically, i.e. when something interesting or unusual is spotted. This behaviour would lead to under-recording of common species (August *et al.*, 2020a). Accounting for observer identity has been shown to improve the performance of spatial distribution models (Johnston *et al.*, 2018), so incorporation of observer effects in the workflow is desirable. At present we are hindered by the fact that observer identities are not regularised in most of the scheme data sets. The increased adoption of online recording technologies (e.g. iNaturalist – [www.inaturalist.org](http://www.inaturalist.org); iRecord; [www.brc.ac.uk/irecord](http://www.brc.ac.uk/irecord)) offers a potential solution in the longer term.

A second important source of heterogeneity in detection probability is the day of year, as most species have seasonal life histories. Van Strien *et al.* (2013) addressed this problem by modelling the phenology of detection as a quadratic function of Julian date. Our experience is that the parameters of this function are not mutually identifiable; hence, we have explored modelling phenology using a Gaussian distribution, in which the mean and standard deviation of detection dates are estimated. The Gaussian function is suitable for many species with annual life cycles, but not for long-lived or multi-voltine species, in which case a different formulation is required. This might involve splines (Crainiceanu, Ruppert & Wand, 2005) or additional levels of the hierarchy (Direnzo, Miller & Grant, 2021). A complementary approach is to include only those visits that occur within the active period of the focal species. Such refinement can also be applied spatially, where visits to sites far beyond the likely range of the focal species can be removed (Guzman *et al.*, 2021).

A third source of heterogeneity in detection probability is abundance. Detection is more likely on sites with abundant populations. Ignoring this variation can lead to biased estimation in occupancy models (Royle & Nichols, 2003).

Many data sets we encounter have few repeat visits to the same site on different dates in the same year, which are necessary for estimating detection probabilities. There has been some debate about whether it is appropriate to model detectability in this situation or whether it is better to estimate occupancy naively (i.e. assuming detectability = 1; Guillerá-Arroita *et al.*, 2014; Welsh, Lindenmayer & Donnelly, 2013). The decision on how to proceed depends on what the analyst considers to be useful information. Where repeat visits are few, estimates of occupancy are likely to be uncertain because the model does not know whether non-detections reflect absences or low detectability (i.e. multiple samples from the joint posterior of the parameters might fit similarly well). On the other hand, estimating occupancy naively will introduce a bias, especially if there is variation in detection probabilities over time (Isaac *et al.*, 2014). We have chosen to estimate detectability but acknowledge this may introduce biases where there is heterogeneity in site selection, recorder behaviour and detectability (see above). In future, we plan to assess further the sensitivity of our outputs to these methodological decisions.

### (c) Model fitting

We fit the occupancy-detection models to the detection histories in a Bayesian framework using Markov Chain Monte Carlo (MCMC) implemented in JAGS (Plummer, 2003) via the R package *sparta* (August *et al.*, 2020b). Outhwaite *et al.* (2019a,b) ran each model on three chains for 20,000 iterations with a burn in of 10,000 iterations and a thinning rate of 3. In our most recent set of models, we used 32,000 iterations with a burn in of 30,000 and a thinning rate of 6, which leads to improved mixing of the MCMC chains. These values were chosen to balance the trade-off between computation time and convergence, recognising that for some species there are insufficient data to achieve convergence for all parameters. Priors and hyperpriors are set to be uninformative [see Outhwaite *et al.* (2018) for details] with two exceptions: (i) the random walk in the state sub-model (see Section III.4.a); and (ii) detection probability for single-species lists is set to have a prior mean of 0.12. The value of 0.12 has no particular significance, but the choice to use this minimally informative value reflects prior knowledge that only one species was recorded, and that the mean of this parameter across species must be  $1/n$ , where  $n$  is the local species richness.

Model run times depend on the properties of the data to which they are fitted. It took around 80 min per species (on average) to fit the latest centipede models. The centipede data set contained 29,311 unique records of 55 species from 9276 1 km grid squares. The most recent ladybird data set, on the other hand, contained 296,279 unique records of 54 species from 43,805 grid squares and model fitting took well over 2 days per species.

### (5) Model evaluation

Having fitted statistical models (Fig. 1C), the next step is to evaluate their performance. Common measures of model performance include precision and goodness-of-fit (i.e. the plausibility of the model given the data; MacKenzie & Bailey, 2004). Goodness-of-fit is typically evaluated using the data to which the model was fitted (training data), but it is often desirable to assess a model using independent data.

For some species, there is insufficient information in the data to derive useful measures of change. Notwithstanding the *a priori* exclusion criteria described above, it is sometimes useful to exclude these species *a posteriori*. Several heuristics are available to assess this information content. One is the degree to which the parameter estimates from the MCMC chains have converged upon a common distribution, e.g. using the Gelman–Rubin ‘Rhat’ statistic (Gelman & Rubin, 1992). Another useful metric is the precision of the posterior distribution (e.g. of occupancy or trend), which captures the degree to which the data have overcome the minimally informative prior (Outhwaite *et al.*, 2018).

In our implementation of the workflow, we assess precision and convergence, but do not generally exclude modelled species based on these criteria. For multispecies indicators, we

reason that it is more transparent to propagate the uncertainty (Outhwaite *et al.*, 2020). In other situations, it might be preferable to remove species with low information content (e.g. when testing against a null hypothesis).

Goodness-of-fit is typically evaluated by comparing some fit statistic (e.g.  $\chi^2$ ) describing the discrepancy between the predictions and observations with those from a reference, or null, distribution (Warton *et al.*, 2017). The reference distribution is calculated by simulating many data sets under the model and calculating the equivalent fit statistics; that is, calculating the fit statistics that would be obtained if the model was a perfect representation of the system. Reference distributions may be constructed *via* bootstrapping for models analysed using classical inference (MacKenzie & Bailey, 2004) or as a natural by-product of the MCMC algorithm for models analysed in a Bayesian framework (Gelman, Meng & Stern, 1996; Royle *et al.*, 2007). The latter approach, often called a ‘posterior predictive check’, can be used to calculate Bayesian *P*-values (Kéry & Royle, 2016).

We have used posterior predictive checks in the past (Outhwaite *et al.*, 2020), but these, and the Bayesian *P*-value, have limited ability to detect a lack of fit (Wright, Irvine & Higgs, 2019). Residual plots, constructed for both the occupancy and detection components of the model (Warton *et al.*, 2017; Wright *et al.*, 2019), are a promising alternative that we aim to implement soon.

Precision and goodness-of-fit are useful measures of model performance, but where the available data contain unmodelled heterogeneity (as in our case), neither necessarily indicates a model’s accuracy. For this reason, it would also be useful to consider independent model evaluation using either independent data or elicitation of expert opinion. We are trialling approaches in which data providers are asked to provide opinions about whether model outputs are plausible (Boyd *et al.*, 2023; Pescott, 2022).

