v.6

ROBITT: a tool for assessing the risk-of-bias in studies of temporal trends in ecology

Boyd, R.J.[1], Powney, G.D.[1], Burns, F.[2], Danet, A.[3], Duchenne, F.[4], Grainger, M.[5], Jarvis, S.G.[6], Martin, G.[7], Nilsen, E.B.[5,10], Porcher, E.[3], Stewart, G.B.[8], Wilson, O.J.[9], Pescott, O.L.[1]

1. UK Centre for Ecology & Hydrology, Wallingford, UK
2. RSPB Centre for Conservation Science, David Attenborough Building, Pembroke Street, Cambridge, CB2 3QZ, UK
3. Centre d'Ecologie et des Sciences de la Conservation (CESCO), Muséum national d'Histoire naturelle, CNRS, Sorbonne Université, CP 135, 57 rue Cuvier 75005 Paris, France
4. Swiss Federal Institute for Forest, Snow and Landscape Research (WSL), 8903 Birmensdorf, Switzerland
5. Norwegian Institute for Nature Research (NINA),P.O. Box 5685, Torgard, NO-7485 Trondheim, Norway
6. UK Centre for Ecology & Hydrology, Lancaster Environment Centre, Lancaster, UK
7. Laboratoire EDB Évolution & Diversité Biologique UMR 5174, Université de Toulouse, Université Toulouse 3 Paul Sabatier, UPS, CNRS, IRD, 118 route de Narbonne, 31062 Toulouse, France
8. Evidence Synthesis Lab, School of Natural and Environmental Science, University of Newcastle, Newcastle-upon-Tyne, NE1 7RU
9. Plantlife, Salisbury, Wiltshire, UK
10. Faculty of Biosciences and Aquaculture, Nord University, Steinkjer, Norway

Abstract

1. Aggregated species occurrence and abundance data from disparate sources are increasingly accessible to ecologists for the analysis of temporal trends in biodiversity. However, sampling biases relevant to any given research question are often poorly explored and infrequently reported; this has the potential to undermine statistical inference. In other disciplines, but particularly medicine, researchers are frequently required to complete “risk-of-bias” assessments to expose and document the potential for biases to undermine inference. The huge growth in available data, and recent controversies surrounding their use to infer temporal trends, indicate that similar tools are urgently needed in ecology.

2. We introduce ROBITT, a structured tool for assessing the “Risk-Of-Bias In studies of Temporal Trends in ecology”. ROBITT has a similar format to its counterparts in other disciplines: it comprises signalling questions designed to elicit information on the potential for bias in key study domains. In answering these, users will define their inferential goal(s) and relevant statistical population. This information is used to assess potential sampling biases across domains relevant to the research question (e.g. geography, taxonomy, environment), and how these vary through time. If assessments indicate likely sampling biases, then the user must explain what mitigating action will be taken.

3. Everything that users need to complete a ROBITT assessment is provided: the tool, a guidance document, and a worked example. Following other disciplines, the tool and guidance document were

developed through a consensus-forming process across experts working in relevant areas of ecology and evidence synthesis.

4. We propose that researchers should be strongly encouraged to include a ROBITT assessment as supplementary information when publishing studies of biodiversity trends. This will help researchers to structure their thinking, clearly acknowledge potential sampling issues, and provides an opportunity to describe data checks that might otherwise not be reported. ROBITT will also enable reviewers, editors, and readers to establish whether research conclusions are supported given a particular dataset combined with some analytical approach. In turn, it should strengthen evidence-based policy and practice, reduce differing interpretations of data, and provide a clearer picture of the uncertainties associated with our understanding of ecological reality.

Key words
risk-of-bias; species occurrence data; temporal trends; Essential Biodiversity Variables; indicators; uncertainty

Introduction

Species occupancy and abundance are two fundamental state variables in ecology. Understanding the rates at which these variables might be changing is required to monitor progress towards international biodiversity targets and the effects of conservation interventions. Ultimately, this information comes from data documenting the sighting of one or more individuals of some taxon; that is, species occurrence data, or, in some countries, "biological records" (note that here we use this term to cover abundance data, as such information may be considered an attribute of an occurrence). Species occurrence data from disparate sources are often combined, and then analysed statistically, to derive measures of biodiversity over large taxonomic, spatial, and temporal extents (e.g. Gregory et al., 2005). Indeed, this is the premise of species population "Essential Biodiversity Variables" (Jetz et al., 2019; Kissling et al., 2018; Pereira et al., 2013). The temporal component of these data products may be averaged over spatial (e.g., regional, national, global) and taxonomic domains to produce indicators (GEO BON, 2015); these have become a key source of information on ecological change for policy makers (Navarro et al., 2017). Frequently then, evidence of temporal trends in aspects of global biodiversity are derived through the statistical analysis of raw data on the occurrence and/or abundance of various taxa.

Species occurrence data vary widely in terms of the ways in which they were recorded, why they were collected, and the information that they subsequently provide. First, presence-only data (sometimes also called primary biodiversity data; e.g. Ball-Damerow et al., 2019) document the sighting of some species, with information on where and when the sighting occurred. These data are derived from a variety of sources, including natural history collections in museums and herbaria, surveys by professional biologists, and various types of data collected by volunteer naturalists or other types of citizen scientist (Collen et al., 2013; Spear et al., 2017). Second, presence-absence data provide additional information on sampling events which did not yield a detection of the focal taxon. These data are most likely to be collected through structured monitoring schemes using specific protocols (but see Sullivan et al., 2014). Third, abundance data provide more information still: they document the number (or other quantity) of individuals. Both structured and unstructured (i.e. opportunistic) sampling schemes may collect abundance data, although this is often more likely to be associated with structured monitoring. All these data can provide information on temporal trends in biodiversity, but only if they are accessible to researchers.

In recent years, species occurrence data have both increased in volume, and have become increasingly accessible. This increase in accessibility can be ascribed to several initiatives: the digitization of historic biological records (Page et al., 2015); the proliferation and growth of citizen science monitoring initiatives (Spear et al., 2017); the launch of online data aggregators such as GBIF and similar regional portals (Nelson and Ellis, 2019); and the compilation of more specialist databases focused on particular types of ecological community (Dengler et al., 2011), monitoring data (Dornelas et al., 2018) or other evidence types (Hudson et al., 2017). Thanks to these initiatives, it is now relatively straightforward for ecologists to access large quantities of data, from disparate sources, and to use them for various analytical or other reporting purposes. However, quantity of data does not necessarily equal quality of scientific insight, and there have been important questions raised concerning the suitability of some biodiversity data types for drawing reliable inferences about true change over time (e.g. Ball-Damerow et al., 2019; Cardinale et al., 2018; Pescott et al., 2019).

To help the reader fully appreciate the potential challenges associated with the analysis of heterogeneous biodiversity data, it is useful to define some key statistical concepts (see Box 1 for a glossary of terms). Whilst there are many possible definitions of statistics (Barnett, 1982), one typical conception is that of reasoning under uncertainty and inherent variability, with classical texts (e.g. Lehmann, 1959) focusing on the use of observed data to make inferences concerning unobserved distributions. For example, monitoring-type investigations can be clearly appreciated as a sample-based approach to understanding features of some broader environment; likewise, smaller-scale experimental work is normally conducted with generalisation to a larger set of situations ultimately in mind. In both these cases, and everything in between, it is rarely feasible to census (or conduct experiments on) an entire population of interest: researchers rely on samples. This leads to questions concerning the validity of the resulting inferences. One assessment of a study’s validity is to ask whether these inferences are well-supported by the data in hand (known as internal validity). For sample-based results to be generalisable, however, they must also be true of the wider population of interest (external validity; Box 1). A study’s claimed external validity is likely to be undermined if the sample is not representative of the population with respect to important features for the desired inferences (Meng, 2018); this is often known as “sampling bias” (Box 1), although in some areas “selection bias” is used as a synonym.

To obtain a representative sample, researchers would ideally select individual units randomly from the population (probability sampling). However, there are many circumstances in which it is not possible to do this. In these cases, researchers might instead make use of nonprobability samples (such as those found in aggregated biodiversity databases); that is, samples that were not necessarily collected to be representative of a clearly defined statistical population, or which end up being unrepresentative due to data loss or subject drop-out. Small samples may also be unrepresentative of important features by chance, even if they are true probability samples. Before researchers can truly understand the representativeness of a sample, they must first formally define their research question and the statistical target population about which they intend to draw inferences. A failure to do this may result in researchers drawing erroneous conclusions about reality.

Box 1. Glossary of terms used in this paper and in the related literature on “risk-of-bias" assessments in other research areas (see Table 1).

Bias (in general): A systematic deviation from the truth in data acquisition, analysis, interpretation, publication etc. Due to the many ways in which such deviations might occur, some disciplines have created catalogues of biases in order to support clarity of communication and focus between researchers (e.g. see https://catalogofbias.org/ for health research).
Bias (in frequentist statistics): The tendency of a sampling distribution-based estimator to under- or overestimate the true population value of a parameter.
Bias (in sampling): Where samples do not match the population in important characteristics (i.e., the sample is not representative of the population). Note the difference between estimator bias in frequentist statistics and sampling bias: frequentist bias is a property evaluated over multiple samples, whereas sampling bias, in the more general sense of a failure of representativeness, can be defined in the context of a single sample. A small sample may be unrepresentative even though it is a true probability sample.

Convenience sampling: A form of nonprobability sampling, whereby units are selected because they are easily accessible. For example, a naturalist might record sightings in their back garden, or near a road, rather than travelling to a remote location. This results in a convenience sample; also sometimes called opportunistic sampling.

Probability sampling: Sampling from a statistical population using probability theory; that is, random sampling of some form. Results in probability samples which will be representative of the statistical population on average (although random variation dictates that this is not always the case for particular samples, especially with small sample sizes).

