

Rethinking bias in biodiversity science: sources and solutions

Mingrui Li¹, Jonathan Bennie⁴, Richard ffrench-Constant¹, Richard Fox^{1,2}, David B. Roy^{1,3}, Emily B. Dennis^{2,5}, Robin J. Boyd¹

1 University of Exeter, Centre for Ecology and Conservation, Penryn Campus,
Penryn, Cornwall, TR10 9FE, United Kingdom

2 Butterfly Conservation, Manor Yard, East Lulworth, Wareham, Dorset, BH20
5QP, United Kingdom

3 UK Centre for Ecology and Hydrology, Benson Lane, Crowmarsh Gifford,
Oxfordshire, OX10 8BB, United Kingdom

4 University of Exeter, Centre for Geography and Environmental Science, Penryn
Campus, Penryn, Cornwall, TR10 9FE, United Kingdom

5 University of Kent, School of Computing, Canterbury, Kent, CT2 7NZ, United
Kingdom

Corresponding author emails: r.j.boyd@exeter.ac.uk; ml902@exeter.ac.uk

Abstract

Bias is increasingly a central concern in biodiversity science, yet the term is used inconsistently. A key source of confusion is the failure to distinguish clearly between the dimension in which bias manifests (e.g. spatial or taxonomic) and/or the stage of the observation process that gives rise to it. This distinction matters because the appropriate method for mitigating bias depends primarily on the latter rather than the former. We introduce a theory of observation for biodiversity science that classifies biases according to whether they arise during sampling or subsequent stages of the observation process (e.g. measurement, databasing). This classification is more informative than describing bias by its dimension alone because it directly identifies the mechanisms responsible and therefore the appropriate mitigation strategies. A key consequence of the framework is that an observation process may be a source of sampling or nonsampling bias depending on the scientific question being addressed.

Keywords: Citizen science; sampling bias; measurement error; estimand; bias correction

Introduction

Biodiversity scientists ask a wide range of questions about ecological systems. They may seek to determine where species might occur (Pimm et al., 2014, Bowler et al., 2022b, Chowdhury et al., 2026), how abundant they are (McGill et al., 2007, Justine et al., 2026), how their distributions or population sizes change over time (Bates and Bertelsmeier, 2025, Boulanger et al., 2025, Boyd et al., 2025), and which environmental or anthropogenic factors drive these patterns (Stolar and Nielsen, 2015, Li et al., 2026b). Other questions concern comparisons among sites, taxa, time periods, or survey methods, or environmental change (Sullivan et al., 2014, La Sorte and Somveille, 2020, Sánchez-Fernández et al., 2021).

To address a question quantitatively, it helps to codify its answer as a mathematical quantity whose value can, in principle, be estimated (Boyd et al., 2026a, Boyd et al., 2026b). In statistical parlance, this quantity is sometimes referred to as the estimand (from the Latin *aestimandus*, meaning “that which is being estimated” (Clark et al., 2022)). An estimand can be thought of as having three components: a target population, a unit-specific quantity and an aggregation rule (Lundberg et al., 2021). The target population is the set of units that is of interest, the unit-specific quantity is the attribute of those units under study, and the aggregation rule is a mathematical operation that summarises the unit-specific quantity in the target population (Table 1). For example, we might want to learn the mean (aggregation rule) abundance of a species (unit-specific quantity) across sites within a landscape (target population).

Table 1 Examples of biodiversity estimands and their three key components: target population, unit-specific quantity, and aggregation rule.

Study	Estimand	Target population	Unit-specific quantity	Aggregation rule
Li et al. (2025)	Rarity-weighted species richness at a given 1-km grid square	Butterfly species in Devon and Cornwall	Species presence at a focal site weighted by inverse recorded frequency	Sum across species
Newbold et al. (2015)	Local total abundance	Taxa within a focal site	Abundance of each taxon	Sum across taxa
Gregory et al. (2005)	National bird population index	Survey locations within a country	Abundance of a focal bird species at each location	Mean across locations

Having defined an estimand, the next step is to collect or assemble some data in the hope that they are informative about its value. In a perfect world, the unit-specific quantity would be measured without error for every unit in the target population. For example, if the estimand is the mean of a species' abundance

across sites within a landscape, then every individual would be observed and recorded at every site. Applying the aggregation rule to this ‘perfect’ dataset would yield the estimand. In practice, however, the analyst does not see the perfect dataset but rather an observed dataset that is produced by the observation process.

Following the Total Survey Error (TSE) framework (Groves and Lyberg, 2010), we split the observation process into sampling and nonsampling components. Sampling comes first and concerns the selection of a subset of units in the target population on which to collect data. A common example is the selection of a sample of sites at which to monitor species. Once a sample has been selected, the observer measures and records the unit-specific quantity on sampled units. These are nonsampling components of the observation process. The resulting dataset differs from the ‘perfect’ dataset because it reflects only a subset of the target population (Lohr, 2021), and, even for that subset, the unit-specific quantity may be measured with error (Buonaccorsi, 2010). Table 2 demonstrates the distinction between sampling and nonsampling observation processes using the UK Butterfly Monitoring Scheme as an example.

Table 2 Sources of sampling and nonsampling error in the UK Butterfly Monitoring Scheme (UKBMS). For each species of butterfly and year, the estimand is essentially the mean relative abundance across sites in the UK.

	Sampling	Nonsampling
Process	Site selection	Species detection and counting throughout the season Imperfect detection and
Source of error	Non-random site selection	measurement error (e.g. species are difficult to detect when they are not on the wing)
Correction method	Impute missing values for non-sampled sites (Boyd et al., 2025)	Account for seasonal variation in detectability using the Generalized Abundance Index method (Dennis et al., 2016)

Since the observed dataset differs from the perfect dataset, the estimand cannot be calculated directly. Instead, an estimator must be constructed that transforms the available data into an estimate of its value. One option is to apply the aggregation rule to the observed dataset as though it was the perfect one. For example, one might use the mean count of a species across sampled sites as an estimator of its mean abundance in the wider landscape. Generally, however, more sophisticated alternatives are preferred (e.g. Royle (2004)).

The discrepancy between the estimate produced by applying the estimator to the available data and the estimand is known as the estimation error. Both sampling and nonsampling components of the observation process can cause estimation error (Table 2), although the distinction is not always appreciated in biodiversity science (Aubry et al., 2026). Rather than the realised error given the data in hand, we are often more interested in the average (or expected) error across the many possible datasets that could have arisen had we repeated the data generating process many times (Aubry et al., 2024). An estimator's expected error is known as its bias.

In biodiversity science, the term “bias” is used inconsistently. Datasets are often described as being biased (Cretois et al., 2021, Boyd et al., 2024a, Díaz-Calafat et al., 2024, La Sorte et al., 2024, Backstrom et al., 2025a, Chowdhury et al., 2026) when in fact bias is a property of the estimator (and estimand). A dataset may induce bias for one estimand and estimator but not another. A second issue is that biases are often categorised by the dimension in which they manifest (e.g. spatial, temporal or taxonomic) (Isaac and Pocock, 2015, Bowler et al., 2024, Díaz-Calafat et al., 2024, Goldstein et al., 2024, La Sorte et al., 2024, Boyd et al., 2025) rather than by the part of the observation process that gives rise to them (e.g. sampling, measurement, etc.). The lack of a standardised framework creates the potential for confusion about how best to conceptualise bias, what gives rise to it and therefore how it can be mitigated.

We present a theory of observation for biodiversity science that decomposes bias into sampling and nonsampling components. The framework applies to any estimand, clarifies the origin of familiar sources of bias, and provides a principled basis for selecting appropriate mitigation strategies. This perspective shifts attention away from describing the characteristics of datasets towards understanding the observation processes that determine whether an estimator recovers the estimand.

Theory

Let $\mathcal{u}$ denote a target population comprising $i = 1, \dots, N$ units. These units might represent sites within a landscape, species within a taxon group, etc. Each unit is characterised by a vector of attributes $\mathbf{X}_i = \{X_{i1}, \dots, X_{ij}\}$. The empirical (relative frequency) distribution of $\mathbf{X}$ in the target population is denoted $P_{\mathbf{X}}$, and the estimand is some functional $f(P_{\mathbf{X}})$. We will represent the observation process as two transformations of the empirical distribution $P_{\mathbf{X}}$.

The first transformation is sampling: that is, the selection of a subset of units in the target population for which data will be collected. Let $R_i \in \{0,1\}$ denote a

sample inclusion indicator taking the value 1 for units included in the sample and 0 otherwise. Conditioning (i.e. focusing) on the sample ($R = 1$) results in the conditional distribution $P_{\mathbf{X}|\mathbf{R}=\mathbf{1}}$. This is simply the empirical distribution of $\mathbf{X}$ if we exclude nonsampled units.

Once a set of units has been selected for inclusion in the sample, data on the variables $\mathbf{X}$ are then measured on those units and stored for later use. Let $\mathbf{X}^* = \mathbf{X} + \mathbf{U}$ denote the recorded version of $\mathbf{X}$, where $\mathbf{U}$ is the realised discrepancy between the recorded and true values. $\mathbf{U}$ might represent, say, imperfect detection, instrument failure, false presences, and reporting errors. Applying this transformation to $P_{\mathbf{X}|\mathbf{R}=\mathbf{1}}$ yields the final observed empirical distribution $P_{\mathbf{X}^*|\mathbf{R}=\mathbf{1}}$.