## (6) Populating the species–space–time cube

Once evaluated, model outputs for each species can be used to populate the final species–space–time cube in Fig. 1, the derived and modelled EBV (Kissling *et al.*, 2018). The minimum level of information required is a point estimate for each cell. Better is to provide information on uncertainty, such as bootstrapped estimates, standard errors, or posterior distributions from models fitted in a Bayesian framework. Ideally, information on model accuracy from the previous step, or on the risk of bias from the data assessment step, would be provided in addition to measures of uncertainty.

In our implementation of the workflow, we populate each cell in the species–space–time cube with 999 estimates of species occupancy. Occupancy is calculated as the proportion of (sampled) sites occupied by the focal species in each region (e.g. England, Scotland, Wales) and year. We also provide summary statistics, such as the mean, the standard deviation and the Rhat statistic for each cell (Outhwaite *et al.*, 2019b). In future, we plan to include information on model fit and the risk of bias.

## (7) Applications

Having populated the species–space–time cube with information on species’ changing distributions, the next step is to apply the cube for scientific research and to inform policy (Jetz *et al.*, 2019). Here, we focus on the applications for which the workflow was designed, which all involve the estimation of *temporal trends* in species’ occupancy.

### (a) Species trends

Estimates of occupancy for each species in each year across some spatial domain can be extracted from the species–space–time cube. These can be used to calculate temporal change as mean annual growth rates in occupancy (Outhwaite *et al.*, 2019a) or linear trends. Species-level trends are useful for identifying correlates of range contractions and expansions (Bowler *et al.*, 2021; Powney *et al.*, 2014), tracking the spread of invasive species and their effects on native taxa (Roy *et al.*, 2012) and conducting species *Red List* assessments (Maes *et al.*, 2015), amongst other applications.

### (b) Multispecies indicators

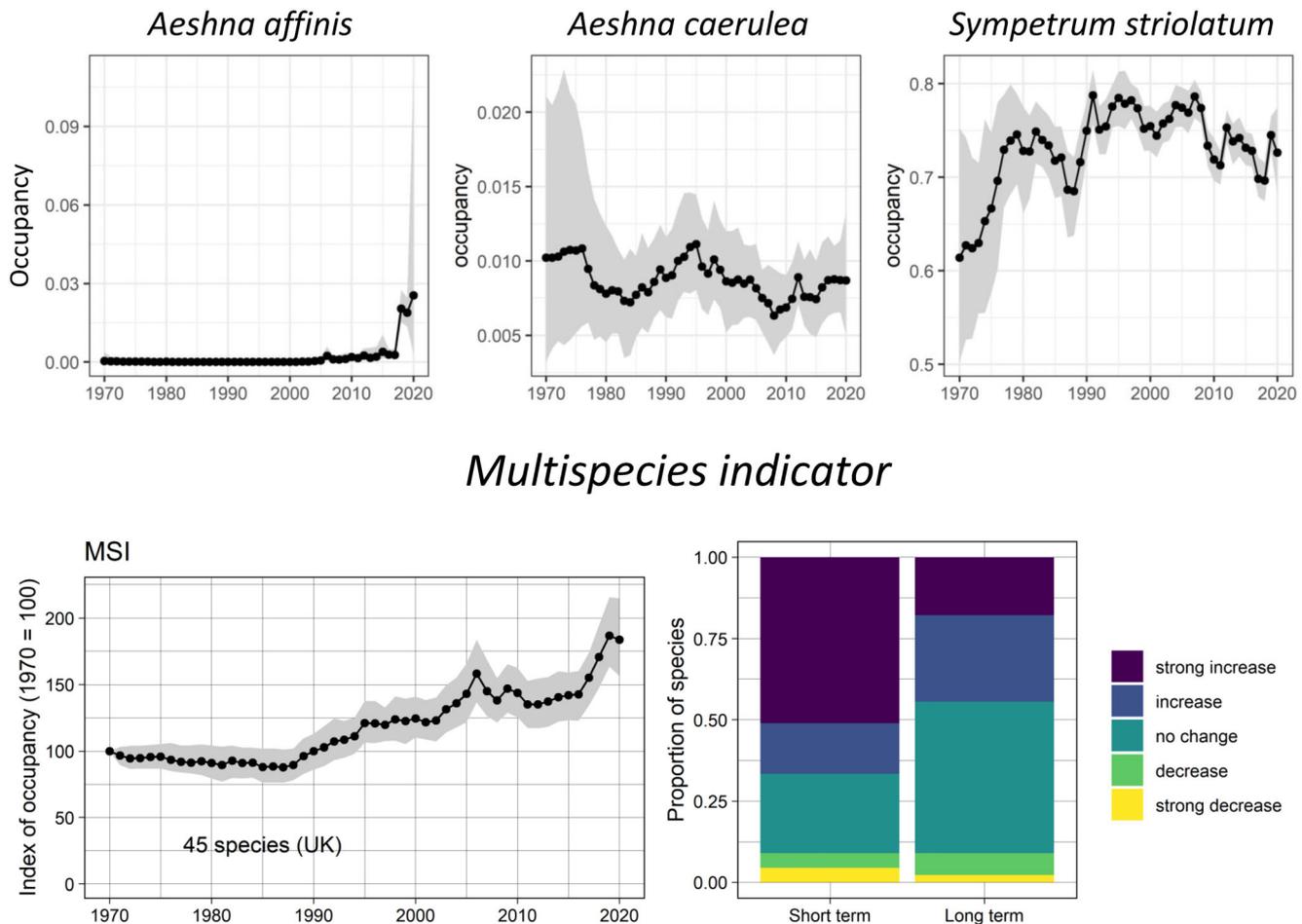
Species’ occupancy or trends thereof can be ‘averaged’ over some set of taxa to produce multispecies indicators. For many applications, the geometric mean is a sufficient summary statistic (Outhwaite *et al.*, 2020). More complex methods that can handle missing values and/or incorporate smoothing (Freeman *et al.*, 2020; Soldaat *et al.*, 2017) are now preferred for several national biodiversity indicators in the UK.

We have produced indicators for several taxonomic groupings and regions (Fig. 3). Our group is funded to produce multispecies trends, which are official government statistics, on pollinating insects and ‘priority species’ in the UK and England (e.g. Department of Environment Food and Rural Affairs, 2021). We have also produced an index of occupancy for ~2000 species in Scotland. These indicators, which are updated regularly, are derived from material in the peer-reviewed literature (Outhwaite *et al.*, 2019a, 2020; Powney *et al.*, 2019). In Fig. 3, we present an example multispecies indicator the 45 species of dragonfly for which we have data in the UK. The code to reproduce this example can be found in the online Supporting Information.

### (c) Comparing trends

As described above, our models include terms for regions within the UK. The regions are typically countries, but in principle, they can be any subset of grid squares. Regionalising the models provides a flexible way to assess variation in trends of specific groups, or to evaluate the impact of differing land-management strategies (e.g. comparing between land cover types, or between grid squares inside *versus* outside protected areas; Cooke *et al.*, 2023). In this way, our data products can be tailored to spatio-temporal resolutions that are most useful for decision-makers and policy creation (Jetz *et al.*, 2019) without the need to go back to the raw data.