Nonprobability sampling: Sampling from a population whereby each member of that population does not have an equal chance of selection (or an equal chance conditional on some set of weights). Results in nonprobability samples which may be unrepresentative of the population in ways that could be important for resulting inferences if there are correlations between the selection mechanism and the properties of the population relevant to the research question (Meng, 2018).

Parameter: A (statistical distribution-based) estimate of some quantity pertaining to a statistical population (e.g., its mean, variance, etc.).

Statistic: An estimator of a population parameter derived from a sample of that population. (Note that a parametric model is implied by this definition).

Statistical inference: The process of drawing inferences about a statistical population from a sample (or set of samples) from that population; the inferences made from any given sample will also depend on the study goals (e.g. prediction, causal understanding, description) and the statistical model or models used to investigate these (Barnett, 1982). (Note that here we distinguish description, as one possible inferential goal, from descriptive statistics, which are normally defined as simple descriptors of a sample).

Statistical population: The complete set of units, with some identified properties, of which a sample is desired. For example, the true presence/absence status of all vascular plants in 1 km landcontaining grid cells in a country. Sometimes also called the target population, or simply "population". We avoid the latter term to avoid confusion with the common use of the term in ecology to refer to species' populations.

Representativeness: The degree to which a sample matches the population of interest in terms of features that are important for inference. This may apply to multiple domains; for example, a sample may be representative of the geographic domain of the population, but not the taxonomic domain. Representative samples are best obtained by probability sampling of a population. Unrepresentativeness is also referred to as **"coverage error"** in some disciplines (as distinct from the frequentist's concept of confidence interval coverage).

Validity: Validity is the degree to which a study's conclusions are likely to be true of the statistical population about which inferences are desired. This may be decomposed into internal and external validity. **Internal validity** is the extent to which the desired inferences can be accurately extracted from the sample, e.g., whether the effect of some intervention is well-identified from some experimental or quasi-experimental approach. **External validity** is the degree to which the results may or may not generalise to the wider population of interest. The concepts of internal and external validity are most frequently encountered in the context of causal inference, where the main aim is typically to estimate the effect of some intervention, e.g., the effects of some new medical treatment or social policy; however, they also apply to descriptive and predictive research. For example, describing national-level temporal trends in a species from a very unrepresentative sample would have low external validity, even if the data themselves were collected in a very robust fashion from the sampled locations (i.e. with high internal validity). In relation to any clearly stated inferential goal then, low and high validity map to high and low risks of bias respectively.

In studies of temporal biodiversity trends in ecology, researchers tend to define their statistical populations along the axes of space, time, and taxonomy (e.g. Dennis et al., 2019; Outhwaite et al., 2019; Powney et al., 2019; van Strien et al., 2019). For example, one might be interested in trends in bird distributions in North America over the period 1950 to the present day, or in pollinator abundance in Great Britain in the 20th century. It is also worth noting that, although they may not always be defined explicitly, other axes may be important for inference. For example, researchers may be more interested in whether samples represent all areas of some multi-dimensional environmental space (e.g. as defined by a set of climatic and/or habitat variables), rather than just being considered representative of geographic space. Likewise, for some purposes, representative coverage of species' traits may be desired along with, or instead of, even taxonomic or phylogenetic coverage. To be representative of populations conceived of in this way, data would ideally be representative of all axes of importance. To illustrate this point using the first of the above examples, data would need to be sampled as close to randomly as possible across North America, across all bird species that occupy North America, and evenly over the period 1950 to the present day. Otherwise, it is possible that the data will be unrepresentative of the statistical populations of interest; for example, particular geographical areas may be over- or under-sampled in particular time periods, leading to a confounding of time and space, and, ultimately, conclusions that may bear little resemblance to the true state of nature that a researcher intended to uncover.

There are many situations in which species occurrence data are unlikely to be representative of the statistical populations as defined (or merely implied) by studies of temporal biodiversity trends. Data collected opportunistically (convenience samples) are highly likely to be non-random along the axes of space, time, and taxonomy (or other important dimensions)—the same axes on which ecologists tend to define their target populations (e.g. Belitz, 2021; Hughes et al., 2020; Pescott et al., 2019). Volunteer naturalists, for example, tend to preferentially sample accessible and attractive locations, and interesting (e.g. rare) species (Barends et al., 2020; Prendergast et al., 1993). Structured data, collected according to some planned sampling design, may well be representative of some particular set of spatial, temporal, and taxonomic domains; however, when multiple datasets, with different aims, extents, and protocols, are aggregated (e.g. as on GBIF), then the target population to which these data pertain becomes unclear. To illustrate this point, imagine several datasets, each derived from structured monitoring of some taxon in some spatial unit at regular time intervals (i.e. time series data). These data might be very informative about change in those units (but see Gonzalez et al., 2016), however, there is no reason to suppose that these data can be combined and used to draw robust inferences about some wider geographic domain, unless the samples happen to resemble a probability sample of the true broader population(s) of interest (Cardinale et al., 2018). The problem of a mismatch between sample and population could be avoided if researchers first assessed their data to inform their choice of population and/or the claimed scope of their inferences. Unfortunately we find that such cases are the exception rather than the rule.

The frequent mismatch between sample and statistical target population in studies of trends in biodiversity has not gone unnoticed. It is a common subject for critical comments on studies in the ecological literature. For example, Sánchez-Bayo and Wyckhuys (2019) and van Klink et al. (2020) were both criticised for extrapolating their claims of insect declines beyond the taxonomic and geographical limits of their data (Desquilbet et al., 2020; Jähnig et al., 2021; Saunders et al., 2020; Simmons et al., 2019). Similarly, Vellend et al. (2013) and Dornelas et al. (2014) were criticised for concluding that local species richness is not in decline globally from meta-analyses of studies that were geographically biased in relation to human disturbance and species richness itself (Cardinale et al., 2018; Gonzalez et al., 2016). Crossley et al. (2020) and van Klink et al. (2020), on the other hand, were taxonomically selective when reporting their conclusions: both sets of authors included non-insect groups in their analyses, but restricted their conclusions (and paper titles) to insects (Desquilbet et al., 2021, 2020). Other studies of insect trends have been criticised with regards to whether particular modelling approaches have appropriately dealt with temporal biases in the data. For example, both Lister and Garcia (2018) and Soroye et al. (2020) have been criticised in this regard (Anon., 2020; Guzman et al., 2021; Willig et al., 2019). This brief overview of some recent disagreements in the ecological literature on biodiversity trends highlights a fundamental problem: potential sampling biases are rarely communicated to the reader (and reviewers) in sufficient detail; instead, they are often addressed with a passing comment, if at all. This problem is particularly evident where the external validity of claims is concerned (Cardinale et al. 2018). Even if models are presented which claim to deal with various types of potential bias, clear descriptions of these biases, explanations of how proposed models address these specifically, and critical exploration of these claims (e.g. via appropriate sensitivity analyses) are only rarely encountered. As a result, inferential “over-reach” in paper titles, abstracts, and conclusions is not uncommon.

In some other disciplines, strategies have been developed to assist researchers in avoiding potentially inappropriate inferential claims. In medicine and related areas for example, researchers are strongly encouraged, if not mandated, to submit “risk-of-bias” (RoB) assessments when publishing their studies (Table 1). The function of these assessments is to clearly expose and document any threats to the validity of a study’s conclusions arising from potential biases in the underlying data. A number of RoB assessment tools have been developed for various types of data and study design (Table 1). These tools have been described as reflecting a “shift in focus from methodological quality to risk of bias” (Sterne et al., 2016)—a shift that has yet to take place in ecology, despite some efforts to provide structured approaches to documenting methodological choices in some areas (Grimm et al., 2010, 2006; Schmolke et al., 2010; Zurell et al., 2020). It is not difficult to appreciate why this shift was needed in the medical sciences: one would not want to approve some pharmaceutical product which had been demonstrated to be safe only for some subset of a population, for example. We argue here that the increasing policy relevance of inferences about temporal trends in biodiversity requires a similar shift in focus in our discipline.

Table 1. A (non-exhaustive) list of risk-of-bias (RoB) tools, and broader checklists with RoB elements, used across scientific disciplines.

Tool	**Field**	**Study/ data type**	**Details**	**Reference(s)**
Cochrane RoB tool	Medicine	Randomized controlled trials of medical interventions	Used where studies are to be included in systematic reviews	Version 1: Higgins et al. (2011) Version 2: Sterne et al. (2019)
Constraints to Generality tool	Psychology	Any inferential study	Engenders clear definition of the statistical population of interest and assesses external validity	Simons et al. (2017)
GRADE	Medicine	Medical interventions	Not exclusively a RoB tool but contains a RoB component.	Schünemann et al. (2013)
PRISMA	Cross-discipline	Systematic reviews and meta-analyses	Contains questions about RoB at both the study and overall review level	2020 version: Page et al. (2021)
PROBAST	Medicine	Predictive modelling studies	Used for predictive studies of diagnoses and prognoses	Wolff et al. (2019)
RoBANS	Medicine	Non-randomized (observational) studies of medical interventions	-	Kim et al. (2013)

ROBINS-E	Public health	Non-randomized (observational) studies of exposure	Akin to ROBINS-I, but for studies of effects of exposure	Bero et al. (2018)
ROBINS-I	Medicine	Non-randomized (observational) studies of medical interventions	Compares data to that of a hypothetical randomized trial	Sterne et al. (2016)

In this paper, we introduce ROBITT, a tool for assessing the "Risk-Of-Bias In studies of Temporal Trends in ecology". The tool has a similar format to its counterparts in other fields: it comprises a number of "signalling" questions (Sterne et al., 2016), each designed to elicit information on the potential for bias in the study being assessed. The user is first asked to define the statistical target population about which they intend to make inferences, and then to assess whether their data are likely to be representative of this population in the geographic, temporal, environmental, and taxonomic domains (the latter defined broadly as covering any organismal space that might be important for inferences for the study, e.g. taxonomic, phylogenetic, trait-based etc.) If the data are found to be potentially biased, then the user is also asked to explain how they will mitigate those biases. We begin by describing the development of the tool, including the solicitation of expert advice. We then provide an overview of the tool, describe its constituent sections, and refer the reader to the supplementary material for a full template, guidance document, and worked example. Finally, we discuss the potential value of the ROBITT tool to the field of ecology, and propose its inclusion as supplementary information for all studies of temporal trends based on all biodiversity datasets of occurrence and/or abundance data, but particularly for those retrieved from aggregated databases where true probability samples of the populations of inferential interest are unlikely to be found.

