

ARTICLE

Methods, Tools, and Technologies

Quantifying effort needed to estimate species diversity from citizen science data

Corey T. Callaghan^{1,2}  | Diana E. Bowler^{1,3,4} | Shane A. Blowes^{1,5}  |
Jonathan M. Chase^{1,5}  | Mitchell B. Lyons⁶  | Henrique M. Pereira^{1,2,7} 

¹German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany

²Institute of Biology, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany

³Institute of Biodiversity, Friedrich Schiller University Jena, Jena, Germany

⁴Department of Ecosystem Services, Helmholtz Center for Environmental Research-UFZ, Leipzig, Germany

⁵Institute of Computer Science, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany

⁶Centre for Ecosystem Science, School of Biological, Earth and Environmental Sciences, UNSW Sydney, Sydney, New South Wales, Australia

⁷CIBIO (Research Centre in Biodiversity and Genetic Resources)–InBIO (Research Network in Biodiversity and Evolutionary Biology), Universidade do Porto, Vairão, Portugal

Correspondence

Corey T. Callaghan
Email: c.callaghan@unsw.edu.au

Handling Editor: Debra P. C. Peters

Abstract

Broad-scale biodiversity monitoring relies, at least in part, on the efforts of citizen, or community, scientists. To ensure robust inferences from citizen science data, it is important to understand the spatial pattern of sampling effort by citizen scientists and how it deviates from an optimal pattern. Here, we develop a generalized workflow to estimate the optimal distribution of sampling effort for inference of species diversity (e.g., species richness, Shannon diversity, and Simpson's diversity) patterns using the relationship between species diversity and land cover. We used data from the eBird citizen science project that was collected across heterogeneous landscapes in Florida (USA) to illustrate this workflow across different grain sizes. We found that a relatively small number of samples are needed to meet 95% sampling completeness when diversity estimation is focused on dominant species: 43, 64, 96, 123, 172, and 176 for 5×5 , 10×10 , 15×15 , 20×20 , 25×25 , and 30×30 -km² grain sizes, respectively. In contrast, three to five times more samples are necessary to infer species diversity when estimation is focused on rare species. However, in both cases, the optimal distribution of effort was spatially heterogeneous, with more effort needed in regions of higher diversity. Our results highlight the potential of citizen science data to make informed comparisons of species diversity in space and time, as well as how sampling effort inherently depends on monitoring goals, such as whether dominant or rare species are targeted. Our general workflow allows for the quantification of sampling effort needed to estimate species diversity with citizen science data and can guide future adaptive sampling by citizen science participants.

KEYWORDS

biodiversity monitoring, biodiversity sampling, citizen science, community science, eBird, species richness

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *Ecosphere* published by Wiley Periodicals LLC on behalf of The Ecological Society of America.

INTRODUCTION

With the development of open science, large-scale biodiversity databases have enabled better estimates of the distribution and abundance of species, enhancing our understanding of biodiversity patterns in time and space. For example, there are currently more than 1.6 billion species observations in the Global Biodiversity Information Facility (GBIF; www.gbif.org) and the amount of data in GBIF has increased 12-fold since 2007 (Heberling et al., 2021). Despite this rapid growth of global biodiversity data, estimates of biodiversity in many parts of the world remain at best imprecise and at worst nonexistent (Boakes et al., 2010; Cornwell et al., 2019; Essl et al., 2013; Scheffers et al., 2012; Stork, 1993). These geographic differences largely stem from variation in sampling effort. Monitoring biodiversity is increasingly important in the face of anthropogenic changes and associated ongoing biodiversity losses (Butchart et al., 2010; IPBES, 2019) and plays a core role in indicator development for conservation decision-making, such as the Post-2020 Global Biodiversity Framework of the Convention on Biological Diversity (Xu et al., 2021). Nevertheless, biodiversity monitoring is expensive and time-consuming, with generally insufficient funding (Bakker et al., 2010).

Future biodiversity monitoring will likely rely, at least in part, on broad-scale citizen science (also called community science) data (Bonney et al., 2009; Chandler et al., 2017; McKinley et al., 2017; Pereira et al., 2017). Already, citizen science data have proven useful for a wide variety of applications, including monitoring marine protected areas (Freiwald et al., 2018), developing conservation strategies for migratory species of birds (Schuster et al., 2019), highlighting the negative impacts of roads on amphibians (Cosentino et al., 2014), enabling continental-scale population trend estimation of birds (Brlík et al., 2021), rapidly mapping the effects of bushfires on biodiversity (Kirchhoff et al., 2021), and understanding the effects of the COVID-19 “anthropause” on biodiversity (Vardi et al., 2021). Clearly, the scope and extent of citizen science projects and the resulting data are rapidly increasing (Pocock et al., 2017). In 2020, for example, the web app iNaturalist (www.inaturalist.org), which allows naturalist observers to take pictures and upload georeferenced observations, averaged about 63,000 observations per day. As the quantity of data continues to increase, the potential of these data for biodiversity research also increases.