## Species 1, species 2, ..., species 45



**Fig. 3.** Example occupancy trends for three species of dragonfly in the UK and a multispecies indicator (MSI) denoting the ‘average’ change over all 45 species for which we have data. In the top row, the  $y$  axes indicate the proportion of sampled sites occupied by the focal species; in the bottom row, the  $y$  axis in the left-hand panel is the geometric mean occupancy, scaled to a value of 100 in the baseline year (1970). In all four plots, the black points and line show the mean of the posterior distribution, with the 95% credible intervals delimited by the grey ribbon. The bar plot in the bottom right-hand panel indicates the proportion of species in various categories of change in both the short and long term. Short term is the final 5 years of the time series, and long term is the entire time-series. See online supporting information Appendix S1 for the worked example.

### (d) Functional diversity

Occupancy estimates can be combined with data on species traits to estimate patterns of functional diversity in space or time. Greenop *et al.* (2021), for example, identified species that pollinate and control pests; they combined trends across species to assess changes in these functions over time. This example demonstrates the potential of the workflow to inform on policy-relevant questions about ecosystem health or to provide for other EBV categories (community composition, ecosystem functioning; Pereira *et al.*, 2013).

### (8) Dissemination

The final step in the workflow is to disseminate the outputs of the preceding stages to the relevant audiences, which might include

policy makers, collaborators and the wider scientific community. This dissemination step should follow two general principles. First, data products should be FAIR (findable, accessible, interoperable and reusable; Wilkinson *et al.*, 2016). FAIR data can be easily found and accessed, use common standards that allow them to be combined with other EBV data sets and have appropriate metadata describing the data and how they were generated (see Section VI). Second, data products should be tailored to the target audiences, both in terms of the details of the use case (e.g. species trends or multispecies indicators) and data format (e.g. data and code *versus* interactive visualisation).

The primary audience for our data sets is colleagues within the same organisation (UK Centre for Ecology and Hydrology). We have built a shared computing environment, accessible to members of the team, which facilitates

collaborative working (see Section V). Controlled access by external collaborators may also be permitted *via* shared Notebooks in DataLabs (Hollaway *et al.*, 2020).

For the wider scientific community, we have published a species–space–time cube, generated using the workflow and pertaining to 5293 species, under an open government license with an accompanying data paper (Outhwaite *et al.*, 2019a,b). This data product comprises 1000 estimates of occupancy for each cell in the species–space–time cube (i.e. for each species/year/region combination), along with information on model convergence (see Section III.5).

For non-technical audiences, such as staff in government agencies, NGOs and some members of the schemes who supplied the raw data, we have developed R Shiny web applications deployed *via* DataLabs (see Section V). These allow users to browse outputs graphically without needing to download the underlying data.

Our data products are usually disseminated to stakeholders and the scientific community *via* reports and peer-reviewed papers. We provide reports accompanying biodiversity indicators to government agencies annually; the data are made openly accessible by the Joint Nature Conservation Committee (JNCC; <https://jncc.gov.uk>). Our data products also contribute to the triennial *State of Nature* reports in the UK (e.g. Hayhow *et al.*, 2019). Members of our team have published several papers using species' trends derived from the workflow (Cooke *et al.*, 2023; Greenop *et al.*, 2021; Outhwaite *et al.*, 2020, 2019a; Powney *et al.*, 2019).

A priority for future development of the workflow is to improve its interoperability. Hardisty *et al.* (2019) presented the 'Bari manifesto', which comprises 10 principles defining interoperability objectives for EBVs. Examples include the provision of human- and machine-readable metadata, data quality assurance, EBV workflows, the use of standard ontologies and the provision of information on data provenance. The workflow described here adheres to most of these principles to the extent that it is possible. However, some principles, such as the use of standard ontologies and the adoption of agreed-upon dimensions for all data products, cannot be acted upon unilaterally. Rather, they require cohesion among research teams and EBV producers, who need to make collective decisions about how to realise them.

Whilst we generally agree with the principles laid out in the Bari manifesto, we propose a slight update to the data quality principle. This principle refers to the adoption of quality assurance procedures and the propagation of the outputs to downstream data products. It is not clear, however, whether this principle covers issues of model accuracy, or whether it simply refers to the quality of the raw data. In our view, it would be preferable for this principle to incorporate model checks explicitly.

#### IV. IMPLEMENTING THE WORKFLOW

There are constraints on the frequency at which developers will be able to update their EBVs and biodiversity indicators.

First, there is the question of how frequently new data become available, which depends on the data that are used. There is also the matter of the computational cost of fitting models. In our experience, even with access to a high-performance computing cluster (see Section V), model fitting is a major time constraint. Complex models will take longer to run, particularly if they are fitted in a Bayesian framework, and this should be borne in mind at the model development stages above. Additional resourcing constraints, such as staff time, must also be considered.

In our case, we implement most stages of the workflow annually. This reflects the fact that we are funded to produce national indicators of species' distributions every year. However, we have neither the resources nor the data to update every taxonomic group each year. Typically, a handful of the ~30 taxonomic groups are updated, so for most groups the data are a few years out of date. Resourcing constraints mean that the Data assessment and Model evaluation steps are implemented less frequently, but we are working to change this.

#### V. COMPUTER INFRASTRUCTURE

Our realisation of the workflow is supported by an extensive computing infrastructure. Inputs to and outputs of most stages in the workflow are stored on what we call the 'Object Store'. The Object Store storage system manages data as objects referenced by a globally unique identifier, with attached metadata. The objects exist in a flat domain, allowing the Object Store to scale out much more easily than a traditional shared file system (i.e. it can handle large amounts of traffic). The object store ensures that our raw data and data products are easy to archive, locate and access through DataLabs. We have stored >2100 GB of data products, across 55 model runs, to date.

Models are fitted on DataLabs, and, where necessary, the Natural Environment Research Council's supercomputer, JASMIN (<https://jasmin.ac.uk/>). We create clusters in DataLabs for smaller tasks, or JASMIN's cluster facility – LOTUS – for larger jobs. LOTUS has direct access to the object store and *vice versa*, so data do not need to be copied between them manually.

#### VI. METADATA

Our system for producing metadata is quite comprehensive. We save EBVs at several points in the workflow (Fig. 1). Each time data are saved, metadata are stored in .rdata or .rds format. The (*Sparta*) model outputs for each species include metadata embedded as attributes in the R object. These metadata include modelled species names, the spatial and temporal extent of the models, the regions modelled, the quantity of data per region per species, the model

formulation and code, the *Sparta* version used, the date the model was fitted, the modeller who fitted it, the R session information and additional information on ‘provenance’. Provenance is a free-text field used to capture the motivation for the model run and/or to summarise data acquisition steps.