ROBITT tool

Development

A prototype version of the ROBITT tool was created by RB, GP, and OP (authors). In a similar way to other RoB tools (e.g. Sterne et al., 2016), this prototype was refined over the course of two online workshops in which participants (the other authors) gave critical feedback. During the first workshop, participants were asked to review the tool, guided, but not restricted, by a broad set of prompting questions (supplementary material **3**). During the second workshop, each participant presented their feedback. Feedback was grouped into broad themes that warranted discussion in terms of establishing a consensus across participants on the content and presentation of the tool. The workshop chair (OP) outlined the perceived consensus after open discussion on each point, and participants were asked for further comments. The ROBITT tool and guidance document were then updated and circulated to participants for additional feedback (this process used an online word processing tool, so that all feedback was visible to all participants). The tool and guidance versions presented here are the final result of this consensus-forming process.

Overview

The ROBITT tool comprises 17 questions designed to elicit information on the potential for bias in a study. The user may answer the questions using text and/or figures. The first section, the "research statement and pre-bias assessments", comprises four questions; the remainder of the tool constitutes the bias assessment itself. The ROBITT tool and supporting guidance document can be found in supplementary materials **1 and 2**. The guidance document follows the PRISMA model (Page et al., 2021b), in that an explanation of the rationale for each question is given, followed by a concise summary of the expected response. A worked example can be found in supplementary material **4**.

Tool sections

Research statement and pre-bias assessment

The purpose of this section is to assemble all the information needed to assess the risk of bias in the focal study. The first step in this section is to formally define the statistical target population about which inferences are desired. This must include a specification of the extents of any relevant domains of that population (e.g. geographic, temporal, taxonomic, environmental). It must also include a statement of the resolutions (grain sizes) at which analyses will be conducted (e.g., 1 km grid cells, annual increments, species level etc.). This is because the scale at which a research question is formulated can influence both data availability and the nature of, and potential for, biases (e.g. Pescott et al., 2019). The next step is to state the inferential goal; for example, "to estimate temporal trends in species' occupancy". In the remainder of this section, the user must document data provenances, and explain and justify any steps that were taken to modify or otherwise clean data.

Bias assessment

The next and main section of the ROBITT tool is the bias assessment. This section begins with a specification of the geographic (i.e. spatial), temporal, and taxonomic resolutions (i.e. grain sizes or scales) at which the assessment will be conducted. Generally, these should match the resolutions at which inferences are desired (as specified in the research statement section). It would very likely be inappropriate, for example, to assess one's data in decadal time periods and 100 km grid cells, and then conclude that they were unbiased for making inferences about annual time periods at the 1 km resolution. We note that there may be limited exceptions to this: for example, one might intend to draw inferences at the taxonomic resolution of the species, but be unable to conduct their assessment at this scale. This is because it can be difficult to assess sampling biases at the species level using presence-only data, because such data say nothing, in isolation, about sampling effort where the focal species was not recorded. In this case, a common approach is to combine all records across species in a taxonomic group and assume that the combined distribution of those records reflects the distribution of sampling (see e.g. Phillips et al., 2009). Here then, the scale of a bias assessment might be subtly different to that at which inferences are desired, because information at one level is assumed (under some model) to provide information for assessing biases at another.

The next three subsections denote our three main domains of potential bias: geographic, environmental and taxonomic (or other relevant organismal axis, such as phylogenetic or functional group). Note that temporal biases are dealt with within each of these three sections (see below for more on this). In each subsection, the user must answer three questions: the first two questions in each subsection are designed to reveal any potential biases relative to the research question (I.e. the inferential goal). The first asks whether the data are representative of that domain; that is to say, do the data cover the whole domain evenly (ideally randomly)? The answer to this question will provide an indication of whether the desired scope of the inferential goal in that domain is justifiable. The second question asks whether the same portion of the focal domain has been sampled over time; that is, is there any indication of temporal changes in coverage in that domain? The answers to the second question in each subsection are crucial for assessing the suitability of the data for estimating temporal trends in biodiversity. To illustrate this point, imagine that species data are collected from one portion of some country in one time period, and then from another part in the next. Using these data to estimate changes in species' distributions or abundances between time periods may then be problematic, because shifts in space are confounded with shifts in time. In one sense, the distinction between the first and second question can be considered equivalent to the distinction between external and internal validity (Box 1): a study might have low external validity (i.e. broad generalisability) if it is not representative of some domain overall; however, for a subset of that domain (e.g. a well-sampled portion of geographic space, perhaps a site for which a longitudinal study was conducted), the data might be very informative about change (i.e. high internal validity). The answers to these first two questions in each domain have important implications for how one goes about answering the third.

The third question in each domain subsection asks the user to state how they will mitigate any potential biases indicated by the preceding two questions. There are several ways in which one might go about mitigating biases, such as through data manipulation or statistical correction procedures. Mitigation might also include simply redefining the statistical target population; this approach changes the question to be more appropriate to the data in hand, rather than attempting to use some model-based fix to support broader inference. We briefly review mitigation options in the Discussion section below. There will also be cases in which the user feels it is not necessary to mitigate a lack of coverage or inconsistent sampling over time, because these are not relevant to their inferential goal. For example, if one is interested in understanding how species' abundances have changed in South Africa, then the data need not be sampled from the same portion of environmental space over time; indeed, if the data are representative of South Africa geographically in each time period, then they will necessarily be sampled from different portions of environmental space as land use and climate changes. Users should not feel compelled to explain a lack of coverage in any domain if it is not relevant to their inferential goal.

The next, and final, subsection in the bias assessment is “Other potential biases”. This subsection is slightly different to the previous three subsections in that it does not relate to a single domain of bias; rather, it provides an opportunity for the user to consider any additional biases that might impact their research and to highlight any mitigation approaches. The first question in this subsection asks whether there are any temporal biases in the data that do not relate to the ecological states of interest (e.g. occupancy, abundance etc.). Often, these biases will relate to the estimation of some parameter in a statistical model. For example, site-occupancy models are frequently used to estimate temporal trends in species' distributions (Kéry and Royle, 2016). These models normally require data from replicate visits to particular sites within short spaces of time (“closure periods”) to estimate species' detection probabilities. Where these models are used, analysts might want to consider whether there is variation in the quantity and type of repeat visits that could result in biases in estimates in these parameters (Royle, 2006). We note that biases of this type may well go unnoticed in the answers to the previous questions. The next question in this subsection asks the user to consider whether there are any other biases that are not covered by the preceding questions that might potentially cause problems for inference. Examples include biases relating to phenology, time of day, temporal baselines, etc. (e.g. Buckland and Johnston, 2017). The final question in this subsection is equivalent to the final question in the other domains: users are asked to explain, in detail, how they plan to mitigate any biases revealed in their answers to the two preceding questions. See the guidance document in **supplementary material 2** for greater detail on the expected content of researcher responses to the ROBITT questions, and other additional background information.

Completing a ROBITT assessment

Whilst the questions that constitute ROBITT require individual answers, it may be the case that researchers prefer to provide such responses in the main text of a research paper or report. This is not a problem; as a point of comparison, PRISMA (Page et al., 2021a) provides a checklist format that allows researchers to direct the reader to the answer to any given question. This could also be the case here; for example, clear subheading or section references could be provided in response to a question, provided it was the case that the manuscript text referenced was a clear and complete answer to that question.

Users may go about answering the 13 questions in the bias assessment section of ROBITT in the best ways they see fit. It is worth noting, however, that there exists a substantial literature of studies which screen species occurrence data for various potential biases and, between them, these provide a suite of heuristics that could be deployed in answering such questions (Barends et al., 2020; Boyd et al., 2021; Meyer et al., 2016; Petersen et al., 2021; Ruete, 2015; Speed et al., 2018; Sumner et al., 2019; Troudet et al., 2018). In the simplest case, for example, a researcher might map their data to assess representativeness in the geographic domain of interest. Indeed, several maps could be produced, each pertaining to some time period, to assess temporal changes in geographic coverage. More formally, one could screen data for potential geographic biases by comparing the nearest neighbor distances of their data to those of a simulated random distribution (Clark and Evans, 1954). This gives a “nearest neighbour index” which indicates the extent to which the data departs from a random distribution in geographic space. Similar heuristics have been proposed to screen data for biases in the taxonomic and environmental dimensions. Instead, see the literature cited earlier in this paragraph, and Boyd et al. (2021) for a review and R package in which many such heuristics are implemented. We note, however, that a ROBITT assessment is not merely intended to be a repository for some set of contextless numbers or figures: sampling bias can strictly only be defined in relation to some inferential goal. The central point of ROBITT is that assessments of potential bias are clearly linked to a research question, and assessed in the context of this and any analytical tools being used to answer that question.

In some cases completing a ROBITT assessment will be an iterative process. For example, researchers might complete a first iteration of the tool and find that data coverage is not sufficient in portions of their geographic domain of interest. In this case, they might decide to redefine their geographical domain to exclude the poorly sampled regions; this would mean completing a second iteration of the ROBITT form using the appropriate subset of the data. If users become aware that their data do not permit inferences across their desired domains midway through completing an iteration of the ROBITT form, there may be no need to continue with that iteration; instead, they might decide to redefine their domain as required and move onto the next iteration without answering every question. In these cases, the user should clearly version control (i.e. tracking and recording changes over time) their documents and provide this history as supporting information to their studies.