Recall that our estimand is some functional f applied to the true empirical distribution $P_{\mathbf{X}}$. Suppose that we estimate the estimand by applying the same functional to the observed empirical distribution $P_{\mathbf{X}^*|\mathbf{R}=\mathbf{1}}$ —that is, $f(P_{\mathbf{X}^*|\mathbf{R}=\mathbf{1}})$. The error of $f(P_{\mathbf{X}^*|\mathbf{R}=\mathbf{1}})$ as an estimator of $f(P_{\mathbf{X}})$ for any given sample and set of measurements is

$$\text{realised error} = f(P_{\mathbf{X}^*|\mathbf{R}=\mathbf{1}}) - f(P_{\mathbf{X}}).$$

By adding and subtracting the same quantity, $f(P_{\mathbf{X}|\mathbf{R}=\mathbf{1}})$, we can decompose the error into sampling and nonsampling components (Figure 1):

$$\underbrace{f(P_{\mathbf{X}^*|\mathbf{R}=\mathbf{1}}) - f(P_{\mathbf{X}})}_{\text{realised error}} = \underbrace{[f(P_{\mathbf{X}|\mathbf{R}=\mathbf{1}}) - f(P_{\mathbf{X}})]}_{\text{sampling error}} + \underbrace{[f(P_{\mathbf{X}^*|\mathbf{R}=\mathbf{1}}) - f(P_{\mathbf{X}|\mathbf{R}=\mathbf{1}})]}_{\text{nonsampling error}}.$$

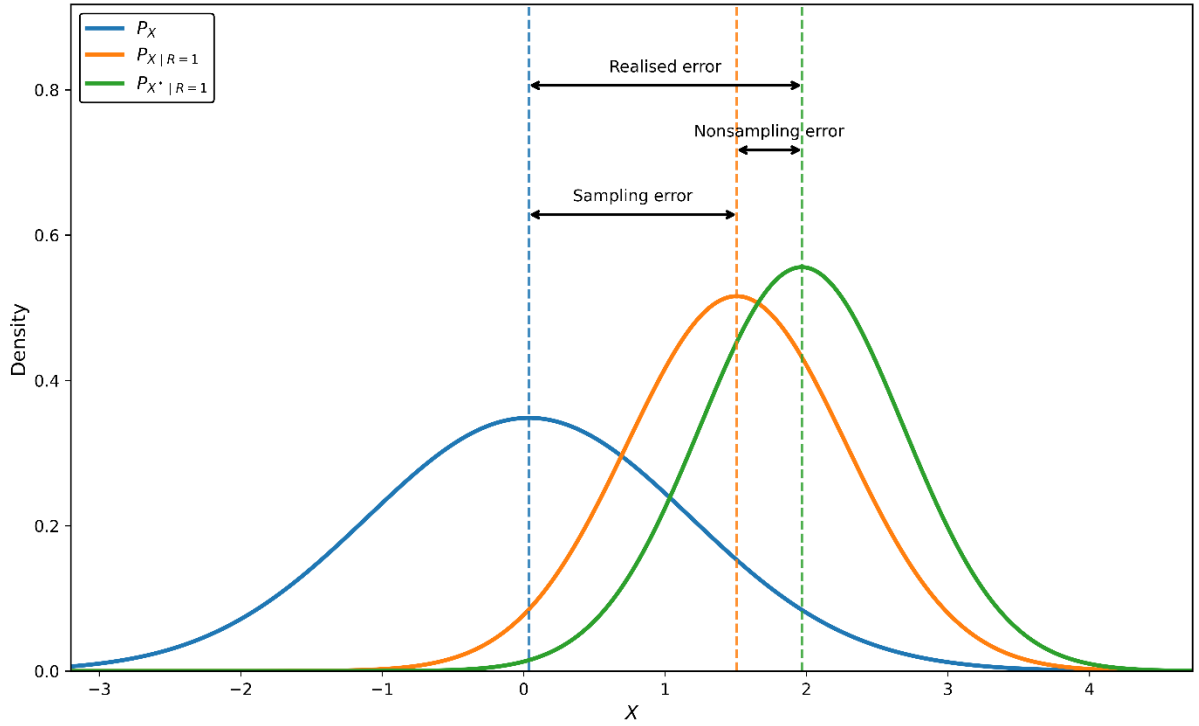


Figure 1. Illustration of realised error decomposition. The blue, orange, and green curves represent the distributions of the target population, the sampled units (true values), and the measured values for the sampled units, respectively. The vertical dashed lines indicate the mean of each distribution. The horizontal arrows illustrate the sampling error (difference between the target population and sample means), nonsampling error (difference between the sample and measured sample means), and realised error (difference between the target population and measured sample means). Gaussian curves are used for illustration only. The arrows indicate the magnitudes of the differences, while the algebraic decomposition retains their signs.

Rather than the realised error of an estimator, we are often interested in its bias. An estimator's bias is its average, or expected, error with respect to whichever components of the system are considered to be stochastic (random). A process is stochastic if it has many possible realisations: for example, under random sampling, we select one sample, but many others were possible. The same logic extends to sampling that isn't explicitly random if we imagine some model determining which units are selected (e.g. citizen scientists selecting sites at which to collect data). Depending on the inferential framework, sampling, measurement, the population itself (i.e. P_X), or any combination thereof may be treated as stochastic. Regardless, linearity of expectation implies that the bias of $f(P_{X^*|R=1})$ as an estimator of $f(P_X)$ is

$$\underbrace{E[f(P_{X^*|R=1}) - f(P_X)]}_{\text{estimator bias}} = \underbrace{E[f(P_{X|R=1}) - f(P_X)]}_{\text{sampling bias}} + \underbrace{E[f(P_{X^*|R=1}) - f(P_{X|R=1})]}_{\text{nonsampling bias}},$$

where the expectations are taken with respect to the random components of the population and/or observation process. Thus, irrespective which components are treated as stochastic, the bias of $f(P_{X^*|R=1})$ decomposes into sampling and

nonsampling components. This decomposition provides the organizing principle for the review that follows by linking bias to the processes through which it arises.

Estimands describing change over time

Many scientific questions do not concern parameters defined at one point in time but rather how some quantity has changed over time. For example, we might be interested in the how a per time-period parameter changes between time periods, the linear trend, etc. When the estimand concerns change over time, what matters is not the bias in any individual time period but how that bias changes through time. Indeed, the bias of any linear functional of the time-series is equal to the same functional applied to the bias itself. For example, if an estimand is calculated for two time periods, the bias of the difference between the two is equal to the difference between the period-specific biases. Likewise, the bias of the coefficient from a simple linear regression of an estimand on time is equal to the coefficient that would be obtained by regressing the bias itself on time. These results follow directly from the linearity of expectation. Since the per-period biases have sampling and non-sampling components, so too does the bias of the estimated trend. That is, change over time in sampling bias can bias estimates of temporal change and likewise for nonsampling bias.

The target population determines the bias dimension

It is common for biases in ecological data to be described as “spatial”, “taxonomic” or “temporal” (Bowler et al., 2024, Goldstein et al., 2024, Alfeus et al., 2025, Boyd et al., 2025, Pescott et al., 2025, Boyd et al., 2026b). Our framework reveals that it is more natural to distinguish biases according to the stage of the observation process at which they arise: sampling bias arises from which units are selected for inclusion in the sample, whereas nonsampling bias arises from how information about sampled units is measured, recorded or stored. From a mathematical perspective, the distinction between spatial and taxonomic sampling bias lies primarily in the definition of the population units. Temporal bias is somewhat different, as discussed above.

To illustrate this point, suppose the target population is a set of sites within a landscape and that the estimand is mean species richness. Data on species richness are measured with error at sampled sites. In this example, any sampling bias arises solely from the selection of which sites end up in the sample. Nonsampling bias, on the other hand, arises from how species richness is measured at sampled sites and how the resulting data are recorded and stored. Although it is tempting to describe the failure to detect certain species as a taxonomic sampling bias, it is in this context a nonsampling bias.

Sources of sampling bias

Sampling bias arises when the surveyed units do not adequately represent the target population specified by the estimand. The central question is therefore whether the surveyed units represent that target population in terms of $\mathbf{X}$.

Spatial populations

When the target population consists of locations, sampling involves selecting sites. For example, to estimate the mean abundance of a species across a region, it is usually impossible to sample every location. Therefore, any factor that affects both sample inclusion and the unit-specific quantity of interest can be considered a potential source of sampling bias

Public access limitations

Different indicators of public accessibility and their effects on recording have been widely studied (Stolar and Nielsen, 2015, Geldmann et al., 2016, Mair and Ruete, 2016, Thornhill et al., 2016, Welvaert and Caley, 2016, Tiago et al., 2017, Millar et al., 2019, Zhang, 2020, Petersen et al., 2021, Guetté et al., 2022). The length and density of public roads or footpaths, as well as the distance from observation points to the nearest public footpath, are commonly used indicators of accessibility (Tiago et al., 2017, Sicacha-Parada et al., 2019, Mandeville et al., 2022, Ramesh et al., 2022, Dimson and Gillespie, 2023, Geurts et al., 2023, Ostermann-Miyashita et al., 2023). Generally, more public footpaths mean more observation opportunities (Li et al., 2026b). At the same time, studies have shown that the number of citizen science records is negatively correlated with the distance from footpaths, with more records occurring closer to footpaths (Sicacha-Parada et al., 2019, Ostermann-Miyashita et al., 2023). The area of publicly accessible land can also directly represent accessibility (Boakes et al., 2010, Millar et al., 2019, Mentges et al., 2020, Dimson et al., 2023, Grady et al., 2026). Larger nature reserves or national parks not only mean a larger extent of potential habitat but also more space that can be accessed and observed, thereby increasing the likelihood of recording.

In addition, human population density can also represent accessibility to a certain extent (Ward, 2014, Fithian et al., 2015, Stolar and Nielsen, 2015, August et al., 2020, Zhang, 2020, Grady et al., 2026). In areas with higher population density, there is usually more infrastructure such as roads and footpaths, as well as more transport options (Phillips et al., 2009, Fourcade et al., 2014, Zhang, 2020, Arazy and Malkinson, 2021, Guetté et al., 2022, Bowler et al., 2024), so observation activities are more frequent. Studies have shown that the number of species records in these areas is often higher than in surrounding rural areas, and the data

show a clustered pattern, indicating a clear sampling bias (Arazy and Malkinson, 2021, Bowler et al., 2024).

Agricultural land, private gardens, and other types of land that are not typically open to the public often contain suitable habitats for species, but due to the lack of accessibility, species records in these areas are usually insufficient, thereby leading to spatial bias in species distributions (Bach et al., 2022, Potgieter et al., 2024, Barreiro et al., 2025, Gillard and Rowley, 2026). Similar situations also occur in other restricted areas, such as railway corridors in cities or military sites (Potgieter et al., 2024, Gillard and Rowley, 2026). Taking Dartmoor Training Area within Dartmoor National Park in the United Kingdom as an example, these areas may contain suitable habitat types, but because the public generally cannot access them (Wang and Wu, 2013), the species distributions within them are difficult to record effectively. Moreover, legal regulations regarding access to private land vary significantly across countries (Nichiforel et al., 2020). For example, in the United States, relevant laws such as the 'castle doctrine' reinforce, to some extent, the exclusivity of private land, which further limits the collection of citizen science data in these areas (O'Brien et al., 2021). Fortunately, targeted citizen science schemes focusing on private land, such as gardens (Wilson et al., 2015, Fontaine et al., 2016, Hordley and Fox, 2024) or farmland (Billaud et al., 2021), together with structured sampling by researchers, may help compensate for the lack of citizen science data from these areas (Dickinson et al., 2010, O'Brien et al., 2021, Agnoux et al., 2026, Gillard and Rowley, 2026). However, the availability of such data depends heavily on the subjective willingness of landowners, as not everyone pays attention to or records the biodiversity on their land (Williams et al., 2015, O'Brien et al., 2021).