In addition to the species-level metadata, we store metadata at the ‘run’ level (i.e. for all species in a taxonomic group). These files are generated and updated using the *createMetadata* function in the *wrappeR* package (Boyd *et al.*, 2022b), which summarises the input and output files from the Object Store. The run-level metadata propagates metadata from the species level, where applicable, whilst also summarising higher level metadata, such as the number of species modelled. The propagation of metadata ensures that the EBV data products shown in Fig. 1 retain information about the raw data from which they were derived and the model configuration, thus being reproducible. Run-level metadata are subsequently used by functions in the *BRCindicators* package to create multispecies indicators based on the latest model outputs.

## VII. GENERALISABILITY OF THE WORKFLOW TO OTHER TAXA AND GEOGRAPHIC LOCATIONS

It is likely that research groups in other parts of the world will need to make some modifications to the workflow as we implement it in the UK. Each group will need to source its own data. Data may be retrieved from data aggregators such as GBIF, local databases, dedicated monitoring schemes, or meta-databases such as PREDICTS (Projecting Responses of Ecological Diversity in Changing Terrestrial Systems; Hudson *et al.*, 2017). The properties of the available data will determine whether and to what extent our specific implementation of the workflow is applicable.

The modelling step of the workflow is also likely to require adaptation. If the available data are severely biased at fine scales, it will be necessary to work at coarser resolutions at which those biases are less evident (Pescott *et al.*, 2019). However, the occupancy-detection model assumes that species’ occupancy at each site does not change within ‘closure periods’ (here 1 year); as the definition of the closure period becomes coarser, this assumption becomes less tenable. Likewise, it becomes less realistic to suppose that repeat visits to a site pertain to the same location. Where alternative analytical approaches are deemed more appropriate, the preceding stages will be affected (in particular, the data manipulation step, which is largely about preparing the data for modelling). In these cases, the general structure of the workflow will still apply, but the detail will differ.

Alternative models might be considered where the data are few, coarse or biased. Simpler models (e.g. Franklin, 1999; Telfer *et al.*, 2002; Roy *et al.*, 2012) require fewer data. Others are less contingent on fine-scale data (e.g. Hill, 2012).

Pescott *et al.* (2019) review the pros and cons of various range change models, albeit in the context of modelling relatively well-sampled plants in the UK. Each of these methods can be implemented in the R package *sparta* (August *et al.*, 2020b).

For some taxa and regions, structured monitoring data will be available in addition to unstructured data, in which case integrated distribution models might be considered (Isaac *et al.*, 2020). Our long-term ambition is to switch to integrated models for suitable taxa. Preliminary testing by our team has indicated that they take longer to converge than our current model, but that their estimates are more precise. Simulations suggest that covariates capturing geographic biasing mechanisms are needed to correct spatial biases in integrated models (Simmonds *et al.*, 2020).

Whilst the modelling step is likely to require some modification for use elsewhere, other steps in the workflow are broadly applicable. For example, ROBITT is applicable to all regions and taxa, so the Risk of bias assessment stage can be implemented as we have described it regardless of the specific application. The model evaluation step might require some adaptation where different models are used, but should generally comprise an assessment of model fit and solicitation of expert opinion. Our paper is aimed at developers of biodiversity monitoring infrastructures, who presumably plan to develop indicators or similar for relevant stakeholders, so the applications and dissemination steps are also likely to be broadly applicable.

## VIII. ‘IDEAL’ VERSUS ‘MINIMAL’ REQUIREMENTS FOR SPECIES DISTRIBUTION EBVs

Kissling *et al.* (2018) set out seven ‘ideal’ and ‘minimal’ criteria for species distribution EBVs, six of which relate to the extents and resolutions of the EBV data products in the spatial, temporal and taxonomic dimensions. In Table 1, we list the extents and resolutions of our Derived and modelled EBV, i.e. the final species–space–time cube in Fig. 1, and whether these meet the ideal or minimal requirements.

It is worth noting that the data products produced using the workflow presented here will differ from those produced using habitat suitability models (which are also used to construct EBVs) in terms of which of Kissling *et al.*’s (2018) criteria they satisfy. Our data products meet the ideal criteria relating to the temporal dimension of the species–space–time cube. This is not surprising because the ability to monitor species’ occupancy at a fine temporal resolution was the principal motivation for the development of the workflow. Most EBVs developed using habitat suitability models do not satisfy these criteria, perhaps because the relevant environmental data are typically not available at fine temporal resolutions (e.g. land cover products) or sufficiently far into the past. An exception are the data products derived using deductive habitat suitability models that are held on the *Map of Life* web platform. These models are projected onto

new land cover data each year, thereby enabling estimation of changes in the area of suitable habitat at a fine temporal resolution. However, these models are not suitable for many taxa, whose habitat requirements are not reflected in broad land cover classes, and the satellite-derived land cover data on which they rely are only available for the recent past. Whilst they are less likely to satisfy the ideal criteria relating to temporal resolutions and extents, EBVs developed using habitat suitability models are more likely to fulfil those relating to spatial resolutions and extents.

Kissling *et al.* (2018) asked the scientific community to develop their criteria. We recommend that the provision of information on model accuracy should be a minimal requirement. One option is to conduct sensitivity analyses with multiple models. Others include soliciting expert opinion on model outputs and including a statement of the risk of bias.

It is worth pointing out that the ‘ideal’ species distribution EBV is likely unattainable, as acknowledged by Kissling *et al.* (2018). First, there are trade-offs between criteria. For example, working at a fine spatial resolution precludes inclusion of species or regions for which such precise data are not available. The suggestion that species distribution EBVs should be global and temporally explicit is also optimistic given current data availability (Hughes *et al.*, 2020; Peterson & Soberón, 2018). For the foreseeable future, species distribution EBVs will be most useful if constrained in spatial or taxonomic domains, and/or if coarse resolutions are employed.

## IX. OUTSTANDING QUESTIONS

There remain outstanding questions at all stages in the workflow, the most pressing of which are outlined below.

### (1) Can we statistically correct for a wider range of biases in the species occurrence data?

At present, we construct our EBV using an occupancy-detection model that does not mitigate all biases. One way to improve the models is to include terms for covariates capturing differences between sampled and non-sampled cells in the species–space–time cube (Sterba, 2009). Others include poststratification and propensity score weighting (Elliott & Valliant, 2017).

### (2) Is the occupancy-detection model appropriate where repeat visits are few?

Occupancy-detection models use information from repeated visits to the same ‘site’ within closure periods (1 year in our case) to estimate detection probabilities. There has been some debate about whether it is sensible to estimate detectability where repeat visits are few, or whether it is better to model species’ occurrences naively [i.e. assuming detectability = 1 (Welsh *et al.*, 2013; Guillera-Arroita *et al.*, 2014)].

### (3) How do we evaluate model adequacy?

Implementing statistical fixes for data biases is one thing; assessing whether these were successful is another. Model evaluation is particularly difficult where the comparison data are biased, because a model with similar biases will appear to fit the data better than an unbiased one. More work is needed to understand which goodness-of-fit measures are most effective and to establish best practices for leveraging independent information (e.g. from experts or structured data).