Biodiversity-focused citizen science projects cover a broad range of biodiversity data collection activities (Isaac & Pocock, 2015; Kelling et al., 2019; Welvaert & Caley, 2016). Generally, projects range from unstructured (e.g., no formal training, very few instructions, and

observations can take place wherever and whenever; e.g., iNaturalist) to semistructured (e.g., minimal training and few instructions, but important metadata are collected as part of the observation process; e.g., eBird) to structured (e.g., formally trained, experienced, participants who conduct repeated surveys in time and space; e.g., Reef Life Survey). For those projects that allow participants to select their own sites to survey biodiversity, sites are rarely selected at random. For instance, there are often hotspots of data contributions from urban areas and their surroundings, highly accessible areas, or sites within nature reserves (Hugo & Altwegg, 2017; Husby et al., 2021; Tiago, Ceia-Hasse, et al., 2017). As the number of participants in broad-scale citizen science projects continues to grow, there is increased potential for redundancy of effort as new participants sample the same set of sites already sampled by others (Callaghan, Poore, et al., 2019). In this light, there is a need to better understand the data collection patterns of citizen science participants, allowing for the development of guidelines for how biodiversity sampling might be improved for the estimation of biodiversity metrics across scales (Callaghan et al., 2021; Callaghan, Rowley, et al., 2019).

Diversity estimates, and in particular, species richness, play an important role in conservation spatial planning. For example, diversity estimates (e.g., species richness, Shannon diversity, and Simpson’s diversity) are often used to identify priority areas for protection in conservation-planning decisions ranging from local (e.g., patch-specific) to landscape (e.g., municipality) scales (Boyd et al., 2008). Species richness is the most intuitive, easily calculated, and probably the most commonly used diversity metric in studies of biodiversity patterns (Gotelli & Colwell, 2001; Magurran & McGill, 2011). Moreover, there are a number of techniques that allow for the extrapolation and comparison of species richness, such as rarefaction (e.g., Chao et al., 2009; Colwell & Coddington, 1994; McGlenn et al., 2019). Therefore, despite potential drawbacks (Fleishman et al., 2006; Hillebrand et al., 2018), species richness remains a useful metric that is possible to estimate with citizen science data. Indeed, species richness measured from citizen science data has been shown to be comparable with professionally collected data (Roman et al., 2017; Van der Wal et al., 2015).

Here, we aimed to demonstrate the applicability of citizen science data to estimate species diversity (e.g., species richness, Shannon diversity, and Simpson’s diversity, additional measures that are less sensitive to rare species) at multiple spatial scales and to show how citizen science sampling effort could be guided toward improving the accuracy of diversity estimates. We developed a generalized workflow to estimate the optimal distribution of sampling effort for inference of species

diversity patterns. This workflow can be continuously updated as data are collected and thus can be used to guide the site selection of participants to improve data collection of citizen science projects. To achieve this goal, we had three specific objectives: (1) test relationships between species diversity, land-cover attributes (e.g., habitat heterogeneity and urban cover), and sampling effort; (2) quantify the relationship between sample completeness (i.e., the probability that a new species would be found if sampling continued; Chao & Jost, 2012) and land-cover attributes; and (3) predict optimal effort (i.e., number of citizen science observations) needed to infer species diversity. Because estimates of biodiversity and its change in space and time are themselves scale-dependent (Chase et al., 2019), we examined how our results vary with spatial grain, testing the ability of citizen science data to assess species diversity at different spatial grains. Together, these analyses provide a useful template for optimizing biodiversity sampling across heterogeneous landscapes.

METHODS

General overview of methods

Our methodological framework can be broken down into three parts, aligning to the objectives mentioned above: (1) test relationships between species diversity, land-cover attributes (e.g., habitat heterogeneity and urban cover), and sampling effort; (2) quantify the relationship between sample completeness and land-cover attributes; and (3) predict optimal effort needed to infer species diversity.

For Objective 1, we first explored the patterns in the sampling effort of citizen science participants. Sampling effort was quantified as the number of samples submitted by citizen science participants. We used piecewise structural equation models (SEMs) to test the effects of land use on real-world effort at a given site and how these effects were mediated by observed species richness.

For Objective 2, we used rarefaction and extrapolation-based tools (e.g., Chao et al., 2014) to quantify sample completeness at a site and then used multiple regression to quantify the relationship between sample completeness and land-cover attributes. Sample completeness is a statistical estimate of the likelihood that a new species would be found if sampling continued at a given site. High values of sample completeness equate to a low probability of a new species being found, and vice versa. We calculated two estimates of sampling completeness: (1) *standardized completeness*, achieved by bootstrapping random subsamples of samples at a site,

thereby standardizing sampling effort among sites prior to calculating completeness, and (2) *total completeness*, calculated using all samples at a site. In addition, to differentiate between the sampling effort required to estimate different types of species diversity, we used three points along the continuum of a sampling completeness profile (Chao et al., 2020): (1) $q = 2$, where dominant, or highly abundant, species were the focus of species diversity estimation; (2) $q = 1$, where each species is weighted by its relative abundance and neither dominant or rare species being favored; and (3) $q = 0$, where rare species were more important to estimate species diversity.