Discussion

Sampling biases have long been recognised as a challenge for inference in ecology (e.g. Peters, 1991), however, no formal tools for assessing these have been produced to date. In this paper, we have designed and introduced ROBITT, a tool for assessing the potential “Risk-Of-Bias in studies of Temporal Trends in ecology”. The tool comprises a number of questions, each designed to clearly elicit the potential for bias in the study under assessment. In answering these questions, users will define their research question (i.e. inferential goal) and statistical target population across relevant domains, and then assess the degree to which their data are likely to be representative of those domains. We propose that researchers be strongly encouraged to include a ROBITT assessment as supporting information when publishing studies of temporal trends in biodiversity. We expect that this will support scientists in writing clear methods sections, strengthen evidence-based policy and practice, help resolve scientific controversies around biodiversity trends, assist editors, reviewers, and readers, and, ultimately, provide a more accurate picture of ecological reality. Accumulated over studies, ROBITT assessments will also highlight where data are required to address pressing questions concerning biodiversity change.

We hope that the completion of a ROBITT form will become a standard requirement where researchers aim to estimate temporal trends in biodiversity from species occurrence data of any type. Many of the tools listed in Table 1 have set similar precedents in other disciplines. Whilst some reporting tools for various subdisciplines of ecology already exist, they do not exclusively focus on risk-of-bias. These include the ODD (Grimm et al., 2010, 2006) and TRACE (Schmolke et al., 2010) protocols for describing and documenting individual- and agent-based models (IBMs), and the ODMAP (Zurell et al., 2020) protocol for documenting the use of species distribution models. ODD and TRACE straddle the line between being strongly encouraged and mandatory: it would be very unusual for a new IBM, or even a study in which an IBM is used, to be published if it was not described and documented using these tools. ODMAP is a much newer tool, but we strongly suspect that this will become a similar requirement. If ROBITT can achieve similar uptake to reporting tools existing elsewhere in ecology and other disciplines, then we suggest that the field will benefit from a much clearer evidence base. In the medical sciences, some reporting tools have evolved from a general focus on methodological quality, to a more specific, and arguably more in-depth, focus on the impacts of potential bias to inference (Sterne et al., 2016). There is no doubt a place for both in ecology (indeed, some tools in medicine combine these aspects, e.g. Page et al., 2021a), however, we agree with Sterne et al. (2016), and others in the medical literature, that in-depth, qualitative, assessments of risk-of-bias across relevant domains are more useful and revealing than simply checking procedural items off a list.

We suggest that researchers will get the greatest benefit from the ROBITT tool if they use it to structure their research. The questions contained in the tool are questions that researchers should be asking themselves regardless; indeed, ROBITT provides an opportunity to demonstrate the large amount of work that goes into studies of temporal trends in biodiversity, but which may not end up being described in a paper. If, on the other hand, a ROBITT form is completed just before, say, the submission of an article for publication, then it may reveal problems that could otherwise have been dealt with earlier. Completing the ROBITT form during the research process therefore has the potential to save researchers' time, by providing a clear framework for structuring thought and making methodological decisions.

Much of the risk-of-bias literature in other disciplines has focused on the effects of interventions (see Table 1). In this type of research, the questions asked are typically explicitly causal, because the desired inference concerns whether some action results in some outcome. This has also been the standard focus of evidence-based conservation, at least historically (e.g. Lortie et al., 2015). ROBITT, on the other hand, is primarily focused on descriptive inference of the type that is often used for ecological indicators (e.g. Gregory et al., 2005), or, more recently, by the Essential Biodiversity Variable literature (e.g. Jetz et al., 2019). However, this distinction is not absolute, and there are many examples of ecological studies that use aggregated, or other, species occurrence data in attempts to reach causal conclusions. For example, Woodcock et al. (2016) split wild bee occurrence data for Britain into two spatial subsets based on the agricultural use of neonicotinoid insecticides, assessing temporal trends in occupancy for the taxa concerned in each subset. Whilst this type of assessment is correlative, there is often clearly a causal motivation (for example, the title of Woodcock et al. 2016 certainly implies causality). Whilst the ROBITT tool has not been designed to deal explicitly with these types of situations, we suggest that the tool will still be useful when researchers are attempting to make causal inferences from observational data. In this example (Woodcock et al., 2016), the domain representativeness of the data in the two spatial subsets could have been assessed separately in order to investigate the potential for confounding; additionally, the full dataset could also have been assessed in order to investigate its overall external validity. In the future however, we anticipate that other risk-of-bias tools will be developed within ecology to fully meet the needs of causal inference research (see Table 1 for potential models).

During the workshops at which we refined the ROBITT tool, several key issues were raised by participants. One issue that was raised by multiple participants was that, whilst it might be easy to define one’s target population, in some cases it will not be easy to determine whether any given sample is representative of that population. For example, a researcher might define their population as wild bees in Chile in the 2010s. They might then download presence-only data from GBIF and attempt to assess its representativeness. Mapping the data might reveal that they are not randomly distributed across the country, but does this reflect the true distribution of wild bees in Chile, or does it reflect non-random sampling? The user might also want to establish whether they have data for all known species of wild bee in Chile; how do they know whether this is the case? The answers to these questions will vary on a case-by-case basis.

Whilst it will not always necessarily be easy to establish whether a sample is representative of a population or not, we propose some simple criteria. First, and most importantly, wherever possible subject-matter experts should be consulted; such experts should be able to tease apart sampling biases from true biological phenomena. For example, an expert might know, or suspect, that a species or taxon group occupies areas where it has not been recorded (e.g. perhaps historic data have not been digitised, or more current data are not being shared effectively); this is likely to be a strong indication of sampling bias. On the other hand, the expert might suggest that the distribution of presence-only data reflects their understanding of the distribution of the focal taxon; this could indicate that the data are unbiased despite being non-randomly distributed across the geographical domain of interest (we note that expert-drawn range maps might be useful in this respect). Second, it might be possible to supplement expert advice with information from other sources. For example, regional or national Floras, Faunas, and other such taxonomic publications may list (undigitised) specimens, or provide information on regional occurrences at a coarse spatio-temporal level. These resources can also sometimes be accessed online (e.g. https://www.discoverlife.org/mp/20q?guide=Apoidea_species). Third, when using presence-only data for a reasonably large number of species in the same taxonomic group (e.g., bees, birds), it may be acceptable to assume that the combined distribution of records for all species approximates the distribution of sampling. Indeed, this is the premise of well-established bias mitigation strategies for fitting species distribution models (Dudík et al., 2005; Phillips et al., 2009). In this case, the combined data would ideally be randomly distributed across the geographical domain of interest. Fourth, presence/absence and abundance data are a direct reflection of the distribution of sampling (i.e. a species might not be detected but a record is still made of the sampling event), therefore such data provide reliable information on absences. If the basis of sampling is known (e.g. random, systematic-random, stratified-random etc.), then data may be representative, at least within the bounds of the original survey; however, even in this situation, such a sample may still be unrepresentative of an analyst's target population (perhaps only part of the geographical area is covered, or some environmental gradient was not targeted by the original scheme design). We can see very few scenarios, then, where it will not be possible to at least approximate the degree to which a dataset is representative of a given population using all the likely available knowledge that could be brought to bear on a question. Indeed, this is the ultimate rationale behind qualitative risk-of-bias tools based on expert assessments (Table 1). In the final instance, if analysts cannot reach an informed conclusion with regards to the likely representativeness of a sample, then broader inference is not likely to be meaningful; in this case simple descriptive (i.e. non-inferential) statistics could be used, and this limitation acknowledged, with paper titles, abstracts, conclusions etc. all clearly reflecting this situation.

Four of the questions in ROBITT provide researchers with an opportunity to think about whether and how they can mitigate biases revealed elsewhere in the tool. It is not possible to review all possible measures that could be taken by researchers here; indeed, a full treatment of adjustments and models for dealing with bias would have to cover many topics within statistics and ecological data management. However, we note here three general approaches to bias mitigation. The first is to modify the data in some way (e.g. spatial, temporal or environmental thinning; Inman et al. 2021). The second is model the biases; typically, this will involve incorporation of variables thought to capture the biasing mechanism in some form of regression analysis (e.g. van Strien et al., 2019), although other approaches are possible (Ahmad Suhaimi et al., 2021). Finally, we suspect that in many cases ROBITT will reveal the need to restrict the desired extent of researcher's inferences. This might include, for example, redefining the spatial extent of an analysis to reflect the fact that data are scarcely available in some portion of geographic space, or coarsening the temporal resolution to "smooth over" temporal biases in geographic or taxonomic coverage (Pescott et al., 2019). Of course, any modifications to the extents of the statistical population should be reflected in paper titles and abstracts, etc. We suspect that using some, or all three, of the general bias mitigation strategies outlined above, researchers will usually be able to proceed with their analyses, even if those analyses relate to more limited statistical populations than initially envisioned.

Bias mitigation strategies may also require sensitivity analyses, particularly if the assessed potential for bias in a domain is uncertain: statistical fixes may make inference worse. For example, thinning datasets obviously removes information (this choice may therefore depend on model aims—if prediction rather than description is key, then exploiting areas with more information using partial pooling may be better than thinning; Gelman, 2006; weighting can increase estimate variance Gelman, 2007; and parameters in more complicated regression models purporting to deal with bias may be unidentifiable (Lele, 2010). The problem of inference from biased samples is, unfortunately, rather difficult, and quick fixes do not exist. Given this, we conclude with a recognition that the version of ROBITT presented in this paper can likely be improved in the future, and we welcome feedback from all who use it. We expect to publish periodic updates to the tool in order to increase its usefulness and clarity.