Site preference

From the perspective of recorders' subjective intentions, the selection of sampling sites itself is also an important source of bias (Williams et al., 2015). Before conducting observations, participants actively or randomly choose survey locations (Thornhill et al., 2016, Millar et al., 2019, Neyens et al., 2019, Cretois et al., 2021, Bullion and Bahlai, 2023, Dimson et al., 2023). Studies have shown that opportunistic recorders tend to visit biodiversity hotspots for recording (Dennis and Thomas, 2000, Boakes et al., 2016, Totterman, 2021, Mondain-Monval et al., 2024, Grady et al., 2026), and this behaviour may be influenced by citizen science projects encouraging the recording of more species or by social media sharing sightings (Dennis and Thomas, 2000, Catlin-Groves, 2012, Callaghan et al., 2023, O'Neill et al., 2023, Pernat et al., 2024, Kc and Sapkota, 2025). In contrast, non-hotspot areas are often overlooked. For example, common species occurring in the broader areas surrounding regions with high data density may be overlooked by citizen scientists (Li et al., 2025), resulting in difficulties in accurately mapping

their distributions (Geldmann et al., 2016, Shirey et al., 2020, Totterman, 2021, Barreiro et al., 2025, Grady et al., 2026), thereby further exacerbating spatial imbalance of citizen science records.

Temporal issues

When a temporal dimension is added to a spatial sampling problem, the same principle still applies. For example, if the aim is to estimate the annual mean abundance of a species across a region, any factor that affects which spatial units are included in the sample remains a source of sampling bias. Some locations may only be accessible during particular periods, causing other sites to be excluded, as in the case of military areas (Wang and Wu, 2013). Unfavourable weather conditions may also reduce recording activity, particularly during extreme weather years, potentially exacerbating temporal variation in sampling bias (Rosário et al., 2025).

Changes in public access or transport availability over time may allow recorders to reach a broader range of locations (Bowler et al., 2022b), thereby altering which sites are included in the sample. During COVID, restrictions on movement caused citizen science sampling to become more geographically biased towards urban or otherwise human modified areas, where people were more likely to record close to home (Hochachka et al., 2021, Sánchez-Clavijo et al., 2021).

Taxonomic populations

When species constitute the population units, sampling bias arises when species differ systematically in their probability of being included in the sample, resulting in a sampled assemblage that is not representative of the target taxonomic population. One important source of such unequal inclusion probabilities is species preference, whereby recorders may preferentially seek, detect, or report particular species over others.

Species preference

For example, in some citizen science databases, birds account for a disproportionately large number of records compared with other taxa (Troudet et al., 2017, La Sorte and Somveille, 2020, Stoudt et al., 2022, Huang et al., 2023, Backstrom et al., 2025b), while Lepidoptera records often greatly outnumber those of other insect orders (Díaz-Calafat et al., 2024) indicating clear taxonomic skew. Within taxa, some recorders exhibit clear bias towards rarer species (Johnston et al., 2023) or larger species (Callaghan et al., 2021) and may invest substantial time and effort to locate and record them, for example travelling long distances to document rare butterfly species (Li et al., 2025). This can lead to overreporting of rarer species (Backstrom et al., 2025b). Tourists tended to record

taxa that are particularly attractive or distinctive while common species or less attractive species are overlooked (Tulloch and Szabo, 2012, Dimson et al., 2023) and this might lead to local underestimation of species richness.

In addition, smartphone apps, websites and the ways in which they are promoted can lead to species preference (Habel et al., 2025). Content on these platforms is often influenced by the preferences of those who post it (Catlin-Groves, 2012). For example, some content creators are more likely to showcase visually attractive or particularly charismatic species, which may encourage viewers to record those species preferentially (Pernat et al., 2024). Social platforms may also encourage competition among recorders to report the first sighting of a species (Bishop et al., 2013), which can further increase taxonomic bias (Belitz et al., 2020). Moreover, species that have already been recorded may be less likely to be reported again, as recorders may preferentially record species that are new to them (Bowler et al., 2022a). This behavioural pattern further leads to sampling completeness issue.

In addition, volunteers may simultaneously consider target species and accessibility when making recording decisions. For example, Li et al. (2025) illustrated that different type of recorders make trade-offs between observation benefits and costs, such as distance, and many volunteers tend to submit records close to their place of residence (Dennis and Thomas, 2000). This 'site preference' further reinforces the clustered spatial pattern of records (Dennis and Thomas, 2000, Arazy and Malkinson, 2021, Totterman, 2021). It also indicates that sampling bias and measurement bias may occur simultaneously and interact with each other (Boakes et al., 2010, Ward, 2014, Boakes et al., 2016, Zhang, 2020, Petersen et al., 2021, Backstrom et al., 2024, Barreiro et al., 2025).

Temporal issues

Similarly, when a temporal dimension is added to a taxonomic sampling problem, the same principle applies. For example, if the aim is to estimate annual species richness within a region, species are the population units, and any factor that affects which species are included in the sample remains a source of sampling bias. Some species may only be detectable during particular periods, causing them to be excluded from the sampled species set if surveys do not cover those periods. In terms of phenology, recorders may prefer to record the adult stage of insects rather than the immature developmental stages, because individuals are relatively more detectable during the flight period (Bishop et al., 2013, Rosário et al., 2025).

Temporal preferences in species occurrence and variation in detectability at particular times are discussed in more detail under the spatial population section

of nonsampling sources of bias; the underlying mechanisms are similar, but the classification differs because the target population and population units are different.

Sources of nonsampling bias

Nonsampling bias arises when, among the units included in the sample, the observed data do not accurately reflect the true variables specified by the estimand. The central question is therefore whether the observations collected from sampled units provide accurate measurements of those variables.

Spatial populations

Species detectability

When spatial locations are the population units, imperfect species detectability is a source of nonsampling bias because it can distort the quantity estimated for each sampled location.

Differences in species characteristics, including body size, colouration, conspicuousness, and behaviour, can influence the probability that individuals are detected and subsequently recorded. A large body of research has shown that detectability error is widespread in some forms of citizen science (Tulloch and Szabo, 2012, Boakes et al., 2016, Xue et al., 2016, Troudet et al., 2017, Robinson et al., 2018, Lobo et al., 2021, Petersen et al., 2021, Stoudt et al., 2022, Huang et al., 2023). For example, large bird species are more frequently recorded because they are easier to observe and identify (Callaghan et al., 2021, Stoudt et al., 2022, Alfeus et al., 2025). Similarly, in opportunistic butterfly datasets, species with bright colours or more visually attractive appearances often receive more records, whereas common species lacking distinctive features may be overlooked by recorders (Dennis et al., 2006, Kral-O'Brien et al., 2020).

Species detectability may also shift abruptly when new technologies or survey methods are introduced, for example through improved optical equipment, enhanced identification support from new field guides or AI-based classifiers (Santoro et al., 2025, Sharpe et al., 2025), and more effective active-sampling tools, including novel pheromone lures and trap designs (Thomaes et al., 2022).

Sampling completeness

When spatial locations constitute the population units, incomplete sampling can introduce nonsampling bias by causing site-level estimates of species richness or community composition to systematically deviate from their true values.

Sampling completeness refers to the extent to which a survey records all species present at a given site (Sánchez-Fernández et al., 2021). At the individual level, this is often reflected in list length. Sampling completeness may be influenced by multiple factors, such as recorders' species preferences, recorders' skills and expertise, or the unintentional overlooking of certain species, all of which can result in incomplete records (Jacobs and Zipf, 2017, Sánchez-Fernández et al., 2021, Backstrom et al., 2024, Alfeus et al., 2025). At larger spatial scales, such incompleteness may manifest as uneven sampling effort, where some locations are heavily surveyed while others receive little or no observation (La Sorte and Somveille, 2020, Chanachai et al., 2024, Barreiro et al., 2025). From a long-term perspective, the completeness of surveys at a given location typically increases as observation time accumulates and more visits are conducted (Backstrom et al., 2024). However, ensuring complete sampling across all sites within a short time frame remains challenging (Backstrom et al., 2024), as it is often uncertain whether additional species may still be detected in the future. Consequently, assessments of sampling completeness might use results from previous structured surveys as a reference (Brown et al., 2020).

Survey methods

Survey methods can introduce nonsampling bias when they systematically distort the quantities estimated at sampled sites. Differences among methods in their ability to detect, capture, identify, or count organisms can cause site-level observations to deviate from the underlying true values. Active sampling methods can present particular challenges because the effects of environmental covariates on the estimand may be complex and non-linear. A classic example is the relationship between pan-trap capture rates and floral resources (Wilson et al., 2008, Shrestha et al., 2019): traps may capture fewer insects in flower-rich habitats than in flower-poor habitats, even when the true abundance of insects is higher, because abundant natural flowers compete with the traps for insect attraction (Wilson et al., 2008).

However, the classification of these effects depends on how the target population is defined. When species rather than spatial locations constitute the population units, survey methods may also generate taxonomic nonsampling bias by systematically distorting the quantities estimated for species that have already been included in the sample.

Temporal issues

When spatial locations constitute the target population and the sampling units have already been selected, nonsampling bias arises when any aspect of the

subsequent observation or measurement process causes the estimated values for those sampled units to systematically deviate from their true values.

One form of temporal issues occurs when recording activity is disproportionately concentrated at particular time points or during specific time periods, causing per-period estimates to overrepresent conditions associated with those times. For example, studies of volunteer recording activity have shown that observations can be concentrated in the evening, late at night, or shortly before sunrise rather than around midnight, potentially reflecting individual recorder preferences or targeted searches for nocturnal species (Arazy et al., 2024, Willig et al., 2024). It has also been reported that some species records are substantially more frequent during daytime than at night, which may be partly attributed to higher species detectability under daylight conditions and that most human recording activity is diurnal (Welvaert and Caley, 2016, Fraisl et al., 2022, Arazy et al., 2024). Approximately 60% of invertebrates are nocturnal, with a similar proportion of mammals, whereas less than 20% of birds are nocturnal (Hölker et al., 2010). This mismatch increases the likelihood of recording diurnal taxa, making birds more likely to be recorded than mammals, and day-flying butterflies more likely to be recorded than night-flying moths (Hölker et al., 2010). In addition, some studies have shown that larger bodied species tend to be more active during the warmest periods of the day, which may further bias records towards these time windows and contribute to the underrepresentation of nocturnal species (Rosário et al., 2024, Idec et al., 2025, Rosário et al., 2025).