### (4) What are the optimal species inclusion criteria and are they generalisable?

For some species, the data are so few that we can say little about their distributions. In this situation there are two options: (i) ignore the poorly recorded species and focus on those with more data; or (ii) accept the uncertainty and include all species to maximise taxonomic coverage. At present, we drop species based on the ‘rules of thumb’ described above, but it might be preferable to take a different approach in other circumstances.

### (5) Is the one-size-fits-all approach appropriate?

We estimate occupancy for each species at the same resolutions and extents using the same model. This ‘one-size-fits’ all approach is relatively simple, easy to implement and produces comparable outputs. However, questions remain about whether more bespoke models that capture taxonomic idiosyncrasies might be more appropriate and how best to combine the outputs of such models.

## X. CONCLUSIONS

- (1) To tackle the ongoing biodiversity crisis, accurate, synthetic, synoptic and interoperable data products are needed.
- (2) The workflow presented herein represents a step towards this ambition, yet as noted, substantial challenges remain.
- (3) We hope that research groups around the world will adopt this workflow, but consider these challenges, which are likely to be more acute in regions sampled less comprehensively than the UK (e.g. Boyd *et al.*, 2022a).
- (4) The operationalisation of regional and/or national EBV workflows will help us understand global biodiversity change better.

## XI. ACKNOWLEDGEMENTS

This work was supported by the Natural Environment Research Council award number NE/R016429/1 as part of the UK-SCAPE programme delivering National Capability. We are grateful to the many volunteer-run recording schemes for making their data available and for

constructive comments on the modelling framework. We would also like to thank Diana Bowler and two anonymous referees for their helpful comments on an earlier version of this manuscript.