For Objective 3, we used the relationships between land use, sampling completeness, and the number of samples to predict the effort (i.e., number of citizen science samples) needed to infer species diversity across the whole landscape. Since our aim here was prediction, we used random forest models, but we note that other forms of models (e.g., generalized linear or additive models) could be used in this step.

eBird citizen science data

We used data from eBird because it is one of the most successful citizen science projects to date, with >1 billion global observations (Sullivan et al., 2009, 2014). eBird is a semistructured project, where volunteer birdwatchers submit their observations in the form of “checklists” and indicate whether or not they recorded every species they were able to identify, that is, a “complete” checklist. In addition, useful information regarding sampling effort, such as the distance traveled, the duration of observation, time of observation, and the spatial coordinates, is recorded. Regional filters are also used to flag unusual bird observations, meaning if an observer sees a species or records a count of species outside of the filters, then the data are thoroughly reviewed by regional reviewers before being added to the database (Gilfedder et al., 2019).

As a test region, we used eBird data from peninsular Florida in the United States (Bird Conservation Region 31; <https://nabci-us.org/resources/bird-conservation-regions-map>). This region is relatively well sampled and includes diverse habitats, including a large urban area and large swaths of protected areas mixed with agriculture, forested, and wetland areas, but also contains regions with few or no eBird samples. We downloaded the eBird basic dataset (version eBd_May_2020) and subsetted the data to include all of 2019 (from 1 January to 31 December). Because the distance and time spent on a given eBird checklist can vary drastically (Kelling et al., 2015), we further subsetted the checklists by (1) including only checklists that were “complete”; (2) including checklists that were <120 min

and >15 min in duration; and (3) including checklists that traveled <1 km or sampled an area <150 ha. Although we focused on 2019 because it had the highest amount of data, we performed analyses for years 2014–2018 and found qualitatively and quantitatively similar patterns.

Because sites are not predefined in eBird, and different locations can correspond to drastically different sized habitat patches, we aggregated eBird checklists into grids. We used grids with six different spatial resolutions corresponding with 5×5 , 10×10 , 15×15 , 20×20 , 25×25 , and 30×30 -km² grain sizes. Only grid cells that had a minimum of 25 eBird checklists were used for analyses.

Land-cover habitat attributes

At each site across our different grain sizes, we quantified four land-cover attributes: habitat heterogeneity, percent cover of urban habitat, percent cover of trees, and percent cover of water. We used the Copernicus Global Land Cover Layers (Buchhorn et al., 2020), which is a 100-m pixel resolution product that maps discrete land-cover classes, as well as fractional cover of major structural types (bare ground, crops, grass, shrub, tree, urban, permanent water, and seasonal water). This provided sufficient detail for our case study site (i.e., Florida) and additionally allows for transferability to other regions in future studies. Habitat heterogeneity was calculated as the variance of the discrete land-cover classes within the 5×5 , 10×10 , 15×15 , 20×20 , 25×25 , and 30×30 -km² grid cells, and the fractional cover of the structural types was calculated as the mean within each grid cell. All processing for these metrics was done using Google Earth Engine (Gorelick et al., 2017).

Understanding patterns of real-world effort

We used the total observed species richness, land-cover habitat attributes (described above), and sampling effort (i.e., number of eBird checklists) in a site to quantify the direct and indirect relationships between these variables in a SEM framework. Our SEM consisted of two a priori linear models with a Gaussian error distribution: (1) log-transformed number of eBird checklists as the response variable and observed species richness, urban cover, and habitat heterogeneity as predictor variables; and (2) observed species richness as the response variable and habitat heterogeneity, urban cover, water cover, and tree cover as predictor variables. The SEM was fitted using the R package piecewiseSEM v2.1.2 (Lefcheck, 2016), and we present the results for the standardized coefficient estimates representing the strength of the various relationships.

Relationship between sample completeness and land-cover attributes

To quantify sample completeness, we used the iNEXT R package (Chao et al., 2014; Chao et al., 2020; Hsieh et al., 2016). As most citizen science platforms document occurrence (i.e., generates presence-only data), rather than abundance of species (e.g., iNaturalist), we calculated sample completeness for incidence data (Chao et al., 2020), by converting eBird data to presence/absence data (though we note that a similar approach could be used with abundance data; Chao & Jost, 2012). We focused on three points along the Hill number continuum: $q = 0$ (i.e., species richness), $q = 1$ (Shannon diversity), and $q = 2$ (i.e., Simpson's diversity). When $q = 0$, sample completeness is the ratio of observed species richness to true species richness; this metric does not consider species frequencies and thus is most sensitive to rare species. When $q = 1$, each species is weighted by its relative abundance and thus does not favor dominant or rare species in its estimation. When $q = 2$, sample completeness is disproportionately sensitive to species that occur in high frequencies (i.e., dominant species) (Chao et al., 2020). Accordingly, throughout our results we refer to our use of $q = 0$, $q = 1$, and $q = 2$ as “rare species sensitive,” “common species sensitive,” and “dominant species sensitive,” respectively, to represent species diversity. Because the results for $q = 1$ and $q = 2$ were qualitatively and quantitatively similar, we visualize the results of $q = 0$ (“rare species sensitive”) and $q = 2$ (“dominant species sensitive”) in the main text and present the results of $q = 1$ in the supporting information.