At the week and seasonal levels, some studies have shown that citizen science recorders tend to record more frequently during weekends and this pattern is likely to reflect greater recorder availability during weekends, when many people have fewer work related commitments and more opportunities to participate in observation activities (Welvaert and Caley, 2016, Zhang, 2020, Di Cecco et al., 2021, Fraisl et al., 2022, Rosário et al., 2024, Rosário et al., 2025). However, others have reported lower submission rates during weekends, which may be explained by recorders collecting data during spare moments on weekdays (Khoi and Casteleyn, 2018). This pattern may also be related to the uneven demographics of citizen science participants, who tend to be older and have higher incomes than the general population (Pateman et al., 2021), and may therefore have more flexible schedules or more free time and be less constrained by weekday versus weekend differences.

Some studies have also shown that citizen science participants tend to conduct sampling more frequently in spring and summer (Bishop et al., 2013, Hadj-Hammou et al., 2017, Rosário et al., 2024, Rosário et al., 2025). However, others have reported an overrepresentation of records in winter, which is mainly related to the characteristics of target species (Milberg et al., 2024).

Specific events may alter the number of records for particular species even when observation locations remain fixed, potentially distorting species-level estimates during the affected period. For example, the COVID 19 pandemic had a significant impact on citizen science data (Basile et al., 2021, Crimmins et al., 2021, Sánchez-Clavijo et al., 2021, Sweet et al., 2022, Dimson and Gillespie, 2023, Zhang et al., 2024). Some studies have illustrated that, under social distancing measures, people showed a significantly increased willingness to participate in outdoor activities and some countries did not impose strict restrictions on outdoor activities, which led to a peak in data submissions during this period (Dimson and Gillespie, 2023, Zhang et al., 2024).

Taxonomic populations

For taxonomic populations, nonsampling bias arises when the values associated with sampled species are systematically mismeasured or misestimated.

Species preference

Species preference can also generate nonsampling bias in taxonomic populations when it alters the estimated values of already sampled species, rather than determining which species enter the sample. Because the underlying mechanisms of species preference have been discussed above, they are not considered in detail again here.

Species detectability

The underlying mechanism is similar to that discussed for spatial populations, but here species rather than locations constitute the population units. Differences in detectability can therefore affect the probability that individual species are included in the sample. For example, visually attractive or conspicuous species may be more likely to be noticed and recorded, resulting in their overrepresentation in the observed species list. As the mechanisms affecting detectability have already been discussed above for spatial populations, they are not repeated in detail here.

Recorder's skill level

Recorder skill level affects the measurement component because differences in observers' ability to detect, identify, and count species can lead to missed detections, misidentifications, and inaccurate abundance estimates (Johnston et al., 2023).

Recorders can be classified according to their level of expertise, such as experts and non-experts (Kosmala et al., 2016, Brown and Williams, 2019, McCarthy et al.,

2021, Huang et al., 2023, Zheng et al., 2025), or according to their residential status, such as tourists and residents (Tulloch and Szabo, 2012, Dimson and Gillespie, 2023). Studies have shown that data provided by tourists and residents may differ in both spatial distribution and the types of species recorded, and their levels of observational skill may also vary (Dimson and Gillespie, 2023, Li et al., 2025, Grady et al., 2026). Some recorders, due to limited taxonomic skills, may be more inclined to record common or easily identifiable species, while overlooking species that are more difficult to identify (Crall et al., 2011, Kelling et al., 2015, van der Wal et al., 2016, Santos-Fernandez and Mengersen, 2021, Goldstein et al., 2024). Incorporating measures of recorder expertise into models can improve analysis of citizen science data (Johnston et al., 2018). However, while attempts have been made to classify recorders (Boakes et al., 2016, August et al., 2020) and assess the impact of recorder behaviour on biodiversity assessments (Pocock et al., 2023), there are practical challenges in characterising recorder expertise, including how expertise should be measured and its likely dependence on context, such as time, space and taxonomic group. Consequently, a widely accepted and unified framework is still lacking.

Recorder skill can also contribute to nonsampling bias in spatial populations, depending on how the population units and estimand are defined. For example, if locations are treated as the population units, differences in recorder skill may affect the accuracy of site-level estimates by altering species detection, identification, or counting within sampled sites. Thus, the same mechanism may be classified differently depending on whether the unit of interest is a site or a species.

Temporal issues

When species constitute the target population and the sampled species have already been selected, nonsampling bias arises when any aspect of the subsequent observation or measurement process causes the estimated values for those species to systematically deviate from their true values. The classification of bias therefore depends on how the target population is defined. Many of the underlying mechanisms are the same as those discussed for spatial populations, so they are not repeated here; instead, we focus on aspects that differ when species are treated as the population units.

For example, if the aim is to estimate annual abundance for a sampled species, the magnitude or direction of bias in that estimate may vary among years. Such temporal variation can arise for several reasons, including changes over time in recorder preference for that species or in recorders' ability to identify it correctly (Kelling et al., 2015). As the season progresses, recorders may gradually lose interest in routine species recording and instead focus on specific time points or

events (Bishop et al., 2013). Consequently, the bias may vary across time points, resulting in differences in both its magnitude and direction over time.

In addition, when species constitute the target population, temporally clustered citizen science recording events may introduce taxonomic nonsampling bias by systematically affecting species-level estimates (Kullenberg and Kasperowski, 2016), such as World Migratory Bird Day on 11 May, bioblitz (Lundmark, 2003), and the Big Butterfly Count held in the UK during summer (Dennis et al., 2024). Although these semi-structured recording activities can alleviate some limitations of unstructured data (often by encouraging complete or constrained taxonomic lists), they may also concentrate observations within specific dates or time windows, thereby reinforcing temporal unevenness in records (Dennis et al., 2017).

Moreover, citizen science participants may compete to become the first to record a species, which may lead to mismatches between phenology estimated from citizen science data and that from structured data (Li et al., 2026a), thereby affecting the prediction of phenological metrics (Bishop et al., 2013); For example, it has been reported that volunteers may record some species earlier than experts (Soroye et al., 2018).

Bias correction

So far, we have defined the estimator as the same functional that defines the estimand but applied to the available data. This is the ‘plug-in’ estimator and is the natural starting point. However, more sophisticated estimators often have lower bias than the plug-in estimator and should therefore be preferred. We argue that many such estimators can be viewed as attempts to ‘undo’ the sampling and measurement processes. Starting with sampling, we review these estimators below.

Estimators that correct sampling bias

Sampling removes units from the target population, thereby changing the relative frequencies of the variables in $\mathbf{X}$. Many estimators of population parameters can be viewed as approaches to restoring the population relative frequencies before applying the plug-in estimator.

One approach is to apply weights to sampled units. The idea is to place more weight on values of $\mathbf{X}$ that are underrepresented in the sample relative to the population and vice versa, thereby reconstructing a “pseudopopulation”. Specifying weights is simple under probability sampling, in which case each unit’s weight is the reciprocal of the probability that it was included in the sample (Boyd

et al., 2023). Under nonprobability sampling, weights must be estimated (Elliott and Valliant, 2017). Having specified weights, one effectively applies the plug-in estimator to the weighted sample. This approach has been used to estimate population means, totals and regression coefficients for decades (Horvitz and Thompson, 1952). In biodiversity science, weighted estimators have been used to estimate trends in species occupancy, correlates of species occupancy, and trends in species abundance (Van Swaay et al., 2002, Irvine et al., 2018, Johnston et al., 2020, Fink et al., 2023, Boyd et al., 2024b).

Closely related to weighting is subsampling. The idea is to create a representative miniature of the target population (Meng, 2022). Some units are removed, which effectively gives them zero weight. The plug-in estimator is then applied to the subsample. This approach has been used to estimate species distributions and temporal trends thereof (Beck et al., 2014, Boyd et al., 2024b).

Rather than reweighting the sample, a second class of estimators reconstructs the population by predicting or imputing values for nonsampled population units (Elliott and Valliant, 2017). The plug-in estimator is then applied to imputed or predicted dataset. This approach has been used in biodiversity science to estimate trends in species abundance (Boyd et al., 2025) and parameters describing species traits (Johnson et al., 2021).

Estimators that correct nonsampling bias

Many estimators of population parameters can be viewed as attempts to recover the latent value of the variable(s) of interest on sampled units, thereby reconstructing the hypothetical error-free sampled population. Canonical examples in biodiversity science include occupancy-detection and N-mixture models (MacKenzie et al., 2002, Royle, 2004). These models exploit replicate data on plant and animal occupancy or abundance to estimate detection probabilities, which are then used to infer the true latent occupancy or abundance. Importantly, as emphasised by MacKenzie et al. (2002), these models do not deal with sampling bias.

Estimators that correct sampling and nonsampling bias

It is also possible to construct estimators that simultaneously correct sampling and nonsampling biases. For example, Irvine et al. (2018) and Johnston et al. (2020) constructed likelihood-weighted occupancy detection-models. The likelihood weights represent an attempt to 'undo' the sampling process, and the detection sub-model uses data on replicate visits to correct for imperfect detection.

Recommendations

We recommend that researchers first define the estimand clearly, including the target population, variables, and parameter of interest, and then determine which component of the inferential process each source of bias affects. Corrections should be selected according to the underlying source of error, with explicit consideration of which limitations can be corrected analytically or through study design and which can only be addressed by narrowing or changing the scope of interpretation. In this way, the estimand framework supports both more appropriate treatment of inferential error and the identification of further questions about how biodiversity data are generated and interpreted.

Acknowledgements

RJB was supported by the NERC Independent Research Fellowship UKRI4865: A new paradigm for biodiversity monitoring. DBR was supported UKCEH National Capability for UK Challenges programme NE/Y006208/1 and NERC award number NE/R016429/1 as part of the UK-SCAPE programme delivering National Capability. RF and EBD were supported by the Heather Corrie Fund from Butterfly Conservation.