References

Ahmad Suhaimi, S.S., Blair, G.S., Jarvis, S.G., 2021. Integrated species distribution models: A comparison of approaches under different data quality scenarios. Divers. Distrib. 1–10. https://doi.org/10.1111/ddi.13255

Anon., 2020. PubPeer comments on Soroye et al. 2019. [WWW Document]. URL https://pubpeer.com/publications/B2A01E04EBC41DAD2C767B2111972C#1 (accessed 10.8.21).

Ball-Damerow, J.E., Brenskelle, L., Barve, N., Soltis, P.S., Sierwald, P., Bieler, R., LaFrance, R., Ariño, A.H., Guralnick, R., 2019. Research applications of primary biodiversity databases in the digital age. PLoS One 1–26. https://doi.org/10.1101/605071

Barends, J.M., Pietersen, D.W., Zambatis, G., Tye, D.R.C., Maritz, B., 2020. Sampling bias in reptile occurrence data for the Kruger National Park. Koedoe 62, 1–9. https://doi.org/10.4102/koedoe.v62i1.1579

Barnett, V., 1982. Comparative Statistical Inference., 2nd ed. Wiley, Chichester.

Belitz, M.W., 2021. A complete inventory of North American butterfly occurrence data: narrowing data gaps, but increasing bias 1–11. https://doi.org/10.1111/ecog.05396

Bero, L., Chartres, N., Diong, J., Fabbri, A., Ghersi, D., Lam, J., Lau, A., McDonald, S., Mintzes, B., Sutton, P., Turton, J.L., Woodruff, T.J., 2018. The risk of bias in observational studies of exposures (ROBINS-E) tool: Concerns arising from application to observational studies of exposures. Syst. Rev. 7, 1–11. https://doi.org/10.1186/s13643-018-0915-2

Boyd, R.J., Powney, G., Carvell, C., Pescott, O.L., 2021. occAssess: An R package for assessing potential biases in species occurrence data. Ecol. Evol. https://doi.org/10.1002/ece3.8299

Buckland, S.T., Johnston, A., 2017. Monitoring the biodiversity of regions: Key principles and possible pitfalls. Biol. Conserv. 214, 23–34. https://doi.org/10.1016/j.biocon.2017.07.034

Cardinale, B.J., Gonzalez, A., Allington, G.R.H., Loreau, M., 2018. Is local biodiversity declining or not? A summary of the debate over analysis of species richness time trends. Biol. Conserv. 219, 175–183. https://doi.org/10.1016/j.biocon.2017.12.021

Clark, P., Evans, F., 1954. Distance to Nearest Neighbour as a Measure of Spatial Relationships in Populations. Ecology 35, 445–453. https://doi.org/10.1007/BF02315373

Collen, B., Pettorelli, N., Baillie, J., Durant, S., 2013. Biodiversity Monitoring and Conservation: Bridging the Gap Between Global Commitment and Local Action - Conservation Science and Practice Series. Wiley.

Crossley, M.S., Meier, A.R., Baldwin, E.M., Berry, L.L., Crenshaw, L.C., Hartman, G.L., Lagos-Kutz, D., Nichols, D.H., Patel, K., Varriano, S., Snyder, W.E., Moran, M.D., 2020. No net insect abundance and diversity declines across US Long Term Ecological Research sites. Nat. Ecol. Evol. 4, 1368–1376. https://doi.org/10.1038/s41559-020-1269-4

Dengler, J., Jansen, F., Glöckler, F., Peet, R.K., de Cáceres, M., Chytrý, M., Ewald, J., Oldeland, J., Lopez-Gonzalez, G., Finckh, M., Mucina, L., Rodwell, J.S., Schaminée, J.H.J., Spencer, N., 2011. The Global Index of Vegetation-Plot Databases (GIVD): A new resource for vegetation science. J. Veg. Sci. 22, 582–597. https://doi.org/10.1111/j.1654-1103.2011.01265.x

Dennis, E.B., Brereton, T.M., Morgan, B.J.T., Fox, R., Shortall, C.R., Prescott, T., Foster, S., 2019. Trends and indicators for quantifying moth abundance and occupancy in Scotland. J. Insect Conserv. 23, 369–380. https://doi.org/10.1007/s10841-019-00135-z

Desquilbet, M., Cornillon, P.A., Gaume, L., Bonmatin, J.M., 2021. Adequate statistical modelling and data selection are essential when analysing abundance and diversity trends. Nat. Ecol. Evol. 5, 592–594. https://doi.org/10.1038/s41559-021-01427-x

Desquilbet, M., Gaume, L., Grippa, M., Céréghino, R., Jean-François, H., Bonmatin, J.-M., Cornillon, P.-A., Maes, D., Van Dyck, H., Goulson, D., 2020. Comment on “Meta-analysis reveals declines in terrestrial but increases in freshwater insect abundances.” Science (80-.). https://doi.org/10.1126/science.abd8947

Dornelas, M., Antão, L.H., Moyes, F., Bates, A.E., Magurran, A.E., Adam, D., Akhmetzhanova, A.A., Appeltans, W., Arcos, J.M., Arnold, H., Ayyappan, N., Badihi, G., Baird, A.H., Barbosa, M., Barreto, T.E., Bässler, C., Bellgrove, A., Belmaker, J., Benedetti-Cecchi, L., Bett, B.J., Bjorkman, A.D., Błażewicz, M., Blowes, S.A., Bloch, C.P., Bonebrake, T.C., Boyd, S., Bradford, M., Brooks, A.J., Brown, J.H., Bruelheide, H., Budy, P., Carvalho, F., Castañeda-Moya, E., Chen, C.A., Chamblee, J.F., Chase, T.J., Siegwart Collier, L., Collinge, S.K., Condit, R., Cooper, E.J., Cornelissen, J.H.C., Cotano, U., Kyle Crow, S., Damasceno, G., Davies, C.H., Davis, R.A., Day, F.P., Degraer, S., Doherty, T.S., Dunn, T.E., Durigan, G., Duffy, J.E., Edelist, D., Edgar, G.J., Elahi, R., Elmendorf, S.C., Enemar, A., Ernest, S.K.M., Escribano, R., Estiarte, M., Evans, B.S., Fan, T.Y., Turini Farah, F., Loureiro Fernandes, L., Farneda, F.Z., Fidelis, A., Fitt, R., Fosaa, A.M., Daher Correa Franco, G.A., Frank, G.E., Fraser, W.R., García, H., Cazzolla Gatti, R., Givan, O., Gorgone-Barbosa, E., Gould, W.A., Gries, C., Grossman, G.D., Gutierréz, J.R., Hale, S., Harmon, M.E., Harte, J., Haskins, G., Henshaw, D.L., Hermanutz, L., Hidalgo, P., Higuchi, P., Hoey, A., Van Hoey, G., Hofgaard, A., Holeck, K., Hollister, R.D., Holmes, R., Hoogenboom, M., Hsieh, C. hao, Hubbell, S.P., Huettmann, F., Huffard, C.L., Hurlbert, A.H., Macedo Ivanauskas, N., Janík, D., Jandt, U., Jażdżewska, A., Johannessen, T., Johnstone, J., Jones, J., Jones, F.A.M., Kang, J., Kartawijaya, T., Keeley, E.C., Kelt, D.A., Kinnear, R., Klanderud, K., Knutsen, H., Koenig, C.C., Kortz, A.R., Král, K., Kuhnz, L.A., Kuo, C.Y., Kushner, D.J., Laguionie-Marchais, C., Lancaster, L.T., Min Lee, C., Lefcheck, J.S., Lévesque, E., Lightfoot, D., Lloret, F., Lloyd, J.D., López-Baucells, A., Louzao, M., Madin, J.S., Magnússon, B., Malamud, S., Matthews, I., McFarland, K.P., McGill, B., McKnight, D., McLarney, W.O., Meador, J., Meserve, P.L., Metcalfe, D.J., Meyer, C.F.J., Michelsen, A., Milchakova, N., Moens, T., Moland, E., Moore, J., Mathias Moreira, C., Müller, J., Murphy, G., Myers-Smith, I.H., Myster, R.W., Naumov, A., Neat, F., Nelson, J.A., Paul Nelson, M., Newton, S.F., Norden, N., Oliver, J.C., Olsen, E.M., Onipchenko, V.G., Pabis, K., Pabst, R.J., Paquette, A., Pardede, S., Paterson, D.M., Pélissier, R., Peñuelas, J., Pérez-Matus, A., Pizarro, O., Pomati, F., Post, E., Prins, H.H.T., Priscu, J.C., Provoost, P., Prudic, K.L., Pulliainen, E., Ramesh, B.R., Mendivil Ramos, O., Rassweiler, A., Rebelo, J.E., Reed, D.C., Reich, P.B., Remillard, S.M., Richardson, A.J., Richardson, J.P., van Rijn, I., Rocha, R., Rivera-Monroy, V.H., Rixen, C., Robinson, K.P., Ribeiro Rodrigues, R., de Cerqueira Rossa-Feres, D., Rudstam, L., Ruhl, H., Ruz, C.S., Sampaio, E.M., Rybicki, N., Rypel, A., Sal, S., Salgado, B., Santos, F.A.M., Savassi-Coutinho, A.P., Scanga, S., Schmidt, J., Schooley, R., Setiawan, F., Shao, K.T., Shaver, G.R., Sherman, S., Sherry, T.W., Siciński, J., Sievers, C., da Silva, A.C., Rodrigues da Silva, F., Silveira, F.L., Slingsby, J., Smart, T., Snell, S.J., Soudzilovskaia, N.A., Souza, G.B.G., Maluf Souza, F., Castro Souza, V., Stallings, C.D., Stanforth, R., Stanley, E.H., Mauro Sterza, J., Stevens, M., Stuart-Smith, R., Rondon Suarez, Y., Supp, S., Yoshio Tamashiro, J., Tarigan, S., Thiede, G.P., Thorn, S., Tolvanen, A., Teresa Zugliani Toniato, M., Totland, Ø., Twilley, R.R., Vaitkus, G., Valdivia, N., Vallejo, M.I., Valone, T.J., Van Colen, C., Vanaverbeke, J., Venturoli, F., Verheye, H.M., Vianna, M., Vieira, R.P., Vrška, T., Quang Vu, C., Van Vu, L., Waide, R.B., Waldock, C., Watts, D., Webb, S., Wesołowski, T., White, E.P., Widdicombe, C.E., Wilgers, D., Williams, R., Williams, S.B., Williamson, M., Willig, M.R., Willis, T.J., Wipf, S., Woods, K.D., Woehler, E.J., Zawada, K., Zettler, M.L., 2018. BioTIME: A database of biodiversity time series for the Anthropocene. Glob. Ecol. Biogeogr. 27, 760–786. https://doi.org/10.1111/geb.12729