Similar to species diversity estimates, estimates of sample completeness are sensitive to sampling effort. Therefore, to quantify the relationship between sample completeness and land-cover characteristics, we standardized the number of checklists per site in our sample completeness calculation. To standardize sampling effort across grid cells of the same size, an equal number of checklists (5, 10, 15, and 20) were resampled ($N = 50$) randomly without replacement from each site, and the sampling completeness was then estimated as described above and averaged across the resamples. The analysis was robust to the number of checklists chosen each time (Appendix S1: Figure S1), and thus, we only present the results using 10 randomly chosen checklists for each sample. This standardized measure of completeness indirectly represents the size of the species pool. For equivalent sampling effort, we expect to sample a smaller proportion of species in a large, compared to a small species pool, meaning that lower values of standardized completeness indicate a larger species pool. Conversely, higher values of standardized completeness indirectly indicate a smaller species pool.

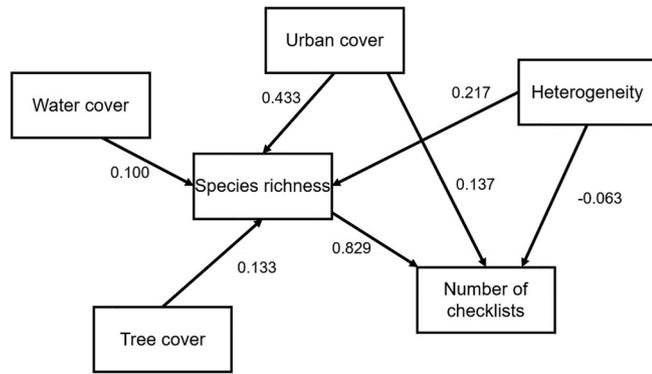


FIGURE 1 The results of our structural equation modeling, where the arrow goes from the predictor to the response variable, and the numbers represent standardized regression coefficients. These results represent the mean standardized regression coefficient across all grain sizes

We used linear models with a Gaussian distribution to assess the relationship between standardized completeness and land-cover habitat predictors, where the mean sampling completeness was the response variable and the predictor variables were the land-cover attributes (i.e., habitat heterogeneity, tree cover, water cover, and urban cover).

Predicting effort needed to infer species richness

To predict how many checklists were required to achieve a target completeness, we used a random forest model, where the response variable was the total number of observed log10-transformed eBird checklists, and the predictor variables were the total completeness estimate,