## XII. REFERENCES

- AIELLO-LAMMENS, M. E., BORJA, R. A., RADOSAVIJEVIC, A., VILELA, B. & ANDERSON, R. P. (2015). spThin: an R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography* **38**, 541–545.
- ALTWEGG, R. & NICHOLS, J. D. (2019). Occupancy models for citizen-science data. *Methods in Ecology and Evolution* **10**, 8–21.
- AMINI TEHRANI, N., NAIMI, B. & JABOYEDOFF, M. (2021). Modeling current and future species distribution of breeding birds as regional essential biodiversity variables (SD EBVs): a bird perspective in Swiss Alps. *Global Ecology and Conservation* **27**, e01596.
- AUGUST, T., FOX, R., ROY, D. B. & POCOCK, M. J. O. (2020a). Data-derived metrics describing the behaviour of field-based citizen scientists provide insights for project design and modelling bias. *Scientific Reports* **10**, 11009.
- AUGUST, T., POWNEY, G., OUTHWAITE, C., HARROWER, C., HILL, M., HATFIELD, J., MANCINI, F. & ISAAC, N. (2020b). sparta: trend analysis for unstructured data. R package version 0.2.18. Electronic file available at <https://github.com/BiologicalRecordsCentre/sparta>. Accessed 20.11.2022.
- AUGUST, T. A., HARVEY, M. C., AUGUST, T. O. M., HARVEY, M., LIGHTFOOT, P., KILBEY, D., PAPADOPOULOS, T. & JEPSON, P. (2015). Emerging technologies for biological recording. *Biological Journal of the Linnean Society* **115**, 731–749.
- AUGUST, T. A., POWNEY, G. D., OUTHWAITE, C. L., HATFIELD, J., LOGIE, M. W. & ISAAC, N. J. B. (2022). BRCindicators: creating multispecies biodiversity indicators. Electronic file available at <https://github.com/BiologicalRecordsCentre/BRCindicators>. Accessed 20.11.2022.
- BAKER, E., DRURY, J. P., JUDGE, J., ROY, D. B., SMITH, G. C. & PHILIP, A. (2021). The verification of ecological citizen science data: current approaches and future possibilities. *Citizen Science: Theory and Practice* **6**(1), 12.
- BOWLER, D. E., KLAUS, D. E., CONZE, J., SUHLING, F., BAUMANN, K., BENKEN, T., BÖNSEL, A., BITTNER, T., DREWS, A., GÜNTHER, A., ISAAC, N. J. B. & PETZOLD, F. (2021). Winners and losers over 35 years of dragonfly and damselfly distributional change in Germany. *Diversity and Distributions* **27**, 1353–1366.
- BOYD, R. J., AIZEN, M. A., PRADO, L. F., FONTÚRBE, F. E., FRANCOY, T. M., MARTINEZ, L., MORALES, C. L., OLLERTON, J., PESCOTT, O. L., POWNEY, G. D., MAURO, A., RETO, S., EDUARDO, S. & CARVELL, C. (2022a). Inferring trends in pollinator distributions across the Neotropics from publicly available data remains challenging despite mobilization efforts. *Diversity and Distributions* **28**, 1404–1415.
- BOYD, R. J., COOKE, R., AUGUST, T. A., POWNEY, G. D., MANCINI, F., LOGIE, M. W. & ISAAC, N. J. B. (2022b). wrappR: a wrapper for producing biodiversity indicators. Electronic file available at <https://github.com/BiologicalRecordsCentre/wrappR>. Accessed 20.11.2022.
- BOYD, R. J., HARVEY, M., ROY, D. B., BARBER, T., HAYSOM, K. A., MACADAM, C. R., BALL, S. G., PESCOTT, O. L. & FK, S. (2023). Causal inference and large-scale expert validation shed light on the drivers of SDM accuracy and variance. *ecoevrxiv*. <https://doi.org/10.1111/ddi.13698>.
- BOYD, R. J., POWNEY, G., CARVELL, C. & PESCOTT, O. L. (2021). occAssess: an R package for assessing potential biases in species occurrence data. *Ecology and Evolution* **11**, 16177–16187.
- BOYD, R. J., POWNEY, G. D., BURNS, F., DANET, A., DUCHENNE, F., GRAINGER, M. J., JARVIS, S. G., MARTIN, G., NILSEN, E. B., PORCHER, E., STEWART, G. B., WILSON, O. J. & PESCOTT, O. L. (2022d). ROBITT: a tool for assessing the risk-of-bias in studies of temporal trends in ecology. *Methods in Ecology and Evolution* **13**, 1497–1507.
- BOYD, R. J. & TURVEY, K. (2023). ROBITT Assessments of UK Recording Scheme Data (0.1). Zenodo.
- CHAMBERLAIN, S., BARVE, V., MCGLINN, D., OLDONI, D., DESMET, P., GEFFERT, L. & RAM, K. (2021). rgbif: Interface to the 3.3.0, Global Biodiversity Information Facility API. R package version. <https://cran.r-project.org/package=rgbif>.
- CHAMBERLAIN, S. A. & SZÖCS, E. (2013). taxize: taxonomic search and retrieval in R. *F1000Research* **2**, 191.
- COOKE, R., MANCINI, F., BOYD, R. J., EVANS, K. L., SHAW, A., WEBB, T. J. & ISAAC, N. J. B. (2023). Protected areas support more species than unprotected areas in Great Britain, but lose them equally rapidly. *Biological Conservation* **278**, 109884.
- CRAINICEANU, C. M., RUPPERT, D. & WAND, M. P. (2005). Bayesian analysis for penalized spline regression using WinBUGS. *Journal of Statistical Software* **14**, 1–24.
- DAMGAARD, C. (2019). A critique of the space-for-time substitution practice in community ecology. *Trends in Ecology and Evolution* **34**, 416–421.
- DEPARTMENT OF ENVIRONMENT FOOD AND RURAL AFFAIRS (2021). UK biodiversity indicators 2021 revised, 1–57.
- DIRENZO, G. V., MILLER, D. A. W. & GRANT, E. H. C. (2021). Ignoring species availability biases occupancy estimates in single-scale occupancy models. *Methods in Ecology and Evolution* **18**, 1790–1804.
- DOSE, J. W., FINLEY, A. O., KÉRY, M. & ZIPKIN, E. F. (2022). spOccupancy: an R package for single-species, multi-species, and integrated spatial occupancy models. *Methods in Ecology and Evolution* **13**, 1670–1678.
- ELLIOTT, M. R. & VALLIANT, R. (2017). Inference for nonprobability samples. *Statistical Science* **32**, 249–264.
- FERNÁNDEZ, N., FERRIER, S., NAVARRO, L. M. & PEREIRA, H. M. (2020). Essential biodiversity variables: integrating in-situ observations and remote sensing through modeling. In *Remote Sensing of Plant Biodiversity*, pp. 485–501. New York City: Springer.
- FRANKLIN, D. C. (1999). Evidence of disarray amongst granivorous bird assemblages in the savannas of northern Australia, a region of sparse human settlement. *Biological Conservation* **90**, 53–68.
- FREEMAN, E. A. & MOISEN, G. (2008). PresenceAbsence: an R package for presence absence analysis. *Journal of Statistical Software* **23**, 1–31.
- FREEMAN, S. N., ISAAC, N. J. B., BESBEAS, P., DENNIS, E. B. & MORGAN, B. J. T. (2020). A generic method for estimating and smoothing multispecies biodiversity indicators using intermittent data. *Journal of Agricultural, Biological, and Environmental Statistics* **26**, 71–89.
- GELMAN, A., MENG, X. L. & STERN, H. (1996). Posterior predictive assessment of model fitness via realized discrepancies. *Statistica Sinica* **6**, 733–807.
- GELMAN, A. & RUBIN, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science* **7**, 457–511.
- GREENOP, A., WOODCOCK, B. A., OUTHWAITE, C. L., CARVELL, C., PYWELL, R. F., MANCINI, F., EDWARDS, F. K., JOHNSON, A. C. & ISAAC, N. J. B. (2021). Patterns of invertebrate functional diversity highlight the vulnerability of ecosystem services over a 45-year period. *Current Biology* **31**, 4627–4634.
- GUILLERA-ARROITA, G., LAHOZ-MONFORT, J. J., MACKENZIE, D. I., WINTLE, B. A. & MCCARTHY, M. A. (2014). Ignoring imperfect detection in biological surveys is dangerous: a response to ‘fitting and interpreting occupancy models’. *PLoS One* **9**, e99571.
- GURALNICK, R., WALLS, R. & JETZ, W. (2018). Humboldt Core – toward a standardized capture of biological inventories for biodiversity monitoring, modeling and assessment. *Ecography* **41**, 713–725.
- GUZMAN, L. M., JOHNSON, S. A., MOOERS, A. O. & GONIGLE, L. K. M. (2021). Using historical data to estimate bumble bee occurrence: variable trends across species provide little support for community-level declines. *Biological Conservation* **257**, 109141.
- HARDISTY, A. R., MICHENER, W. K., AGOSTI, D., ALONSO GARCÍA, E., BASTIN, L., BELBIN, L., BOWSER, A., BUTTIGIEG, P. L., CANHOS, D. A. L., EGLOFF, W., DE GIOVANNI, R., FIGUEIRA, R., GROOM, Q., GURALNICK, R. P., HOBERN, D., ET AL. (2019). The Bari Manifesto: an interoperability framework for essential biodiversity variables. *Ecological Informatics* **49**, 22–31.
- HAYHOW, D. B., EATON, M. A., STANBURY, A. J., BURNS, F., KIRBY, W. B., BAILEY, N., BECKMANN, B., BEDFORD, J., BOERSCH-SUPAN, P. H., COOMBER, F., DENNIS, E. B., DOLMAN, S. J., DUNN, E., HALL, J., HARROWER, C., HATFIELD, J. H., ET AL. (2019). The state of nature 2019. State of Nature Partnership. Electronic file available at <https://nbn.org.uk/wp-content/uploads/2019/09/State-of-Nature-2019-UK-full-report.pdf> Accessed 10.01.2023.
- HICKLING, R., ROY, D. B., HILL, J., FOX, R. & THOMAS, C. (2006). The distributions of a wide range of taxonomic groups are expanding polewards. *Global Change Biology* **12**, 450–455.
- HILL, M. O. (2012). Local frequency as a key to interpreting species occurrence data when recording effort is not known. *Methods in Ecology and Evolution* **3**, 195–205.
- HOLLOWAY, M. J., DEAN, G., BLAIR, G. S., BROWN, M., HENRYS, P. A. & WATKINS, J. (2020). Tackling the challenges of 21st-century open science and beyond: a data science lab approach. *Patterns* **1**, 100103.
- HUDSON, L. N., NEWBOLD, T., CONTU, S., HILL, S. L. L., LYSENKO, I., DE PALMA, A., PHILLIPS, H. R. P., ALHUSSEINI, T. I., BEDFORD, F. E., BENNETT, D. J., BOOTH, H., BURTON, V. J., CHNG, C. W. T., CHOIMES, A., CORREIA, D. L. P., ET AL. (2017). The database of the PREDICTS (projecting responses of ecological diversity in changing terrestrial systems) project. *Ecology and Evolution* **7**, 145–188.
- HUGHES, A., ORR, M., MA, K., COSTELLO, M., WALLER, J., PROVOOST, P., ZHU, C. & QIAO, H. (2020). Sampling biases shape our view of the natural world. *Ecography* **44**, 1259–1269.
- ISAAC, N. J. B., JARZYNA, M. A., KEEL, P., DAMBLY, L. I., BOERSCH-SUPAN, P. H., BROWNING, E., FREEMAN, S. N., GOLDING, N., GUILLERA-ARROITA, G.,