Dornelas, M., Gotelli, N.J., McGill, B., Shimadzu, H., Moyes, F., Sievers, C., Magurran, A.E., 2014. Assemblage time series reveal biodiversity change but not systematic loss. Science (80-.). 344, 296–299. https://doi.org/10.1126/science.1248484

Dudík, M., Schapire, R.E., Phillips, S.J., 2005. Correcting sample selection bias in maximum entropy density estimation. Adv. Neural Inf. Process. Syst. 323–330.

Gelman, A., 2007. Struggles with survey weighting and regression modeling. Stat. Sci. 22, 153–164. https://doi.org/10.1214/088342306000000691

Gelman, A., 2006. Multilevel (hierarchical) modeling: What It can and cannot do. Technometrics 48, 432–435. https://doi.org/10.1198/004017005000000661

GEO BON Secretariat, 2015. Global Biodiversity Change Indicators Version 1.2 1–20.

Gonzalez, A., Cardinale, B.J., Allington, G.R.H., Byrnes, J., Endsley, K.A., Brown, D.G., Hooper, D.U., Isbell, F., O'Connor, M.I., Loreau, M., 2016. Estimating local biodiversity change: A critique of papers claiming no net loss of local diversity. Ecology 97, 1949–1960. https://doi.org/10.1890/15-1759.1

Gregory, R.D., Van Strien, A., Vorisek, P., Meyling, A.W.G., Noble, D.G., Foppen, R.P.B., Gibbons, D.W., 2005. Developing indicators for European birds. Philos. Trans. R. Soc. B Biol. Sci. 360, 269–288. https://doi.org/10.1098/rstb.2004.1602

Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., Goss-Custard, J., Grand, T., Heinz, S.K., Huse, G., Huth, A., Jepsen, J.U., Jørgensen, C., Mooij, W.M., Müller, B., Pe'er, G., Piou, C., Railsback, S.F., Robbins, A.M., Robbins, M.M., Rossmanith, E., Rüger, N., Strand, E., Souissi, S., Stillman, R. a., Vabø, R., Visser, U., DeAngelis, D.L., 2006. A standard protocol for describing individual-based and agent-based models. Ecol. Modell. 198, 115–126. https://doi.org/10.1016/j.ecolmodel.2006.04.023

Grimm, V., Berger, U., DeAngelis, D.L., Polhill, J.G., Giske, J., Railsback, S.F., 2010. The ODD protocol: A review and first update. Ecol. Modell. 221, 2760–2768. https://doi.org/10.1016/j.ecolmodel.2010.08.019

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. Biol. Conserv. 257, 109141. https://doi.org/10.1016/j.biocon.2021.109141

Higgins, J.P.T., Altman, D.G., Gøtzsche, P.C., Jüni, P., Moher, D., Oxman, A.D., Savović, J., Schulz, K.F., Weeks, L., Sterne, J.A.C., 2011. The Cochrane Collaboration’s tool for assessing risk of bias in randomised trials. BMJ 343, 1–9. https://doi.org/10.1136/bmj.d5928

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., Day, J., Echeverría-Londoño, S., Emerson, S.R., Gao, D., Garon, M., Harrison, M.L.K., Ingram, D.J., Jung, M., Kemp, V., Kirkpatrick, L., Martin, C.D., Pan, Y., Pask-Hale, G.D., Pynegar, E.L., Robinson, A.N., Sanchez-Ortiz, K., Senior, R.A., Simmons, B.I., White, H.J., Zhang, H., Aben, J., Abrahamczyk, S., Adum, G.B., Aguilar-Barquero, V., Aizen, M.A., Albertos, B., Alcala, E.L., del Mar Alguacil, M., Alignier, A., Ancrenaz, M., Andersen, A.N., Arbeláez-Cortés, E., Armbrecht, I., Arroyo-Rodríguez, V., Aumann, T., Axmacher, J.C., Azhar, B., Azpiroz, A.B., Baeten, L., Bakayoko, A., Báldi, A., Banks, J.E., Baral, S.K., Barlow, J., Barratt, B.I.P., Barrico, L., Bartolommei, P., Barton, D.M., Basset, Y., Batáry, P., Bates, A.J., Baur, B., Bayne, E.M., Beja, P., Benedick, S., Berg, Å., Bernard, H., Berry, N.J., Bhatt, D., Bicknell, J.E., Bihn, J.H., Blake, R.J., Bobo, K.S., Bóçon, R., Boekhout, T., Böhning-Gaese, K., Bonham, K.J., Borges, P.A.V., Borges, S.H., Boutin, C., Bouyer, J., Bragagnolo, C., Brandt, J.S., Brearley, F.Q., Brito, I., Bros, V., Brunet, J., Buczkowski, G., Buddle, C.M., Bugter, R., Buscardo, E., Buse, J., Cabra-García, J., Cáceres, N.C., Cagle, N.L., Calviño-Cancela, M., Cameron, S.A., Cancello, E.M., Caparrós, R., Cardoso, P., Carpenter, D., Carrijo, T.F., Carvalho, A.L., Cassano, C.R., Castro, H., Castro-Luna, A.A., Rolando, C.B., Cerezo, A., Chapman, K.A., Chauvat, M., Christensen, M., Clarke, F.M., Cleary, D.F.R., Colombo, G., Connop, S.P., Craig, M.D., Cruz-López, L., Cunningham, S.A., D’Aniello, B., D’Cruze, N., da Silva, P.G., Dallimer, M., Danquah, E., Darvill, B., Dauber, J., Davis, A.L.V., Dawson, J., de Sassi, C., de Thoisy, B., Deheuvels, O., Dejean, A., Devineau, J.L., Diekötter, T., Dolia, J. V., Domínguez, E., Dominguez-Haydar, Y., Dorn, S., Draper, I., Dreber, N., Dumont, B., Dures, S.G., Dynesius, M., Edenius, L., Eggleton, P., Eigenbrod, F., Elek, Z., Entling, M.H., Esler, K.J., de Lima, R.F., Faruk, A., Farwig, N., Fayle, T.M., Felicioli, A., Felton, A.M., Fensham, R.J., Fernandez, I.C., Ferreira, C.C., Ficetola, G.F., Fiera, C., Filgueiras, B.K.C., Fırıncıoğlu, H.K., Flaspohler, D., Floren, A., Fonte, S.J., Fournier, A., Fowler, R.E., Franzén, M., Fraser, L.H., Fredriksson, G.M., Freire, G.B., Frizzo, T.L.M., Fukuda, D., Furlani, D., Gaigher, R., Ganzhorn, J.U., García, K.P., Garcia-R, J.C., Garden, J.G., Garilleti, R., Ge, B.M., Gendreau-Berthiaume, B., Gerard, P.J., Gheler-Costa, C., Gilbert, B., Giordani, P., Giordano, S., Golodets, C., Gomes, L.G.L., Gould, R.K., Goulson, D., Gove, A.D., Granjon, L., Grass, I., Gray, C.L., Grogan, J., Gu, W., Guardiola, M., Gunawardene, N.R., Gutierrez, A.G., Gutiérrez-Lamus, D.L., Haarmeyer, D.H., Hanley, M.E., Hanson, T., Hashim, N.R., Hassan, S.N., Hatfield, R.G., Hawes, J.E., Hayward, M.W., Hébert, C., Helden, A.J., Henden, J.A., Henschel, P., Hernández, L., Herrera, J.P., Herrmann, F., Herzog, F., Higuera-Diaz, D., Hilje, B., Höfer, H., Hoffmann, A., Horgan, F.G., Hornung, E., Horváth, R., Hylander, K., Isaacs-Cubides, P., Ishida, H., Ishitani, M., Jacobs, C.T., Jaramillo, V.J., Jauker, B., Hernández, F.J., Johnson, M.F., Jolli, V., Jonsell, M., Juliani, S.N., Jung, T.S., Kapoor, V., Kappes, H., Kati, V., Katovai, E., Kellner, K., Kessler, M., Kirby, K.R., Kittle, A.M., Knight, M.E., Knop, E., Kohler, F., Koivula, M., Kolb, A., Kone, M., Kőrösi, Á., Krauss, J., Kumar, A., Kumar, R., Kurz, D.J., Kutt, A.S., Lachat, T., Lantschner, V., Lara, F., Lasky, J.R., Latta, S.C., Laurance, W.F., Lavelle, P., Le Féon, V., LeBuhn, G., Légaré, J.P., Lehouck, V., Lencinas, M. V., Lentini, P.E., Letcher, S.G., Li, Q., Litchwark, S.A., Littlewood, N.A., Liu, Y., Lo-Man-Hung, N., López-Quintero, C.A., Louhaichi, M., Lövei, G.L., Lucas-Borja, M.E., Luja, V.H., Luskin, M.S., MacSwiney G, M.C., Maeto, K., Magura, T., Mallari, N.A., Malone, L.A., Malonza, P.K., Malumbres-Olarte, J., Mandujano, S., Måren, I.E., Marin-Spiotta, E., Marsh, C.J., Marshall, E.J.P., Martínez, E., Martínez Pastur, G., Moreno Mateos, D., Mayfield, M.M., Mazimpaka, V., McCarthy, J.L., McCarthy, K.P., McFrederick, Q.S., McNamara, S., Medina, N.G., Medina, R., Mena, J.L., Mico, E., Mikusinski, G., Milder, J.C., Miller, J.R., Miranda-Esquivel, D.R., Moir, M.L., Morales, C.L., Muchane, M.N., Muchane, M., Mudri-Stojnic, S., Munira, A.N., Muoñz-Alonso, A., Munyekenye, B.F., Naidoo, R., Naithani, A., Nakagawa, M., Nakamura, A., Nakashima, Y., Naoe, S., Nates-Parra, G., Navarrete Gutierrez, D.A., Navarro-Iriarte, L., Ndang'ang'a, P.K., Neuschulz, E.L., Ngai, J.T., Nicolas, V., Nilsson, S.G., Noreika, N., Norfolk, O., Noriega, J.A., Norton, D.A., Nöske, N.M., Nowakowski, A.J., Numa, C., O'Dea, N., O'Farrell, P.J., Oduro, W., Oertli, S., Ofori-Boateng, C., Oke, C.O., Oostra, V., Osgathorpe, L.M., Otavo, S.E., Page, N. V., Paritsis, J., Parra-H, A., Parry, L., Pe'er, G., Pearman, P.B., Pelegrin, N., Pélissier, R., Peres, C.A., Peri, P.L., Persson, A.S., Petanidou, T., Peters, M.K., Pethiyagoda, R.S., Phalan, B., Philips, T.K., Pillsbury, F.C., Pincheira-Ulbrich, J., Pineda, E., Pino, J., Pizarro-Araya, J., Plumptre, A.J., Poggio, S.L., Politi, N., Pons, P., Poveda, K., Power, E.F., Presley, S.J., Proença, V., Quaranta, M., Quintero, C., Rader, R., Ramesh, B.R., Ramirez-Pinilla, M.P., Ranganathan, J., Rasmussen, C., Redpath-Downing, N.A., Reid, J.L., Reis, Y.T., Rey Benayas, J.M., Rey-Velasco, J.C., Reynolds, C., Ribeiro, D.B., Richards, M.H., Richardson, B.A., Richardson, M.J., Ríos, R.M., Robinson, R., Robles, C.A., Römbke, J., Romero-Duque, L.P., Rös, M., Rosselli, L., Rossiter, S.J., Roth, D.S., Roulston, T.H., Rousseau, L., Rubio, A. V., Ruel, J.C., Sadler, J.P., Sáfián, S., Saldaña-Vázquez, R.A., Sam, K., Samnegård, U., Santana, J., Santos, X., Savage, J., Schellhorn, N.A., Schilthuizen, M., Schmiedel, U., Schmitt, C.B., Schon, N.L., Schüepp, C., Schumann, K., Schweiger, O., Scott, D.M., Scott, K.A., Sedlock, J.L., Seefeldt, S.S., Shahabuddin, G., Shannon, G., Sheil, D., Sheldon, F.H., Shochat, E., Siebert, S.J., Silva, F.A.B., Simonetti, J.A., Slade, E.M., Smith, J., Smith-Pardo, A.H., Sodhi, N.S., Somarriba, E.J., Sosa, R.A., Soto Quiroga, G., St-Laurent, M.H., Starzomski, B.M., Stefanescu, C., Steffan-Dewenter, I., Stouffer, P.C., Stout, J.C., Strauch, A.M., Struebig, M.J., Su, Z., Suarez-Rubio, M., Sugiura, S., Summerville, K.S., Sung, Y.H., Sutrisno, H., Svenning, J.C., Teder, T., Threlfall, C.G., Tiitsaar, A., Todd, J.H., Tonietto, R.K., Torre, I., Tóthmérész, B., Tscharntke, T., Turner, E.C., Tylianakis, J.M., Uehara-Prado, M., Urbina-Cardona, N., Vallan, D., Vanbergen, A.J., Vasconcelos, H.L., Vassilev, K., Verboven, H.A.F., Verdasca, M.J., Verdú, J.R., Vergara, C.H., Vergara, P.M., Verhulst, J., Virgilio, M., Vu, L. Van, Waite, E.M., Walker, T.R., Wang, H.F., Wang, Y., Watling, J.I., Weller, B., Wells, K., Westphal, C., Wiafe, E.D., Williams, C.D., Willig, M.R., Woinarski, J.C.Z., Wolf, J.H.D., Wolters, V., Woodcock, B.A., Wu, J., Wunderle, J.M., Yamaura, Y., Yoshikura, S., Yu, D.W., Zaitsev, A.S., Zeidler, J., Zou, F., Collen, B., Ewers, R.M., Mace, G.M., Purves, D.W., Scharlemann, J.P.W., Purvis, A., 2017. The database of the PREDICTS (Projecting Responses of Ecological Diversity In Changing Terrestrial Systems) project. Ecol. Evol. 7, 145–188. https://doi.org/10.1002/ece3.2579