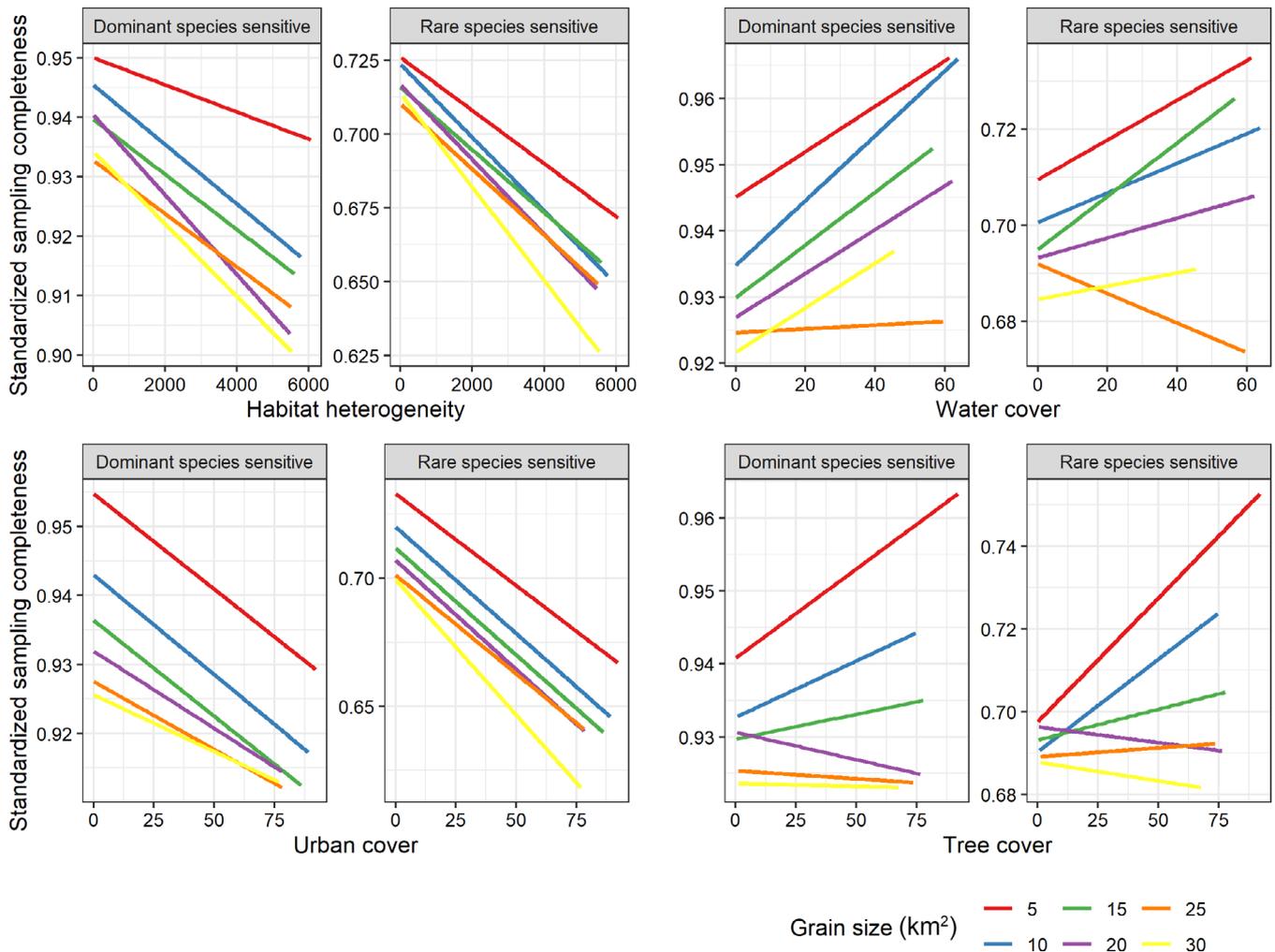


FIGURE 2 The relationship between bootstrapped sampling completeness (y-axis) and four land-cover habitat attributes (x-axis), with habitat heterogeneity (top left), water cover (top right), urban cover (bottom left), and tree cover (bottom right) for our two measures of completeness (i.e., dominant species sensitive and rare species sensitive). The relationship represents a linear model fit for each grain size from 5 × 5 to 30 × 30 km². Note the different scales for the y-axis

habitat heterogeneity, tree cover, water cover, and urban cover. For this objective, we used the total completeness, repeating the analysis above, but calculating only one measure of sampling completeness, where all available eBird checklists were included for each site (i.e., grid cell) to differentiate the difference between well-sampled sites and poorly sampled sites, irrespective of sampling effort. This differs from Objective 1 (see above) that used standardized completeness (Appendix S1: Figure S2).

We chose a random forest model because our aim here was prediction rather than understanding the effects of the covariates. Our goal was to make sites comparable for the purpose of comparing species diversity in space and not necessarily to estimate the total diversity (Chao et al., 2020). Therefore, we predicted the number of eBird checklists necessary to reach a completeness of 0.95 using our fitted random forest models. Exploratory analyses showed that there were marginal differences between completeness values of 0.95, 0.9, 0.85, and 0.8. This analysis was repeated three times: once for completeness sensitive to rare species ($q = 0$), once for completeness sensitive to common species ($q = 1$), and once for completeness sensitive to dominant species ($q = 2$). To assess the predictive performance of our random forest models, we report the mean pseudo- R^2 value, across all decision trees, extracted from the model fit. This pseudo- R^2 value represents the mean squared error divided by the variance of our response variable. Random forest analysis was done using the R package randomForest v4.6-14 (Liaw & Wiener, 2002). Models were fit without replacement and using the recommended settings for regression trees, which was 500 trees and a node size of 5.

Data analysis and availability

Analyses were conducted in R (R Core Team, 2020) and relied heavily on the tidyverse (Wickham et al., 2019). Statistical significance, in the case of multiple linear regressions and SEMs, was inferred at $\alpha < 0.05$. Code and data to reproduce these analyses are available here: <https://doi.org/10.5281/zenodo.5734200>.

RESULTS

We used a total of 138,703 eBird checklists comprising 450 species' observations for 2019 throughout peninsular Florida. The number of grid cells included in analysis was 933, 537, 343, 243, 175, and 132 for 5×5 , 10×10 , 15×15 , 20×20 , 25×25 , and 30×30 -km² grain sizes, respectively.

Our SEM showed a strong influence of land-cover in predicting species richness, with urban cover the strongest supported, followed by habitat heterogeneity, tree cover, and water cover (Figure 1, Appendix S1: Figure S3). The number of checklists at a site was also predicted by the percentage of urban cover and habitat heterogeneity, suggesting that these two land-cover attributes influence where people submit eBird checklists. The number of checklists was also higher in grids with higher species richness. These patterns showed some variation with grain size (Appendix S1: Figure S4).

Among these predictors, habitat heterogeneity and urban cover were consistently the strongest predictors of standardized completeness for both common and rare species diversity analyses, and across grain sizes. Standardized completeness was negatively correlated with habitat heterogeneity, water cover, and urban cover across grain sizes, indicating consistent patterns. In other words, standardized completeness tended to be higher at sites with lower heterogeneity, water cover, and urban cover. However, urban cover seemed to be a stronger