References

- AGNOUX, S., FONTAINE, B., DOZIÈRES, A., JEANMOUGIN, M., SCHMUCKI, R., VIMONT, M. & MARTIN, G. 2026. Contribution of citizen sciences to the production of butterfly species long-term trends. *Insect Conservation and Diversity*, n/a.
- ALFEUS, M., IRISH, J. & BIRKHOFER, K. 2025. Recognition and completeness metrics from iNaturalist and GBIF can inform future citizen science and research projects: a case study on arthropods in Namibia. *Biodiversity and Conservation*, 34, 467–480.
- ARAZY, O., KAPLAN-MINTZ, K., MALKINSON, D. & NAGAR, Y. 2024. A local community on a global collective intelligence platform: A case study of individual preferences and collective bias in ecological citizen science. *PLOS ONE*, 19.
- ARAZY, O. & MALKINSON, D. A Framework of Observer-Based Biases in Citizen Science Biodiversity Monitoring: Semi-Structuring Unstructured Biodiversity Monitoring Protocols. *Frontiers in Ecology and Evolution*, 2021.
- AUBRY, P., FRANCESIAZ, C. & GUILLEMAIN, M. 2024. On the impact of preferential sampling on ecological status and trend assessment. *Ecological Modelling*, 492, 110707.
- AUBRY, P., QUANTENNE, G., GUILLEMAIN, M. & MATUTINI, F. 2026. When n is not enough: Assessing the interplay of spatial sampling error and sample selection in abundance estimation. *Ecological Modelling*, 522, 111793.
- AUGUST, T., FOX, R., ROY, D. B. & POCOCK, M. J. 2020. Data-derived metrics describing the behaviour of field-based citizen scientists provide insights for project design and modelling bias. *Scientific reports*, 10, 11009.
- BACH, H. M., ESCOUBET, H. & MAYER, M. 2022. Co-occurrence patterns and habitat selection of the mountain hare, European hare, and European rabbit in urban areas of Sweden. *Mammalian Biology*, 103, 187–203.
- BACKSTROM, L. J., CALLAGHAN, C. T., LESEBERG, N. P., SANDERSON, C., FULLER, R. A. & WATSON, J. E. M. 2024. Assessing adequacy of citizen science datasets for biodiversity monitoring. *Ecology and Evolution*, 14, e10857.
- BACKSTROM, L. J., CALLAGHAN, C. T., WORTHINGTON, H., FULLER, R. A. & JOHNSTON, A. 2025a. Estimating sampling biases in citizen science datasets. *Ibis*, 167, 73–87.
- BACKSTROM, L. J., DRAKE, R. L., WORTHINGTON, H. & JOHNSTON, A. 2025b. Detections of Rare Species Lead Citizen Scientists to Initiate Data Recording. *Diversity and Distributions*, 31, e70103.
- BARREIRO, K., BENESTAN, L., MORITZ, C., DUCATEZ, S., GAERTNER, J.-C., LE LUYER, J. & MONACO, C. J. 2025. Species richness variation in marine and terrestrial fauna across widespread, fragmented territories: assessing inherent challenges of data scarcity at local and regional scales. *Scientific Reports*, 15, 21043.

- BASILE, M., RUSSO, L. F., RUSSO, V. G., SENESE, A. & BERNARDO, N. 2021. Birds seen and not seen during the COVID-19 pandemic: The impact of lockdown measures on citizen science bird observations. *Biological Conservation*, 256, 109079.
- BATES, O. K. & BERTELSMEIER, C. 2025. Predictions of future insect distributions under climate change. *Diversity and Distributions*, 31, e70106.
- BECK, J., BÖLLER, M., ERHARDT, A. & SCHWANGHART, W. 2014. Spatial bias in the GBIF database and its effect on modeling species' geographic distributions. *Ecological Informatics*, 19, 10-15.
- BELITZ, M. W., LARSEN, E. A., RIES, L. & GURALNICK, R. P. 2020. The accuracy of phenology estimators for use with sparsely sampled presence-only observations. *Methods in Ecology and Evolution*, 11, 1273-1285.
- BILLAUD, O., VERMEERSCH, R.-L. & PORCHER, E. 2021. Citizen science involving farmers as a means to document temporal trends in farmland biodiversity and relate them to agricultural practices. *Journal of Applied Ecology*, 58, 261-273.
- BISHOP, T. R., BOTHAM, M. S., FOX, R., LEATHER, S. R., CHAPMAN, D. S. & OLIVER, T. H. 2013. The utility of distribution data in predicting phenology. *Methods in Ecology and Evolution*, 4, 1024-1032.
- BOAKES, E. H., GLIOZZO, G., SEYMOUR, V., HARVEY, M. C., SMITH, C. W., ROY, D. B. & HAKLAY, M. 2016. Patterns of contribution to citizen science biodiversity projects increase understanding of volunteers' recording behaviour. *Scientific Reports*, 6.
- BOAKES, E. H., MCGOWAN, P. J. K., FULLER, R. A., DING, C.-Q., CLARK, N. E., O'CONNOR, K. & MACE, G. M. 2010. Distorted Views of Biodiversity: Spatial and Temporal Bias in Species Occurrence Data. *PLoS Biology*, 8.
- BOULANGER, Y., DESAINT, A., MARTEL, V., MARCHAND, M., MASSODA TONYE, S., SAINT-AMANT, R. & RÉGNIÈRE, J. 2025. Recent climate change strongly impacted the population dynamic of a North American insect pest species. *PLOS Climate*, 4, e0000488.
- BOWLER, D. E., BHANDARI, N., REPKE, L., BEUTHNER, C., CALLAGHAN, C. T., EICHENBERG, D., HENLE, K., KLENKE, R. A., RICHTER, A., JANSEN, F., BRUELHEIDE, H. & BONN, A. 2022a. Decision-making of citizen scientists when recording species observations. *Scientific Reports*, 12.
- BOWLER, D. E., BOYD, R. J., CALLAGHAN, C. T., ROBINSON, R. A., ISAAC, N. J. B. & POCOCK, M. J. O. 2024. Treating gaps and biases in biodiversity data as a missing data problem. *Biological Reviews of the Cambridge Philosophical Society*, 100, 50 - 67.
- BOWLER, D. E., CALLAGHAN, C. T., BHANDARI, N., HENLE, K., BENJAMIN BARTH, M., KOPPITZ, C., KLENKE, R., WINTER, M., JANSEN, F. & BRUELHEIDE, H. 2022b. Temporal trends in the spatial bias of species occurrence records. *Ecography*, 2022, e06219.
- BOYD, R. J., BOTHAM, M., DENNIS, E., FOX, R., HARROWER, C., MIDDLEBROOK, I., ROY, D. B. & PESCOTT, O. L. 2025. Using causal diagrams and superpopulation models to correct geographic biases in biodiversity monitoring data. *Methods in Ecology and Evolution*, 16, 332-344.
- BOYD, R. J., BOWLER, D. E., ISAAC, N. J. B. & PESCOTT, O. L. 2024a. On the trade-off between accuracy and spatial resolution when estimating species occupancy from geographically biased samples. *Ecological Modelling*, 493, 110739.
- BOYD, R. J., COOKE, R., POWNEY, G. D. & PESCOTT, O. L. 2026a. Why your DAG should probably incorporate sampling and measurement processes. *Current Opinion in Insect Science*, 78, 101578.
- BOYD, R. J., JARVIS, S., MENG, X.-L., POWNEY, G. D., SPAKE, R. & PESCOTT, O. L. 2026b. How to reduce sampling error in species population monitoring: from theory to methods. *Ecography*, 2026, e08103.
- BOYD, R. J., POWNEY, G. D. & PESCOTT, O. L. 2023. We need to talk about nonprobability samples. *Trends in Ecology & Evolution*, 38, 521-531.
- BOYD, R. J., STEWART, G. B. & PESCOTT, O. L. 2024b. Descriptive inference using large, unrepresentative nonprobability samples: An introduction for ecologists. *Ecology*, 105, e4214.
- BROWN, E. D. & WILLIAMS, B. K. 2019. The potential for citizen science to produce reliable and useful information in ecology. *Conservation Biology*, 33, 561-569.
- BROWN, N., PÉREZ-SIERRA, A., CROW, P. & PARNELL, S. 2020. The role of passive surveillance and citizen science in plant health. *CABI Agriculture and Bioscience*.
- BULLION, C. M. & BAHLAI, C. A. 2023. Data gap or biodiversity gap? Evaluating apparent spatial biases in community science observations of Odonata in the east-central United States. *PeerJ*, 12.
- BUONACCORSI, J. P. 2010. *Measurement error: Models, methods, and applications*, CRC press Boca Raton, FL.
- CALLAGHAN, C. T., POORE, A. G. B., HOFMANN, M., ROBERTS, C. J. & PEREIRA, H. M. 2021. Large-bodied birds are over-represented in unstructured citizen science data. *Scientific Reports*, 11, 19073.
- CALLAGHAN, C. T., THOMPSON, M., WOODS, A., POORE, A. G. B., BOWLER, D. E., SAMONTE, F., ROWLEY, J. J. L., ROSLAN, N., KINGSFORD, R. T., CORNWELL, W. K. & MAJOR, R. E. 2023. Experimental evidence that behavioral nudges in citizen science projects can improve biodiversity data. *BioScience*, 73, 302-313.
- CATLIN-GROVES, C. L. 2012. The Citizen Science Landscape: From Volunteers to Citizen Sensors and Beyond. *International Journal of Zoology*, 2012, 349630.
- CHANACHAI, J., ASAMOAH, E. F., MAINA, J. M., WILSON, P. D., NIPPERESS, D. A., ESPERON-RODRIGUEZ, M. & BEAUMONT, L. J. 2024. What remains to be discovered: A global assessment of tree species inventory completeness. *Diversity and Distributions*, 30, e13862.
- CHOWDHURY, S., ROLL, U., AHMED, S., ROKONUZZAMAN, M., MANNA, M. U., TISHA, M. N., DAIYAN, M., WONG, B. B. & MEIRI, S. 2026. Bias and complementarity in species occurrence data from citizen science, museum, and social media collections. *EcoEvoRxiv*.