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 (Cop.). 1–11. https://doi.org/10.1111/ecog.05926

Inman, R., Franklin, J., Esque, T., Nussear, K., 2021. Comparing sample bias correction methods for species distribution modeling using virtual species. Ecosphere 12. https://doi.org/10.1002/ecs2.3422

Jähnig, S.C., Baranov, V., Altermatt, F., Cranston, P., Friedrichs-Manthey, M., Geist, J., He, F., Heino, J., Hering, D., Hölker, F., Jourdan, J., Kalinkat, G., Kiesel, J., Leese, F., Maasri, A., Monaghan, M.T., Schäfer, R.B., Tockner, K., Tonkin, J.D., Domisch, S., 2021. Revisiting global trends in freshwater insect biodiversity. Wiley Interdiscip. Rev. Water 8, 1–5. https://doi.org/10.1002/wat2.1506

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., Turak, E., 2019. Essential biodiversity variables for mapping and monitoring species populations. Nat. Ecol. Evol. 3, 539–551. https://doi.org/10.1038/s41559-019-0826-1

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.

Kim, S.Y., Park, J.E., Lee, Y.J., Seo, H.J., Sheen, S.S., Hahn, S., Jang, B.H., Son, H.J., 2013. Testing a tool for assessing the risk of bias for nonrandomized studies showed moderate reliability and promising validity. J. Clin. Epidemiol. 66, 408–414. https://doi.org/10.1016/j.jclinepi.2012.09.016

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. Biol. Rev. 93, 600–625. https://doi.org/10.1111/brv.12359

Lehmann, E., 1959. Testing Statistical Hypotheses. John Wiley & Sons Inc., New York.

Lele, S.R., 2010. Model complexity and information in the data: Could it be a house built on sand? Ecology 91, 3493–3496. https://doi.org/10.1890/10-0099.1

Lister, B.C., Garcia, A., 2018. Climate-driven declines in arthropod abundance restructure a rainforest food web. Proc. Natl. Acad. Sci. U. S. A. 115, E10397–E10406. https://doi.org/10.1073/pnas.1722477115

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. Ann. Appl. Stat. 12, 685–726. https://doi.org/10.1214/18-AOAS1161SF

Meyer, C., Weigelt, P., Kreft, H., 2016. Multidimensional biases, gaps and uncertainties in global plant occurrence information. Ecol. Lett. 19, 992–1006. https://doi.org/10.1111/ele.12624

Navarro, L.M., Fernández, N., Guerra, C., Guralnick, R., Kissling, W.D., Londoño, M.C., Muller-Karger, F., Turak, E., Balvanera, P., Costello, M.J., Delavaud, A., El Serafy, G.Y., Ferrier, S., Geijzendorffer, I., Geller, G.N., Jetz, W., Kim, E.S., Kim, H.J., Martin, C.S., McGeoch, M.A., Mwampamba, T.H., Nel, J.L., Nicholson, E., Pettorelli, N., Schaepman, M.E., Skidmore, A., Sousa Pinto, I., Vergara, S., Vihervaara, P., Xu, H., Yahara, T., Gill, M., Pereira, H.M., 2017. Monitoring biodiversity change through effective global coordination. Curr. Opin. Environ. Sustain. 29, 158–169. https://doi.org/10.1016/j.cosust.2018.02.005

Nelson, G., Ellis, S., 2019. The history and impact of digitization and digital data mobilization on biodiversity research. Philos. Trans. R. Soc. B Biol. Sci. 374, 2–10. https://doi.org/10.1098/rstb.2017.0391

Outhwaite, C.L., 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., Beckmann, B.C., Cook, T., Flanagan, J., Fowles, A., Hammond, P., Harvey, P., 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) 1–12. https://doi.org/10.5285/0ec7e549-57d4-4e2d-b2d3-2199e1578d84

Page, L.M., Macfadden, B.J., Fortes, J.A., Soltis, P.S., Riccardi, G., 2015. Digitization of Biodiversity Collections Reveals Biggest Data on Biodiversity. Bioscience 65, 841–842. https://doi.org/10.1093/biosci/biv104

Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., Chou, R., Glanville, J., Grimshaw, J.M., Hróbjartsson, A., Lalu, M.M., Li, T., Loder, E.W., Mayo-Wilson, E., McDonald, S., McGuinness, L.A., Stewart, L.A., Thomas, J., Tricco, A.C., Welch, V.A., Whiting, P., Moher, D., 2021a. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ 372. https://doi.org/10.1136/bmj.n71

Page, M.J., Moher, D., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., Chou, R., Glanville, J., Grimshaw, J.M., Hróbjartsson, A., Lalu, M.M., Li, T., Loder, E.W., Mayo-Wilson, E., Mcdonald, S., Mcguinness, L.A., Stewart, L.A., Thomas, J., Tricco, A.C., Welch, V.A., Whiting, P., Mckenzie, J.E., 2021b. PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. BMJ 372. https://doi.org/10.1136/bmj.n160

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., Heip, C., Höft, R., Hurtt, G., Jetz, W., Karp, D.S., McGeoch, M.A., Obura, D., Onoda, Y., Pettorelli, N., Reyers, B., Sayre, R., Scharlemann, J.P.W., Stuart, S.N., Turak, E., Walpole, M., Wegmann, M., 2013. Essential biodiversity variables. Science (80-.). 339, 277–278. https://doi.org/10.1126/science.1229931

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? Br. Irish Bot. 1, 250–282. https://doi.org/10.33928/bib.2019.01.250

Peters, R.H., 1991. A critique for ecology. Cambridge University Press, Cambridge.