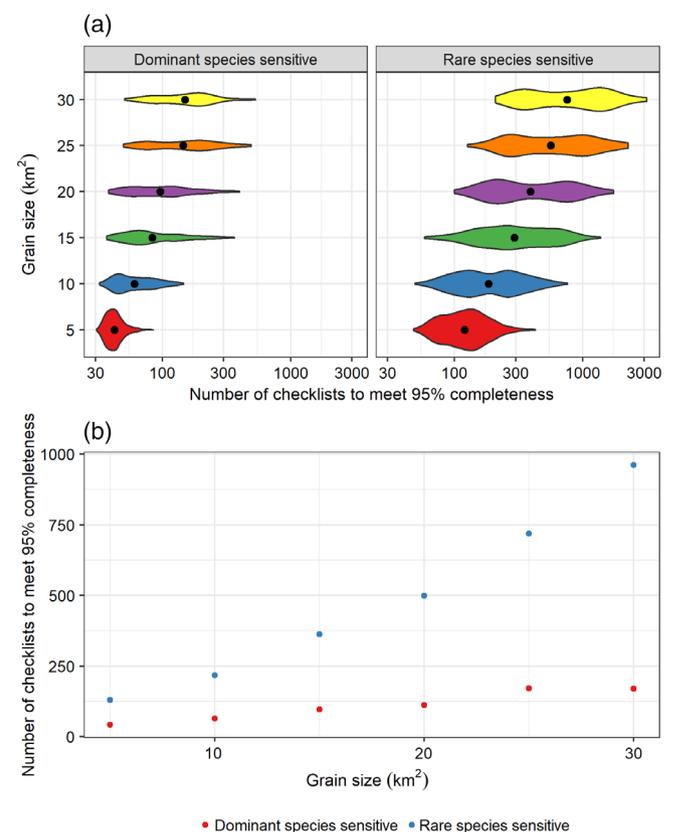


FIGURE 3 (a) The predicted number of checklists (log₁₀ scale) necessary to meet 95% completeness when completeness is calculated sensitive to common species (left) and sensitive to rare species (right). The black dot represents the mean value. (b) The mean number of checklists necessary to meet 95% completeness (y-axis) against grain size (x-axis) where grain size is not on the log₁₀ scale. Grain size represents 5×5 to 30×30 -km² grid cells