- CLARK, T. P., KAHAN, B. C., PHILLIPS, A., WHITE, I. & CARPENTER, J. R. 2022. Estimands: bringing clarity and focus to research questions in clinical trials. *BMJ Open*, 12, e052953.
- CRALL, A. W., NEWMAN, G. J., STOHLGREN, T. J., HOLFELDER, K. A., GRAHAM, J. & WALLER, D. M. 2011. Assessing citizen science data quality: an invasive species case study. *Conservation Letters*, 4, 433-442.
- CRETOIS, B., SIMMONDS, E. G., LINNELL, J. D. C., VAN MOORTER, B., ROLANDSEN, C. M., SOLBERG, E. J., STRAND, O., GUNDERSEN, V., ROER, O. & RØD, J. K. 2021. Identifying and correcting spatial bias in opportunistic citizen science data for wild ungulates in Norway. *Ecology and Evolution*, 11, 15191 - 15204.
- CRIMMINS, T. M., POSTHUMUS, E., SCHAFFER, S. & PRUDIC, K. L. 2021. COVID-19 impacts on participation in large scale biodiversity-themed community science projects in the United States. *Biological Conservation*, 256, 109017.
- DENNIS, E. B., MORGAN, B. J., BRERETON, T. M., ROY, D. B. & FOX, R. 2017. Using citizen science butterfly counts to predict species population trends. *Conservation biology*, 31, 1350-1361.
- DENNIS, E. B., MORGAN, B. J. T., FREEMAN, S. N., BRERETON, T. M. & ROY, D. B. 2016. A Generalized Abundance Index for Seasonal Invertebrates. *Biometrics*, 72, 1305-1314.
- DENNIS, E. B., MORGAN, B. J. T., HARROWER, C. A., BOURN, N. A. D. & FOX, R. 2024. Incorporating Phenology to Estimate Species' Population Trends from Snapshot Citizen-Science Data. *Journal of Agricultural, Biological and Environmental Statistics*.
- DENNIS, R. L. H., SHREEVE, T. G., ISAAC, N. J. B., ROY, D. B., HARDY, P. B., FOX, R. & ASHER, J. 2006. The effects of visual apparency on bias in butterfly recording and monitoring. *Biological Conservation*, 128, 486-492.
- DENNIS, R. L. H. & THOMAS, C. D. 2000. Bias in Butterfly Distribution Maps: The Influence of Hot Spots and Recorder's Home Range. *Journal of Insect Conservation*, 4, 73-77.
- DI CECCO, G. J., BARVE, V., BELITZ, M. W., STUCKY, B. J., GURALNICK, R. P. & HURLBERT, A. H. 2021. Observing the Observers: How Participants Contribute Data to iNaturalist and Implications for Biodiversity Science. *BioScience*, 71, 1179-1188.
- DÍAZ-CALAFAT, J., JAUME-RAMIS, S., SOACHA, K., ÁLVAREZ, A. & PIERA, J. 2024. Revealing biases in insect observations: A comparative analysis between academic and citizen science data. *PLOS ONE*, 19, e0305757.
- DICKINSON, J. L., ZUCKERBERG, B. & BONTER, D. N. 2010. Citizen Science as an Ecological Research Tool: Challenges and Benefits. *Annual Review of Ecology, Evolution, and Systematics*, 41, 149-172.
- DIMSON, M., BERIO FORTINI, L., TINGLEY, M. W. & GILLESPIE, T. W. 2023. Citizen science can complement professional invasive plant surveys and improve estimates of suitable habitat. *Diversity and Distributions*, 29, 1141 - 1156.
- DIMSON, M. & GILLESPIE, T. W. 2023. Who, where, when: Observer behavior influences spatial and temporal patterns of iNaturalist participation. *Applied Geography*, 153, 102916.
- ELLIOTT, M. R. & VALLIANT, R. 2017. Inference for nonprobability samples. *Statistical science*, 249-264.
- FINK, D., JOHNSTON, A., STRIMAS-MACKEY, M., AUER, T., HOCHACHKA, W. M., LIGOCKI, S., OLDHAM JAROMCZYK, L., ROBINSON, O., WOOD, C., KELLING, S. & RODEWALD, A. D. 2023. A Double machine learning trend model for citizen science data. *Methods in Ecology and Evolution*, 14, 2435-2448.
- FITHIAN, W., ELITH, J., HASTIE, T. & KEITH, D. A. 2015. Bias correction in species distribution models: pooling survey and collection data for multiple species. *Methods in ecology and evolution*, 6, 424-438.
- FONTAINE, B., BERGEROT, B., LE VIOL, I. & JULLIARD, R. 2016. Impact of urbanization and gardening practices on common butterfly communities in France. *Ecology and Evolution*, 6, 8174-8180.
- FOURCADE, Y., ENGLER, J. O., RÖDDER, D. & SECONDI, J. 2014. Mapping species distributions with MAXENT using a geographically biased sample of presence data: a performance assessment of methods for correcting sampling bias. *PloS one*, 9, e97122.
- FRAISL, D., HAGER, G., BEDESSEM, B., GOLD, M., HSING, P. Y., DANIELSEN, F., HITCHCOCK, C. B., HULBERT, J. M., PIERA, J., SPIERS, H., THIEL, M. & HAKLAY, M. 2022. Citizen science in environmental and ecological sciences. *Nature Reviews Methods Primers*, 2.
- GELDMANN, J., HEILMANN-CLAUSEN, J., HOLM, T. E., LEVINSKY, I., MARKUSSEN, B., OLSEN, K., RAHBK, C. & TØTTRUP, A. P. 2016. What determines spatial bias in citizen science? Exploring four recording schemes with different proficiency requirements. *Diversity and Distributions*, 22.
- GEURTS, E. M., REYNOLDS, J. D. & STARZOMSKI, B. M. 2023. Not all who wander are lost: Trail bias in community science. *PLOS ONE*, 18.
- GILLARD, G. L. & ROWLEY, J. J. L. 2026. The importance of private land for biodiversity conservation is highlighted by rapid, broad-scale sampling using citizen science. *Conservation Science and Practice*, 8, e70215.
- GOLDSTEIN, B. R., STOUT, S., LEWTHWAITE, J. M. M., SHIREY, V., MENDOZA, E. & GUZMAN, L. M. 2024. Logistical and preference bias in participatory science butterfly data. *Frontiers in Ecology and the Environment*, 22, e2783.
- GRADY, E. L., CAMPBELL, C. J., CALLAGHAN, C. T. & GURALNICK, R. P. 2026. iNaturalist Users Exhibit Distinct Spatiotemporal Sampling Preferences, with Implications for Biodiversity Science and Project Planning. *Citizen Science: Theory and Practice*, 11.
- GREGORY, R. D., VAN STRIEN, A., VORISEK, P., GMEIG MEYLING, A. W., NOBLE, D. G., FOPPEN, R. P. & GIBBONS, D. W. 2005. Developing indicators for European birds. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360, 269-288.
- GROVES, R. M. & LYBERG, L. 2010. Total Survey Error: Past, Present, and Future. *Public Opinion Quarterly*, 74, 849-879.
- GUETTÉ, A., CAILLAULT, S., PITHON, J., PAIN, G., DANIEL, H., MARCHADOUR, B. & BEAUJOUAN, V. 2022. Who and Where Are the Observers behind Biodiversity Citizen Science Data? Effect of Landscape Naturalness on the Spatial Distribution of French Birdwatching Records. *Land*, 11, 2095.

846 HABEL, J. C., HUEMER, P., SCHMITT, T., RÜDISSE, J. & ULRICH, W. 2025. Strengths and shortcomings of citizen science
847 data: lessons from Austrian butterflies. *Journal of Insect Conservation*, 29, 60.

848 HADJ-HAMMOU, J., LOISELLE, S. A., OPHOF, D. & THORNHILL, I. 2017. Getting the full picture: Assessing the
849 complementarity of citizen science and agency monitoring data. *PLoS ONE*, 12.

850 HOCHACHKA, W. M., ALONSO, H., GUTIÉRREZ-EXPÓSITO, C., MILLER, E. & JOHNSTON, A. 2021. Regional variation in the
851 impacts of the COVID-19 pandemic on the quantity and quality of data collected by the project eBird. *Biological*
852 *Conservation*, 254, 108974.

853 HÖLKER, F., WOLTER, C., PERKIN, E. K. & TOCKNER, K. 2010. Light pollution as a biodiversity threat. *Trends in Ecology &*
854 *Evolution*, 25, 681-682.

855 HORDLEY, L. A. & FOX, R. 2024. Wildlife-friendly garden practices increase butterfly abundance and species richness in
856 urban and arable landscapes. *Science of The Total Environment*, 929, 171503.

857 HORVITZ, D. G. & THOMPSON, D. J. 1952. A generalization of sampling without replacement from a finite universe. *Journal*
858 *of the American statistical Association*, 47, 663-685.

859 HUANG, X., LIN, C., JI, L. & FENG, G. 2023. Species inventories from different data sources "shaping" slightly different
860 avifauna diversity patterns. *Frontiers in Ecology and Evolution*, Volume 11 - 2023.

861 IDEC, J., CAMPBELL, C. J., BELITZ, M., AKSHAY, V. & GURALNICK, R. 2025. Using citizen science data to estimate trait and
862 climate drivers of daily activity patterns in temperate butterflies. *PLoS One*, 20, e0335856.

863 IRVINE, K. M., RODHOUSE, T. J., WRIGHT, W. J. & OLSEN, A. R. 2018. Occupancy modeling species-environment
864 relationships with non-ignorable survey designs. *Ecological Applications*, 28, 1616-1625.

865 ISAAC, N. J. & POCOCK, M. J. 2015. Bias and information in biological records. *Biological Journal of the Linnean Society*, 115,
866 522-531.

867 JACOBS, C. & ZIPF, A. 2017. Completeness of citizen science biodiversity data from a volunteered geographic information
868 perspective. *Geo-spatial Information Science*, 20, 3-13.

869 JOHNSON, T. F., ISAAC, N. J. B., PAVIOLO, A. & GONZÁLEZ-SUÁREZ, M. 2021. Handling missing values in trait data. *Global*
870 *Ecology and Biogeography*, 30, 51-62.

871 JOHNSTON, A., FINK, D., HOCHACHKA, W. M. & KELLING, S. 2018. Estimates of observer expertise improve species
872 distributions from citizen science data. *Methods in Ecology and Evolution*, 9, 88-97.

873 JOHNSTON, A., MATECHOU, E. & DENNIS, E. B. 2023. Outstanding challenges and future directions for biodiversity
874 monitoring using citizen science data. *Methods in Ecology and Evolution*, 14, 103-116.

875 JOHNSTON, A., MORAN, N., MUSGROVE, A., FINK, D. & BAILLIE, S. R. 2020. Estimating species distributions from spatially
876 biased citizen science data. *Ecological Modelling*, 422, 108927.

877 JUSTINE, J., AHIRBUDHNYAN, M., ASHIK, M. E. & BALAKRISHNAN, P. 2026. Citizen science in pollinator monitoring: current
878 approaches, challenges and recommendations. *Oikos*, 2026, e11594.

879 KC, S. & SAPKOTA, A. 2025. Can gamification save the planet? Revolutionizing citizen science for biodiversity conservation.
880 *Biological Conservation*, 302, 111001.

881 KELLING, S., JOHNSTON, A., HOCHACHKA, W. M., ILIFF, M., FINK, D., GERBRACHT, J., LAGOZE, C., LA SORTE, F. A., MOORE,
882 T., WIGGINS, A., WONG, W.-K., WOOD, C. & YU, J. 2015. Can Observation Skills of Citizen Scientists Be Estimated
883 Using Species Accumulation Curves? *PLOS ONE*, 10, e0139600.