- HENRY, P. A., JARVIS, S., LAHOZ-MONFORT, J., PAGEL, J., PESCOTT, O. L., SCHMUCKI, R., *ET AL.* (2020). Data integration for large-scale models of species distributions. *Trends in Ecology and Evolution* **35**, 56–67.
- ISAAC, N. J. B. & POCOCK, M. J. O. (2015). Bias and information in biological records. *Biological Journal of the Linnean Society* **115**, 522–531.
- ISAAC, N. J. B., VAN STRIEN, A. J., AUGUST, T. A., DE ZEEUW, M. P. & ROY, D. B. (2014). Statistics for citizen science: extracting signals of change from noisy ecological data. *Methods in Ecology and Evolution* **5**, 1052–1060.
- JETZ, W., MCGEOCH, M. A., GURALNICK, R., FERRIER, S., BECK, J., COSTELLO, M. J., FERNANDEZ, M., GELLER, G. N., KEIL, P., MEROW, C., MEYER, C., MULLER-KARGER, F. E., PEREIRA, H. M., REGAN, E. C., SCHMELLER, D. S., *ET AL.* (2019). Essential biodiversity variables for mapping and monitoring species populations. *Nature Ecology and Evolution* **3**, 539–551.
- JETZ, W., MCPHERSON, J. M. & GURALNICK, R. P. (2012). Integrating biodiversity distribution knowledge: toward a global map of life. *Trends in Ecology and Evolution* **27**, 151–159.
- JOHNSTON, A., FINK, D., HOCHACHKA, W. M. & KELLING, S. (2018). Estimates of observer expertise improve species distributions from citizen science data. *Methods in Ecology and Evolution* **9**, 88–97.
- JÖNSSON, G. M., BROAD, G. R. & UMNER, S. S. (2021). A century of social wasp occupancy trends from natural history collections: spatiotemporal resolutions have little effect on model performance. *Insect Conservation and Diversity* **14**, 543–555.
- KAMP, J., OPPEL, S., HELDBJERG, H., NYEGAARD, T. & DONALD, P. F. (2016). Unstructured citizen science data fail to detect long-term population declines of common birds in Denmark. *Diversity and Distributions* **22**, 1024–1035.
- KÉRY, M. & ROYLE, J. A. (2016). *Applied Hierarchical Modelling in Ecology: Analysis of Species Distribution, Abundance and Species Richness in R and BUGS*. Academic press, London.
- 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., *ET AL.* (2018). Building essential biodiversity variables (EBVs) of species distribution and abundance at a global scale. *Biological Reviews* **93**, 600–625.
- LEWANDOWSKI, E. & SPECHT, H. (2015). Influence of volunteer and project characteristics on data quality of biological surveys. *Conservation Biology* **29**, 713–723.
- MACKENZIE, D. I. & BAILEY, L. L. (2004). Assessing the fit of site-occupancy models. *Journal of Agricultural, Biological, and Environmental Statistics* **9**, 300–318.
- MAES, D., ISAAC, N. J. B., HARROWER, C. A., COLLEN, B., VAN STRIEN, A. J. & ROY, D. B. (2015). The use of opportunistic data for IUCN Red List assessments. *Biological Journal of the Linnean Society* **115**, 690–706.
- MENG, X. L. (2018). Statistical paradises and paradoxes in big data (I): law of large populations, big data paradox, and the 2016 us presidential election. *Annals of Applied Statistics* **12**, 685–726.
- MEYER, C., KREFT, H., GURALNICK, R. & JETZ, W. (2015). Global priorities for an effective information basis of biodiversity distributions. *Nature Communications* **6**, 8221.
- OLIVER, R. Y., MEYER, C., RANIPETA, A., WINNER, K. & JETZ, W. (2021). Global and national trends, gaps, and opportunities in documenting and monitoring species distributions. *PLoS Biology* **19**, 1–14.
- OUTHWAITE, C., CHANDLER, R., POWNEY, G., COLLEN, B., GREGORY, R. D. & ISAAC, N. J. B. (2018). Prior specification in Bayesian occupancy modelling improves analysis of species occurrence data. *Ecological Indicators* **93**, 333–343.
- OUTHWAITE, C., GREGORY, R. D., CHANDLER, R. E., COLLEN, B. & ISAAC, N. J. B. (2020). Complex long-term biodiversity change among invertebrates, bryophytes and lichens. *Nature Ecology & Evolution* **4**, 384–392.
- OUTHWAITE, C., POWNEY, G., AUGUST, T., CHANDLER, R., RORKE, S., PESCOTT, O. L., HARVEY, M., ROY, H. E., FOX, R., ROY, D. B., ALEXANDER, K., BALL, S., BANTOCK, T., BARBER, T., BECKMANN, B. C., *ET AL.* (2019a). Annual estimates of occupancy for bryophytes, lichens and invertebrates in the UK, 1970–2015. *Scientific Data* **6**, 259.
- OUTHWAITE, C., POWNEY, G. D., AUGUST, T. A., CHANDLER, R. E., RORKE, S., PESCOTT, O., HARVEY, M., ROY, H. E., FOX, R., WALKER, K., ROY, D. B., ALEXANDER, K., BALL, S., BANTOCK, T., BARBER, T., *ET AL.* (2019b). *Annual estimates of occupancy for bryophytes, lichens and invertebrates in the UK (1970–2015)*. Wallingford, UK: EIDC, Electronic file available at <https://catalogue.ceh.ac.uk/documents/0ec7e549-57d4-4e2d-b2d3-2199e1578d84> Accessed 10.01.2023.
- PEREIRA, H. M., FERRIER, S., WALTERS, M., GELLER, G. N., JONGMAN, R. H. G., SCHOLES, R. J., BRUFORD, M. W., BRUMMITT, N., BUTCHART, S. H. M., CARDOSO, A. C., COOPS, N. C., DULLOO, E., FAITH, D. P., FREYHOF, J., GREGORY, R. D., *ET AL.* (2013). Essential biodiversity variables. *Science* **339**, 277–278.
- PESCOTT, O. L. (2022). A Google Sheets-Linked R Shiny App for the Expert Validation of Species Distribution Models (Version 1). Zenodo.
- PESCOTT, O. L., HUMPHREY, T. A., STROH, P. A. & WALKER, K. J. (2019). Temporal changes in distributions and the species atlas: how can British and Irish plant data shoulder the inferential burden? *British and Irish Botany* **1**, 250–282.
- PETERSON, A. T. & SOBERÓN, J. (2018). Essential biodiversity variables are not global. *Biodiversity and Conservation* **27**, 1277–1288.
- PHILLIPS, S. J., DUDÍK, M., ELITH, J., GRAHAM, C. H., LEHMANN, A., LEATHWICK, J. & FERRIER, S. (2009). Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological Applications* **19**, 181–197.
- PLUMMER, M. (2003). JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling. In Proceedings of the third international workshop on distributed statistical computing (DSC 2003), Vienna pp. 1–10.
- POCOCK, M. J. O., LOGIE, M. W., ISAAC, N. J. B., OUTHWAITE, C. L. & AUGUST, T. (2019). Rapid assessment of the suitability of multi-species citizen science datasets for occupancy trend analysis. *bioRxiv* 813626.
- POCOCK, M. J. O., ROY, H. E., PRESTON, C. D. & ROY, D. B. (2015). The Biological Records Centre: a pioneer of citizen science. *Biological Journal of the Linnean Society* **115**, 475–493.
- POWNEY, G. D., CARVELL, C., EDWARDS, M., MORRIS, R. K. A., ROY, H. E., WOODCOCK, B. A. & ISAAC, N. J. B. (2019). Widespread losses of pollinating insects in Britain. *Nature Communications* **10**, 1–6.
- POWNEY, G. D., RAPACCIUOLO, G., PRESTON, C. D., PURVIS, A. & ROY, D. B. (2014). A phylogenetically-informed trait-based analysis of range change in the vascular plant flora of Britain. *Biodiversity and Conservation* **23**, 171–185.
- RAPACCIUOLO, G., YOUNG, A. & JOHNSON, R. (2021). Deriving indicators of biodiversity change from unstructured community-contributed data. *Oikos* **130**, 1225–1239.
- ROY, H. E., ADRIAENS, T., ISAAC, N. J. B., KENIS, M., MARTIN, G. S., BROWN, P. M. J., HAUTIER, L., FROST, R., ROY, D. B., COMONT, R., ZINDEL, R., VAN VLAENDEREN, J., LANE, B. & GIFFORD, C. (2012). Invasive alien predator causes rapid declines of native European ladybirds. *Diversity and Distributions* **18**, 717–725.
- ROYLE, J. A. (2006). Site occupancy models with heterogeneous detection probabilities. *Biometrics* **62**, 97–102.
- ROYLE, J. A., KÉRY, M., GAUTIER, R. & SCHMID, H. (2007). Hierarchical spatial models of abundance and occurrence from imperfect survey data. *Ecological Monographs* **77**, 465–481.
- ROYLE, J. A. & NICHOLS, J. D. (2003). Estimating abundance from repeated presence-absence data or point counts. *Ecology* **84**, 777–790.
- RUETE, A. (2015). Displaying bias in sampling effort of data accessed from biodiversity databases using ignorance maps. *Biodiversity Data Journal* **3**, 1–15.
- 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.
- SIMMONDS, E. G., JARVIS, S. G., HENRY, P. A., ISAAC, N. J. B. & HARA, R. B. O. (2020). Is more data always better? A simulation study of benefits and limitations of integrated distribution models. *Ecography* **43**, 1413–1422.
- 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.
- STEEN, V. A., TINGLEY, M. W., PATON, P. & ELPHICK, C. (2020). Spatial thinning and class balancing: key choices lead to variation in the performance of species distribution models with citizen science data. *Methods in Ecology and Evolution* **12**, 216–226.
- STERBA, S. K. (2009). Alternative model-based and design-based frameworks for inference from samples to populations: from polarization to integration. *Multivariate Behavioral Research* **44**, 711–740.
- SULLIVAN, B. L., AYCRRIGG, J. L., BARRY, J. H., BONNEY, R. E., BRUNS, N., COOPER, C. B., DAMOULAS, T., DHONDT, A. A., DIETTERICH, T., FARNSWORTH, A., FINK, D., FITZPATRICK, J. W., FREDERICKS, T., GERBRACHT, J., GOMES, C., *ET AL.* (2014). The eBird enterprise: an integrated approach to development and application of citizen science. *Biological Conservation* **169**, 31–40.
- SZABO, J. K., VESK, P. A., BAXTER, P. W. J. & POSSINGHAM, H. P. (2011). Paying the extinction debt: woodland birds in the mount lofty ranges, South Australia. *Emu* **111**, 59–70.
- SZABO, J. U. K. S., VESK, P. E. A. V., BAXTER, P. E. W. J. B. & POSSINGHAM, H. (2010). Regional avian species declines estimated from volunteer-collected long-term data using list length analysis. *Ecological Applications* **20**, 2157–2169.
- TELFER, M., PRESTON, C. & ROTHERY, P. (2002). A general method for measuring relative change in range size from biological atlas data. *Biological Conservation* **107**, 99–109.
- VAN STRIEN, A. J., VAN SWAAY, C. A. M. & TERMAAT, T. (2013). Opportunistic citizen science data of animal species produce reliable estimates of distribution trends if analysed with occupancy models. *Journal of Applied Ecology* **50**, 1450–1458.
- VIHERVAARA, P., AUVINEN, A. P., MONONEN, L., TÖRMÄ, M., AHLROTH, P., ANTTLA, S., BÖTTCHER, K., FORSUUS, M., HEINO, J., HELIÖLÄ, J., KOSKELAINEN, M., KUUSSAARI, M., MEISSNER, K., OJALA, O., TUOMINEN, S., *ET AL.* (2017). How essential biodiversity variables and remote sensing can help national biodiversity monitoring. *Global Ecology and Conservation* **10**, 43–59.