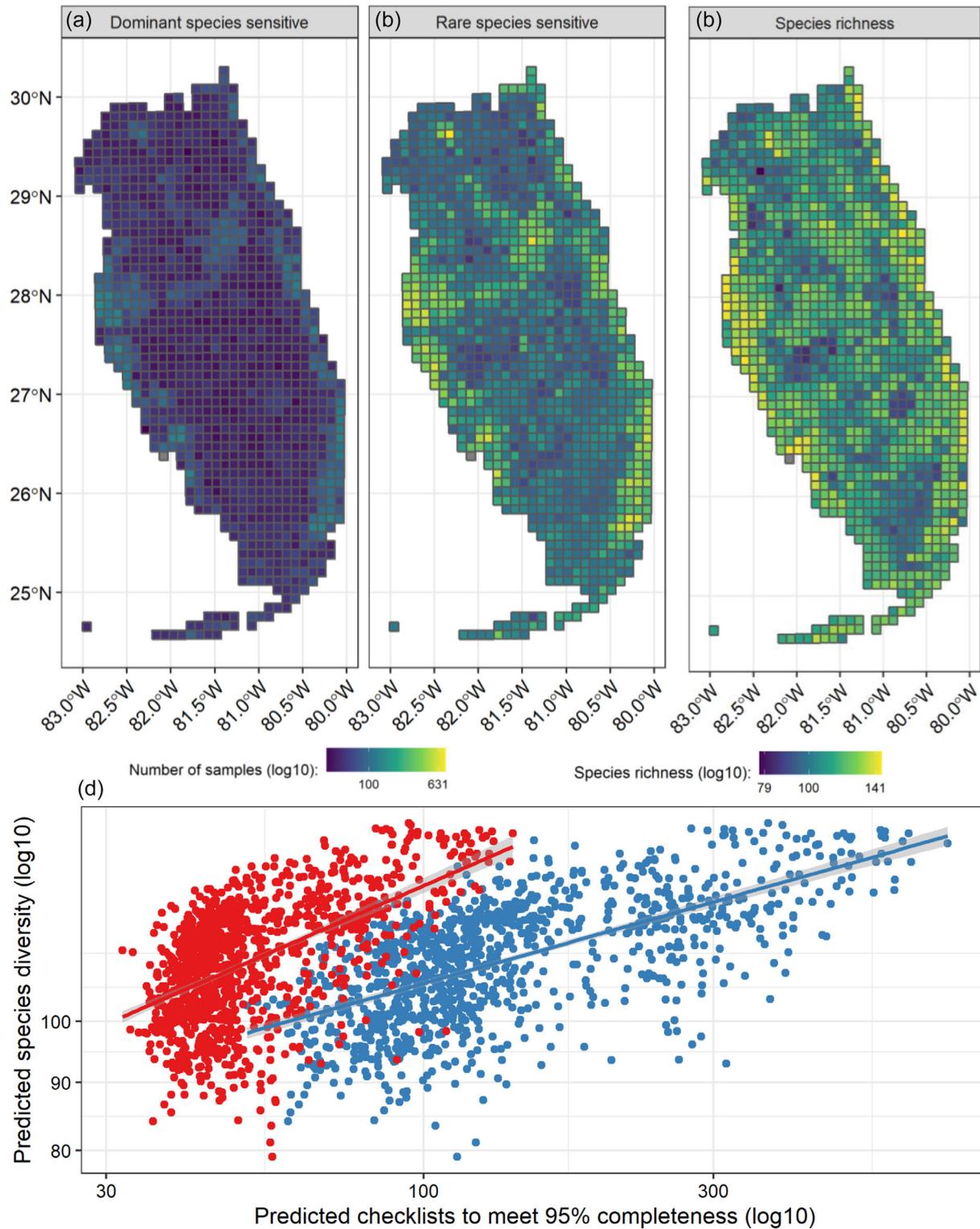


FIGURE 4 Spatial representation of the number of samples necessary to meet 95% completeness for (a) when completeness is calculated toward dominant species sensitive and (b) toward rare species sensitive. (c) A map of species richness throughout peninsular Florida, where species richness was predicted from a random forest model as described in the methods, where observed species richness was the response variable and the land-cover attributes were the predictor variables. (d) There was a generally positive relationship between species diversity and the necessary number of samples needed to reach 95% completeness for dominant species (red) and rare species (blue) diversity measures. Results are shown for the 10-km² grain size only

predictor for the analysis sensitive to rare species compared with the analysis sensitive to common species (Appendix S1: Figure S5). While other predictors showed consistent patterns among grain sizes, tree cover was negatively correlated with standardized completeness at smaller grain sizes, but the strength of this relationship weakened as grain size increased (Figure 2).

We found that on average, the number of checklists necessary to meet 95% sample completeness when calculating total completeness at $q = 2$ (dominant species sensitive) was 43, 64, 96, 123, 172, and 176 for 5×5 , 10×10 , 15×15 , 20×20 , 25×25 , and 30×30 -km² grain sizes, respectively. However, when calculating total completeness at $q = 0$ (rare species sensitive), the average number of eBird checklists was 129, 214, 362, 498, 712, and 946 at 5×5 , 10×10 , 15×15 , 20×20 , 25×25 , and 30×30 -km² grain sizes (Figure 3a). The results for $q = 1$ (common species sensitive) were quantitatively similar to that of $q = 2$ (Appendix S1: Figures S6 and S7). The uncertainty (i.e., standard deviation) in these average estimates was higher when calculating total completeness at $q = 0$ (Figure 3, Appendix S1: Table S1). The relationship between grain size and mean number of samples necessary to meet 95% total completeness increased much faster for rare species compared with dominant species, where it appeared to level off at larger grain sizes (Figure 3b).

Our models to predict the number of checklists necessary to sample in space performed relatively well (mean $R^2 = 0.83$; range = 0.79–0.86; Appendix S1: Table S2) for dominant species sensitive measures ($q = 2$), but less well (mean $R^2 = 0.39$; range = 0.07–0.58; Appendix S1: Table S2) for rare species sensitive measures. Our unsampled sites generally had a narrower distribution of species diversity values than our sampled sites, where our sampled sites included some of the sites that needed the highest number of checklists to meet 95% completeness, suggesting that data contributed to eBird are already originating from the most diverse sites (Appendix S1: Figure S8). We found a strong relationship between the predicted number of eBird checklists and the predicted species richness, confirming that our separate random forest models correlated well, as expected. The highest number of necessary samples was along the coast, and the lowest number was inland in large homogenous areas such as throughout the Florida Everglades, a large wetland complex (Figure 4).

DISCUSSION

We used more than 100,000 citizen science sampling events (i.e., eBird checklists) to quantify the relationships between sampling effort, species diversity, land-cover attributes, and sampling completeness. We found that

observed species richness was strongly modified by land cover, with a strong effect of urban cover and habitat heterogeneity (Figure 1). Moreover, we demonstrate how it is possible to predict necessary sampling effort needed to quantify species diversity in space. When the species diversity estimation is focused on dominant species (i.e., $q = 2$), a relatively small number of samples are needed to meet 95% sampling completeness, with an average of 44 samples in a 5×5 -km² grid cell and 203 samples in a 30×30 -km² grid cell. Yet, if the species diversity estimation is focused on rare species (i.e., $q = 0$ or species richness), substantially more samples are needed, 322 to 530% more samples at 5×5 and 30×30 -km² grids, respectively. Ultimately, our results highlight the potential of citizen science data to monitor species diversity and we provide a framework for future adaptive sampling by citizen science participants.