884 KHOI, N. M. & CASTELEYN, S. 2018. Analyzing Spatial and Temporal User Behavior in Participatory Sensing. *ISPRS Int. J. Geo*
885 *Inf.*, 7, 344.

886 KOSMALA, M., WIGGINS, A., SWANSON, A. & SIMMONS, B. 2016. Assessing data quality in citizen science. *Frontiers in*
887 *Ecology and the Environment*, 14, 551-560.

888 KRAL-O'BRIEN, K. C., KARASCH, B. M., HOVICK, T. J., MORANZ, R. A. & HARMON, J. P. 2020. Morphological traits determine
889 detectability bias in North American grassland butterflies. *Ecosphere*, 11, e03304.

890 KULLENBERG, C. & KASPEROWSKI, D. 2016. What Is Citizen Science? – A Scientometric Meta-Analysis. *PLOS ONE*, 11,
891 e0147152.

892 LA SORTE, F. A. & SOMVEILLE, M. 2020. Survey completeness of a global citizen-science database of bird occurrence.
893 *Ecography*, 43, 34-43.

894 LA SORTE, F. A., COHEN, J. M. & JETZ, W. 2024. Data coverage, biases, and trends in a global citizen-science resource for
895 monitoring avian diversity. *Diversity and Distributions*, 30.

896 LI, M., BENNIE, J., ROY, D. B., FOX, R., BOYD, R. J., SMITH, C. & FFRENCH-CONSTANT, R. H. 2026a. Systematic differences in
897 phenology estimates from unstructured and structured biodiversity datasets. *Insect Conservation and Diversity*.

898 LI, M., BOYD, R. J., SMITH, C., FOX, R., ROY, D. & BENNIE, J. 2025. Citizen scientists as butterfly predators: using foraging
899 theory to understand individual recorder behaviour. *Ecological Modelling*, 510, 111344.

900 LI, M., FFRENCH-CONSTANT, R. H., BENNIE, J., FOX, R., ROY, D. B., SMITH, C. & BOYD, R. J. 2026b. Causal Analysis Reveals
901 That Public Access to Land Can Shape Our View of Species' Geographic Distributions. *Diversity and Distributions*,
902 32, e70242.

903 LOBO, J. M., GUTIÉRREZ-CÁNOVAS, C., FOX, R., DENNIS, R. L. H. & SÁNCHEZ-FERNÁNDEZ, D. 2021. Species traits
904 influence the process of biodiversity inventorying: a case study using the British butterfly database. *Insect*
905 *Conservation and Diversity*, 14, 748-755.

906 LOHR, S. L. 2021. *Sampling: design and analysis*, Chapman and Hall/CRC.

907 LUNDBERG, I., JOHNSON, R. & STEWART, B. M. 2021. What is your estimand? Defining the target quantity connects
908 statistical evidence to theory. *American Sociological Review*, 86, 532-565.

909 LUNDMARK, C. 2003. BioBlitz: Getting into backyard biodiversity. *BioScience*, 53, 329-329.

- MACKENZIE, D. I., NICHOLS, J. D., LACHMAN, G. B., DROEGE, S., ANDREW ROYLE, J. & LANGTIMM, C. A. 2002. Estimating site occupancy rates when detection probabilities are less than one. *Ecology*, 83, 2248-2255.
- MAIR, L. & RUETE, A. 2016. Explaining Spatial Variation in the Recording Effort of Citizen Science Data across Multiple Taxa. *PLoS ONE*, 11.
- MANDEVILLE, C. P., NILSEN, E. B. & FINSTAD, A. G. 2022. Spatial distribution of biodiversity citizen science in a natural area depends on area accessibility and differs from other recreational area use. *Ecological Solutions and Evidence*, 3, e12185.
- MCCARTHY, M. S., STEPHENS, C., DIEGUEZ, P., SAMUNI, L., DESPRÉS-EINSPENNER, M.-L., HARDER, B., LANDSMANN, A., LYNN, L. K., MALDONADO, N., ROČKAIOVÁ, Z., WIDNESS, J., WITTIG, R. M., BOESCH, C., KÜHL, H. S. & ARANDJELOVIC, M. 2021. Chimpanzee identification and social network construction through an online citizen science platform. *Ecology and Evolution*, 11, 1598-1608.
- MCGILL, B. J., ETIENNE, R. S., GRAY, J. S., ALONSO, D., ANDERSON, M. J., BENECHA, H. K., DORNELAS, M., ENQUIST, B. J., GREEN, J. L., HE, F., HURLBERT, A. H., MAGURRAN, A. E., MARQUET, P. A., MAURER, B. A., OSTLING, A., SOYKAN, C. U., UGLAND, K. I. & WHITE, E. P. 2007. Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. *Ecology Letters*, 10, 995-1015.
- MENG, X.-L. 2022. Comments on "Statistical inference with non-probability survey samples" - Miniaturizing data defect correlation: A versatile strategy for handling non-probability samples. STATISTICS CANADA, 100 TUNNEYS PASTURE DRIVEWAY, OTTAWA, ONTARIO K1A 0T6, CANADA.
- MENTGES, A., BLOWES, S. A., HODAPP, D., HILLEBRAND, H. & CHASE, J. M. 2020. Effects of site-selection bias on estimates of biodiversity change. *Conservation Biology*, 35.
- MILBERG, P., FRANZÉN, M., KARPATY WICKBOM, A., SVELANDER, S. & JOHANSSON, V. 2024. Pollinator activity and flowering in agricultural weeds in Sweden. *Ecology and Evolution*, 14.
- MILLAR, E. E., HAZELL, E. C. & MELLES, S. J. 2019. The 'cottage effect' in citizen science? Spatial bias in aquatic monitoring programs. *International Journal of Geographical Information Science*, 33, 1612 - 1632.
- MONDAIN-MONVAL, T. O., POCOCK, M. J. O., ROLPH, S., AUGUST, T., WRIGHT, E. & JARVIS, S. G. 2024. Adaptive sampling by citizen scientists improves species distribution model performance: A simulation study. *Methods in Ecology and Evolution*, 15, 1206 - 1220.
- NEWBOLD, T., HUDSON, L. N., HILL, S. L. L., CONTU, S., LYSENKO, I., SENIOR, R. A., BÖRGER, L., BENNETT, D. J., CHOIMES, A., COLLEN, B., DAY, J., DE PALMA, A., DÍAZ, S., ECHEVERRÍA-LONDOÑO, S., EDGAR, M. J., FELDMAN, A., GARON, M., HARRISON, M. L. K., ALHUSSEINI, T., INGRAM, D. J., ITESCU, Y., KATTGE, J., KEMP, V., KIRKPATRICK, L., KLEYER, M., CORREIA, D. L. P., MARTIN, C. D., MEIRI, S., NOVOSOLOV, M., PAN, Y., PHILLIPS, H. R. P., PURVES, D. W., ROBINSON, A., SIMPSON, J., TUCK, S. L., WEIHER, E., WHITE, H. J., EWERS, R. M., MACE, G. M., SCHARLEMANN, J. P. W. & PURVIS, A. 2015. Global effects of land use on local terrestrial biodiversity. *Nature*, 520, 45-50.
- NEYENS, T., DIGGLE, P. J., FAES, C., BEENAERTS, N., ARTOIS, T. & GIORGI, E. 2019. Mapping species richness using opportunistic samples: a case study on ground-floor bryophyte species richness in the Belgian province of Limburg. *Scientific Reports*, 9, 19122.
- NICHIFOREL, L., DEUFFIC, P., THORSEN, B. J., WEISS, G., HUJALA, T., KEARY, K., LAWRENCE, A., AVDIBEGOVIĆ, M., DOBŠ INSKÁ, Z., FELICIANO, D., GÓRRIZ-MIFSUD, E., HOOGSTRA-KLEIN, M., HRIB, M., JARSKÝ, V., JODŁOWSKI, K., LUKMINE, D., PEZDEVŠEK MALOVRH, Š., NEDELJKOVIĆ, J., NONIĆ, D., KRAJTER OSTOIĆ, S., PUKALL, K., RONDEUX, J., SAMARA, T., SARVAŠOVÁ, Z., SCRIBAN, R. E., ŠILINGIENĖ, R., SINKO, M., STOJANOVSKA, M., STOJANOVSKI, V., STOYANOV, T., TEDER, M., VENNESLAND, B., WILHELMSSON, E., WILKES-ALLEMANN, J., ŽIVOJINOVIĆ, I. & BOURIAUD, L. 2020. Two decades of forest-related legislation changes in European countries analysed from a property rights perspective. *Forest Policy and Economics*, 115, 102146.
- O'BRIEN, R. S. M., DAYER, A. A. & HOPKINS, W. A. 2021. Understanding landowner decisions regarding access to private land for conservation research. *Conservation Science and Practice*, 3, e522.
- O'NEILL, D., HÄKKINEN, H., NEUMANN, J., SHAFFREY, L. C., CHEFFINGS, C. M., NORRIS, K. & PETTORELLI, N. 2023. Investigating the potential of social media and citizen science data to track changes in species' distributions. *Ecology and Evolution*, 13.
- OSTERMANN-MIYASHITA, E. F., BLUHM, H., DOBIÁŠ, K., GANDL, N., HIBLER, S., LOOK, S., MICHLER, F.-U., WELTGEN, L., SMAGA, A., KÖNIG, H. J., KUEMMERLE, T. & KIFFNER, C. 2023. Opportunities and challenges for monitoring a recolonizing large herbivore using citizen science. *Ecology and Evolution*, 13.
- PATEMAN, R. M., DYKE, A. & WEST, S. E. 2021. The diversity of participants in environmental citizen science. *Citizen Science: Theory and Practice*.
- PERNAT, N., CANAVAN, S., GOLIVETS, M., HILLAERT, J., ITESCU, Y., JARIĆ, I., MANN, H. M. R., PIPEK, P., PREDÁ, C., RICHARDSON, D. M., TEIXEIRA, H., VAZ, A. S. & GROOM, Q. 2024. Overcoming biodiversity blindness: Secondary data in primary citizen science observations. *Ecological Solutions and Evidence*, 5, e12295.
- PESCOTT, O. L., POWNEY, G. D. & BOYD, R. J. 2025. Adaptive sampling for ecological monitoring using biased data: a stratum-based approach. *Oikos*, 2025, e11115.
- PETERSEN, T. K., SPEED, J. D. M., GRØTAN, V. & AUSTRHEIM, G. 2021. Species data for understanding biodiversity dynamics: The what, where and when of species occurrence data collection. *Ecological Solutions and Evidence*.
- PHILLIPS, S. J., DUDÍK, M., ELITH, J., GRAHAM, C. H., LEHMANN, A., LEATHWICK, J. R. & FERRIER, S. 2009. Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological applications : a publication of the Ecological Society of America*, 19 1, 181-97.