- WARTON, D. I., STOKLOSA, J., GUILLERA-ARROITA, G., MACKENZIE, D. I. & WELSH, A. H. (2017). Graphical diagnostics for occupancy models with imperfect detection. *Methods in Ecology and Evolution* **8**, 408–419.
- WELSH, A. H., LINDENMAYER, D. B. & DONNELLY, C. F. (2013). Fitting and interpreting occupancy models. *PLoS One* **8**, 1–21.
- WIECZOREK, J., BÁNKI, O., BLUM, S., DECK, J., DÖRING, M., DRÖGE, G., ENDRESEN, D., GOLDSTEIN, P., LEARY, P., KRISHTALKA, L., TUAMA, É. Ó., ROBBINS, R. J., ROBERTSON, T. & YILMAZ, P. (2014). Meeting report: GBIF hackathon-workshop on Darwin core and sample data (22–24 may 2013). *Standards in Genomic Sciences* **9**, 585–598.
- WILKINSON, M. D., DUMONTIER, M., AALBERSBERG, I. J., APPLETON, G., AXTON, M., BAAK, A., BLOMBERG, N., BOITEN, J. W., DA SILVA SANTOS, L. B., BOURNE, P. E., BOUWMAN, J., BROOKES, A. J., CLARK, T., CROSAS, M., DILLO, I., *ET AL.* (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* **3**, 160018.
- WRIGHT, W. J., IRVINE, K. M. & HIGGS, M. D. (2019). Identifying occupancy model inadequacies: can residuals separately assess detection and presence? *Ecology* **100**, 1–9.
- ZIZKA, A., ANTONELLI, A. & SILVESTRO, D. (2021). Sampbias, a method for quantifying geographic sampling biases in species distribution data. *Ecography* **44**, 25–32.
- ZIZKA, A., SILVESTRO, D., ANDERMANN, T., AZEVEDO, J., DUARTE RITTER, C., EDLER, D., FAROOQ, H., HERDEAN, A., ARIZA, M., SCHARN, R., SVANTESSON, S., WENGSTRÖM, N., ZIZKA, V. & ANTONELLI, A. (2019). CoordinateCleaner: standardized cleaning of occurrence records from biological collection databases. *Methods in Ecology and Evolution* **10**, 744–751.

### XIII. SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**Appendix S1.** Code to produce multispecies indicators.

(Received 29 July 2022; revised 31 March 2023; accepted 3 April 2023; published online 16 April 2023)