Petersen, T.K., Austrheim, G., Speed, J.D.M., Grøtan, V., 2021. Species data for understanding biodiversity dynamics : The what , where and when of species occurrence data collection 1–17. https://doi.org/10.1002/2688-8319.12048

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. Ecol. Appl. 19, 181–197. https://doi.org/10.1890/07-2153.1

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. Nat. Commun. 1–6. https://doi.org/10.1038/s41467-019-08974-9

Prendergast, J.. R.., Wood, S.. N.., Lawton, J.. H.., Eversham, B.. C.., 1993. Correcting for Variation in Recording Effort in Analyses of Diversity Hotspots. Biodivers. Lett. 1, 39–53.

Royle, J.A., 2006. Site occupancy models with heterogeneous detection probabilities. Biometrics 62, 97–102. https://doi.org/10.1111/j.1541-0420.2005.00439.x

Ruete, A., 2015. Displaying bias in sampling effort of data accessed from biodiversity databases using ignorance maps. Biodivers. Data J. 3. https://doi.org/10.3897/BDJ.3.e5361

Sánchez-Bayo, F., Wyckhuys, K.A.G., 2019. Worldwide decline of the entomofauna: A review of its drivers. Biol. Conserv. 232, 8–27. https://doi.org/10.1016/j.biocon.2019.01.020

Saunders, M.E., Janes, J.K., O'Hanlon, J.C., 2020. Moving on from the Insect Apocalypse Narrative: Engaging with Evidence-Based Insect Conservation. Bioscience 70, 80–89. https://doi.org/10.1093/biosci/biz143

Schmolke, A., Thorbek, P., DeAngelis, D.L., Grimm, V., 2010. Ecological models supporting environmental decision making: A strategy for the future. Trends Ecol. Evol. 25, 479–486. https://doi.org/10.1016/j.tree.2010.05.001

Schünemann, H., Brożek, J., Guyatt, G., Oxman, A., 2013. GRADE handbook.

Simmons, B.I., Balmford, A., Bladon, A.J., Christie, A.P., De Palma, A., Dicks, L. V., Gallego-Zamorano, J., Johnston, A., Martin, P.A., Purvis, A., Rocha, R., Wauchope, H.S., Wordley, C.F.R., Worthington, T.A., Finch, T., 2019. Worldwide insect declines: An important message, but interpret with caution. Ecol. Evol. 9, 3678–3680. https://doi.org/10.1002/ece3.5153

Simons, D.J., Shoda, Y., Lindsay, D.S., 2017. Constraints on Generality (COG): A Proposed Addition to All Empirical Papers. Perspect. Psychol. Sci. 12, 1123–1128. https://doi.org/10.1177/1745691617708630

Soroye, P., Newbold, T., Kerr, J., 2020. Among Bumble Bees Across Continents. Science (80-.). 367, 685–688.

Spear, D.M., Pauly, G.B., Kaiser, K., 2017. Citizen science as a tool for augmenting museum collection data from urban areas. Front. Ecol. Evol. 5, 1–12. https://doi.org/10.3389/fevo.2017.00086

Speed, J.D.M., Bendiksby, M., Finstad, A.G., Hassel, K., Kolstad, A.L., Prestø, T., 2018. Contrasting spatial, temporal and environmental patterns in observation and specimen based species occurrence data. PLoS One 13, 1–17. https://doi.org/10.1371/journal.pone.0196417

Sterne, J.A., Hernán, M.A., Reeves, B.C., Savović, J., Berkman, N.D., Viswanathan, M., Henry, D., Altman, D.G., Ansari, M.T., Boutron, I., Carpenter, J.R., Chan, A.W., Churchill, R., Deeks, J.J., Hróbjartsson, A., Kirkham, J., Jüni, P., Loke, Y.K., Pigott, T.D., Ramsay, C.R., Regidor, D., Rothstein, H.R., Sandhu, L., Santaguida, P.L., Schünemann, H.J., Shea, B., Shrier, I., Tugwell, P., Turner, L., Valentine, J.C., Waddington, H., Waters, E., Wells, G.A., Whiting, P.F., Higgins, J.P., 2016. ROBINS-I: A tool for assessing risk of bias in non-randomised studies of interventions. BMJ 355, 4–10. https://doi.org/10.1136/bmj.i4919

Sterne, J.A.C., Savović, J., Page, M.J., Elbers, R.G., Blencowe, N.S., Boutron, I., Cates, C.J., Cheng, H.Y., Corbett, M.S., Eldridge, S.M., Emberson, J.R., Hernán, M.A., Hopewell, S., Hróbjartsson, A., Junqueira, D.R., Jüni, P., Kirkham, J.J., Lasserson, T., Li, T., McAleenan, A., Reeves, B.C., Shepperd, S., Shrier, I., Stewart, L.A., Tilling, K., White, I.R., Whiting, P.F., Higgins, J.P.T., 2019. RoB 2: A revised tool for assessing risk of bias in randomised trials. BMJ 366, 1–8. https://doi.org/10.1136/bmj.l4898

Sullivan, B.L., Aycrigg, 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., Hochachka, W.M., Iliff, M.J., Lagoze, C., La Sorte, F.A., Merrifield, M., Morris, W., Phillips, T.B., Reynolds, M., Rodewald, A.D., Rosenberg, K. V., Trautmann, N.M., Wiggins, A., Winkler, D.W., Wong, W.K., Wood, C.L., Yu, J., Kelling, S., 2014. The eBird enterprise: An integrated approach to development and application of citizen science. Biol. Conserv. 169, 31–40. https://doi.org/10.1016/j.biocon.2013.11.003

Sumner, S., Bevan, P., Hart, A.G., Isaac, N.J.B., 2019. Mapping species distributions in 2 weeks using citizen science. Insect Conserv. Divers. 12, 382–388. https://doi.org/10.1111/icad.12345

Troudet, J., Vignes-Lebbe, R., Grandcolas, P., Legendre, F., 2018. The increasing disconnection of primary biodiversity data from specimens: How does it happen and how to handle it? Syst. Biol. 67, 1110–1119. https://doi.org/10.1093/sysbio/syy044

van Klink, R., Bowler, D., Gongalsky, K., Swengel, A., Gentile, A., Chase, J., 2020. Meta-analysis reveals declines in terrestrial but increases in freshwater insect abundances. Science (80-.). 368, 417–420. https://doi.org/10.1126/science.aax9931

van Strien, A.J., van Swaay, C.A.M., van Strien-van Liempt, W.T.F.H., Poot, M.J.M., WallisDeVries, M.F., 2019. Over a century of data reveal more than 80% decline in butterflies in the Netherlands. Biol. Conserv. 234, 116–122. https://doi.org/10.1016/j.biocon.2019.03.023

Vellend, M., Baeten, L., Myers-Smith, I.H., Elmendorf, S.C., Beauséjour, R., Brown, C.D., De Frenne, P., Verheyen, K., Wipf, S., 2013. Global meta-analysis reveals no net change in local-scale plant biodiversity over time. Proc. Natl. Acad. Sci. U. S. A. 110, 19456–19459. https://doi.org/10.1073/pnas.1312779110

Willig, M.R., Woolbright, L., Presley, S.J., Schowalter, T.D., Waide, R.B., Heartsill Scalley, T., Zimmerman, J.K., González, G., Lugo, A.E., 2019. Populations are not declining and food webs are not collapsing at the Luquillo Experimental Forest. Proc. Natl. Acad. Sci. U. S. A. 116, 12143–12144. https://doi.org/10.1073/pnas.1820456116

Wolff, R.F., Moons, K.G.M., Riley, R.D., Whiting, P.F., Westwood, M., Collins, G.S., Reitsma, J.B., Kleijnen, J., Mallett, S., 2019. PROBAST: A tool to assess the risk of bias and applicability of prediction model studies. Ann. Intern. Med. 170, 51–58. https://doi.org/10.7326/M18-1376

Woodcock, B.A., Isaac, N.J.B., Bullock, J.M., Roy, D.B., Garthwaite, D.G., Crowe, A., Pywell, R.F., 2016. Impacts of neonicotinoid use on long-term population changes in wild bees in England. Nat. Commun. 7. https://doi.org/10.1038/ncomms12459

Zurell, D., Franklin, J., König, C., Bouchet, P.J., Dormann, C.F., Elith, J., Fandos, G., Feng, X., Guillera-Arroita, G., Guisan, A., Lahoz-Monfort, J.J., Leitão, P.J., Park, D.S., Peterson, A.T., Rapacciuolo, G., Schmatz, D.R., Schröder, B., Serra-Diaz, J.M., Thuiller, W., Yates, K.L., Zimmermann, N.E., Merow, C., 2020. A standard protocol for reporting species distribution models. Ecography (Cop.). 43, 1261–1277. https://doi.org/10.1111/ecog.04960