- PIMM, S. L., JENKINS, C. N., ABELL, R., BROOKS, T. M., GITTLEMAN, J. L., JOPPA, L. N., RAVEN, P. H., ROBERTS, C. M. & SEXTON, J. O. 2014. The biodiversity of species and their rates of extinction, distribution, and protection. *Science*, 344, 1246752.
- POCOCK, M. J. O., LOGIE, M., ISAAC, N. J. B., FOX, R. & AUGUST, T. 2023. The recording behaviour of field-based citizen scientists and its impact on biodiversity trend analysis. *Ecological Indicators*, 151, 110276.
- POTGIETER, L. J., CADOTTE, M. W., ROETS, F. & RICHARDSON, D. M. 2024. Monitoring urban biological invasions using citizen science: the polyphagous shot hole borer (*Euwallacea fornicatus*). *Journal of Pest Science*, 97, 2073 - 2085.
- RAMESH, V., GUPTA, P. R., TINGLEY, M. W., ROBIN, V. V. & DEFRIES, R. S. 2022. Using citizen science to parse climatic and land cover influences on bird occupancy in a tropical biodiversity hotspot. *Ecography*.
- ROBINSON, O. J., RUIZ-GUTIERREZ, V. & FINK, D. 2018. Correcting for bias in distribution modelling for rare species using citizen science data. *Diversity and Distributions*, 24, 460-472.
- ROSÁRIO, I. T., TIAGO, P., CHOZAS, S. & CAPINHA, C. 2024. Drivers of temporal bias in biodiversity recording by citizen scientists. *bioRxiv*.
- ROSÁRIO, I. T., TIAGO, P., CHOZAS, S. & CAPINHA, C. 2025. When do citizen scientists record biodiversity? Non-random temporal patterns of recording effort and associated factors. *People and Nature*, 7, 860-870.
- ROYLE, J. A. 2004. N-Mixture Models for Estimating Population Size from Spatially Replicated Counts. *Biometrics*, 60, 108-115.
- SÁNCHEZ-CLAVIJO, L. M., MARTÍNEZ-CALLEJAS, S. J., ACEVEDO-CHARRY, O., DIAZ-PULIDO, A., GÓMEZ-VALENCIA, B., OCAMPO-PEÑUELA, N., OCAMPO, D., OLAYA-RODRÍGUEZ, M. H., REY-VELASCO, J. C., SOTO-VARGAS, C. & OCHOA-QUINTERO, J. M. 2021. Differential reporting of biodiversity in two citizen science platforms during COVID-19 lockdown in Colombia. *Biological Conservation*, 256, 109077.
- SÁNCHEZ-FERNÁNDEZ, D., FOX, R., DENNIS, R. L. H. & LOBO, J. M. 2021. How complete are insect inventories? An assessment of the british butterfly database highlighting the influence of dynamic distribution shifts on sampling completeness. *Biodiversity and Conservation*, 30, 889-902.
- SANTORO, S., GUTIÉRREZ-ZAPATA, S., CALZADA, J., SELVA, N., MARÍN-SANTOS, D., BEERY, S., BRANDIS, K., FERNÁNDEZ DE VIANA, I., MEEK, P., MORTELLITI, A., REVILLA, E., RODRÍGUEZ, J. P., STRAKOVÁ, L., TENAN, S. & GEGÚNDEZ, M. E. 2025. Essential tools but overlooked bias: Artificial intelligence and citizen science classification affect camera trap data. *Methods in Ecology and Evolution*, 16, 2638-2652.
- SANTOS-FERNANDEZ, E. & Mengersen, K. 2021. Understanding the reliability of citizen science observational data using item response models. *Methods in Ecology and Evolution*, 12, 1533-1548.
- SHARPE, C. R., HILL, R. A., CHAPPELL, H. M., GREEN, S. E., HOLDEN, K., FERGUS, P., CHALMERS, C. & STEPHENS, P. A. 2025. Increasing citizen scientist accuracy with artificial intelligence on UK camera-trap data. *Remote Sensing in Ecology and Conservation*, 11, 641-655.
- SHIREY, V., BELITZ, M. W., BARVE, V. & GURALNICK, R. 2020. Closing Gaps But Increasing Bias In North American Butterfly Inventory Completeness. *bioRxiv*, 2020.07.20.212381.
- SHRESTHA, M., GARCIA, J. E., CHUA, J. H. J., HOWARD, S. R., TSCHULIN, T., DORIN, A., NIELSEN, A. & DYER, A. G. 2019. Fluorescent Pan Traps Affect the Capture Rate of Insect Orders in Different Ways. *Insects* [Online], 10.
- SICACHA-PARADA, J., STEINSLAND, I., CRETOIS, B. & BORGELT, J. 2019. Accounting for spatial varying sampling effort due to accessibility in Citizen Science data: A case study of moose in Norway. *Spatial Statistics*, 42, 100446.
- SOROYE, P., AHMED, N. & KERR, J. T. 2018. Opportunistic citizen science data transform understanding of species distributions, phenology, and diversity gradients for global change research. *Global change biology*, 24, 5281-5291.
- STOLAR, J. & NIELSEN, S. E. 2015. Accounting for spatially biased sampling effort in presence-only species distribution modelling. *Diversity and Distributions*, 21.
- STOUDT, S., GOLDSTEIN, B. R. & DE VALPINE, P. 2022. Identifying engaging bird species and traits with community science observations. *Proceedings of the National Academy of Sciences*, 119, e2110156119.
- 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. *Biological Conservation*, 169, 31-40.
- SWEET, F. S., RÖDL, T. & WEISSER, W. W. 2022. COVID-19 lockdown measures impacted citizen science hedgehog observation numbers in Bavaria, Germany. *Ecology and Evolution*, 12, e8989.
- THOMAS, A., ONKELINCKX, T., WESTRA, T., VAN DE POEL, S., DE KEERSMAEKER, L., LEDEGEN, H. & NEUKERMANS, A. 2022. A cost-efficient and accurate pheromone-baited monitoring protocol for *Elatér ferrugineus*, a biodiversity indicator of hollow trees. *Journal of Insect Conservation*, 26, 97-106.
- THORNHILL, I., LOISELLE, S., LIND, K. & OPHOF, D. 2016. The citizen science opportunity for researchers and agencies. *BioScience*, 66, 720-721.
- TIAGO, P., CEIA-HASSE, A., MARQUES, T. A., CAPINHA, C. & PEREIRA, H. M. 2017. Spatial distribution of citizen science casuistic observations for different taxonomic groups. *Scientific Reports*, 7.
- TOTTERMAN, S. L. 2021. A local scale evaluation of spatial sampling bias in the Atlas of Australian Birds. *bioRxiv*.
- TROUDET, J., GRANDCOLAS, P., BLIN, A., VIGNES-LEBBE, R. & LEGENDRE, F. 2017. Taxonomic bias in biodiversity data and societal preferences. *Scientific Reports*, 7, 9132.

- TULLOCH, A. I. T. & SZABO, J. K. 2012. A behavioural ecology approach to understand volunteer surveying for citizen science datasets. *Emu*, 112, 313-325.
- VAN DER WAL, R., SHARMA, N., MELLISH, C., ROBINSON, A. & SIDDHARTHAN, A. 2016. The role of automated feedback in training and retaining biological recorders for citizen science. *Conservation Biology*, 30, 550-561.
- VAN SWAAY, C., PLATE, C. L. & VAN STRIEN, A. 2002. Monitoring butterflies in the Netherlands: How to get unbiased indices. *Proc Exp Appl Entomol*, 13, 21-27.
- WANG, X. H. & WU, W. 2013. A Review of Environmental Management Systems in Global Defence Sectors. *American Journal of Environmental Sciences*, 9.
- WARD, D. F. 2014. Understanding sampling and taxonomic biases recorded by citizen scientists. *Journal of Insect Conservation*, 18, 753-756.
- WELVAERT, M. & CALEY, P. 2016. Citizen surveillance for environmental monitoring: combining the efforts of citizen science and crowdsourcing in a quantitative data framework. *SpringerPlus*, 5.
- WILLIAMS, R. L., STAFFORD, R. & GOODENOUGH, A. E. 2015. Biodiversity in urban gardens: Assessing the accuracy of citizen science data on garden hedgehogs. *Urban Ecosystems*, 18, 819-833.
- WILLIG, M. R., ROJAS-SANDOVAL, J. & PRESLEY, S. J. 2024. Phenological patterns in ecology: Problems using circular statistics and solutions based on simulations. *Methods in Ecology and Evolution*, 15, 868 - 885.
- WILSON, J. F., BAKER, D., COOK, M., DAVIS, G., FREESTONE, R., GARDNER, D., GRUNDY, D., LOWE, N., ORRIDGE, S. & YOUNG, H. 2015. Climate association with fluctuation in annual abundance of fifty widely distributed moths in England and Wales: a citizen-science study. *Journal of Insect Conservation*, 19, 935-946.
- WILSON, J. S., GRISWOLD, T. & MESSINGER, O. J. 2008. Sampling bee communities (Hymenoptera: Apiformes) in a desert landscape: are pan traps sufficient? *Journal of the Kansas Entomological Society*, 81, 288-300.
- XUE, Y., DAVIES, I., FINK, D., WOOD, C. & GOMES, C. P. Avicaching: A two stage game for bias reduction in citizen science. Proceedings of the 2016 international conference on autonomous agents & multiagent systems, 2016. 776-785.
- ZHANG, E. Y., BALDWIN, A., HUNDLEY, C., CHANG, E., AUDERSET, S., BAWENDI, M. & KRISTENSEN, T. V. 2024. Influence of an extreme event—the COVID-19 pandemic—On establishment of and data collection by a citizen science project. *PLOS ONE*, 19.
- ZHANG, G. 2020. Spatial and Temporal Patterns in Volunteer Data Contribution Activities: A Case Study of eBird. *ISPRS Int. J. Geo Inf.*, 9, 597.
- ZHENG, K., WONG, M. K. L., TSANG, T. P. N. & LEONG, C. M. 2025. Bridging Citizen Science and Expert Surveys in urban biodiversity monitoring: Insights from insect diversity in Macao. *Biodiversity Data Journal*, 13, e153402.