Biodiversity monitoring frequently relies on stratified sampling designs, where samples are spread across the landscape in an “equal” fashion (Bibby, 2004) to achieve sampling to inform biodiversity estimation. Sometimes, the number of samples is dependent on the size of a region of interest, with larger geographic regions receiving proportionately more sampling effort (Pavlacky Jr et al., 2017; Van Wilgenburg et al., 2020). Although funding and accessibility may be an issue for professionally designed biodiversity monitoring schemes, biodiversity is not spread equally throughout the landscape. Indeed, our results showed that species richness varies considerably in space (Figure 4c), and consequently, sampling effort should spatially match that of the underlying species diversity (Figure 4). Moreover, monitoring schemes often differ in terms of whether they target common, widespread species, or rare species. The optimal sampling strategy can vary depending on whether common or rare species are the goal of the monitoring scheme (e.g., Pacifici et al., 2016; Sanderlin et al., 2014; Sgarbi et al., 2020). For example, Specht et al. (2017) recommend random sampling for common species and conditional sampling for rare species, where more replicates are conducted when rare species are detected. These findings are consistent with our results since more spatially uniform patterns of effort were effective when species diversity estimation was sensitive to common species, while more heterogeneous effort was needed when species diversity estimation was more sensitive to rare species (Figure 4). The challenge for biodiversity monitoring is aligning these efforts for sampling of the entire community (i.e., both common and rare species simultaneously). Quantifying where more surveys are needed is an important first step, after which data gaps can be filled from different sources of data, such as those from professionals or volunteers, for example, by incentivizing

volunteers to sample for the common species in addition to the rare species (Tulloch et al., 2013).

Our analysis demonstrates the unequal sampling that takes place across the landscape, either because of biases in site selection, but also potentially because citizen science participants are preselecting sites with the highest biodiversity and habitat heterogeneity. This suggests that the pattern of citizen science sampling is already directed toward sites most likely to represent the highest levels of species diversity. Repeat sampling by citizen science participants at the same sites is not necessarily redundant sampling if these sites are the most species rich, and repeated sampling at these hotspots can be important for improving the robustness of statistical models (Callaghan, Rowley, et al., 2019) and even necessary for occupancy-detection model estimation (Sanderlin et al., 2014).

A future goal of citizen science projects might be to encourage sampling, for example, through gamification (Callaghan, Poore, et al., 2019), in locations where the samples are most valuable for a given statistical objective, for example, species diversity estimation. By understanding and quantifying the necessary amount of sampling, we can begin to reduce redundancies in the data (Boakes et al., 2010; Courter et al., 2013), for example, by “directing” participants. As an illustration of the potential of adaptive sampling, using a $10 \times 10\text{-km}^2$ grain size ($N = 1113$ grid cells) and focusing on an estimate of the number of relatively dominant species, on average, an average grid cell needs 68 samples to meet 95% completeness. With a total of 138,703 eBird checklists submitted during 2019, each grid cell *would* have received 125 eBird samples if effort were distributed equally across space. Yet, only 537 grids had 25 or more eBird samples, illustrating the highly spatially skewed effort, but also highlighting the potential of citizen science data if this effort was more optimally spread throughout the landscape—by directing effort to align with underlying species diversity. Of course, it is important to acknowledge that to effectively direct citizen science participants, a social science approach will be necessary to devise the best strategies to interact and engage with the participants of a project (Pocock et al., 2019). Nevertheless, some participants’ motivations involve contributing to conservation, suggesting that a subset of users would be willing to alter their sampling efforts (Tiago, Gouveia, et al., 2017) to inform a specific conservation-related goal.

Most conservation planning takes place at relatively small spatial scales, and citizen science is often seen as a coarse source of information on species occurrences. Beyond our general patterns of citizen science effort needed to quantify species diversity, we found significant

differences among grain sizes and depending on whether species diversity estimation was weighted toward dominant or rare species. Unsurprisingly, we found a general increase in the mean number of samples needed to estimate species diversity with increasing grain size, but this relationship was much stronger when species diversity was focused on rare species (Figure 3b). For both dominant and rare diversity measures, the necessary sampling effort grows slower than the area of sampling, following a power law with an exponent of around 0.3–0.4. This may be a consequence of the species–area relationship, $S \sim A^z$, where S is the number of species in a cell, A is the area of the cell, and z is an exponent (Rosenzweig et al., 1995), with the species–area relationship being steeper for rare species than for common species (Sanderlin et al., 2014). This supports an approach of planning sampling effort of fragments or land-cover units based on the species–area relationship (Borges et al., 2009; Proença et al., 2010). In general, our workflow performed better for diversity estimates that focused on dominant species, with better model fit for random forests fit to the common species sensitive compared to the rare species sensitive metric (Appendix S1: Table S2). This could be a result of the increased confidence in documenting the common species with eBird data, but also a result of participants who are inclined to look for new species (e.g., vagrants), such that species accumulation continues to increase indefinitely at popular birding locations due to continued search effort by birders (see the Patagonia Picnic Table Effect; https://en.wikipedia.org/wiki/Patagonia_picnic_table_effect). Or a biological explanation could be that diversity patterns of rare species are intrinsically less predictable than those of dominant species (Lennon et al., 2011), and thus, our model fits are poorer for diversity estimates most sensitive to rare species.

While our case study focused on Florida, the approach could be generalized to other parts of the world, given the global increase in citizen science data (Pocock et al., 2017). We focused on an annual temporal resolution, but smaller temporal windows (e.g., seasonally or monthly) could similarly be implemented to quantify effort needed to estimate species diversity (Figure 4). In the case of Florida, for example, the effort to estimate species richness would be highest during spring and fall migration when bird diversity is highest in the region. The temporal window will ultimately depend on the quantity of data in a given locale. An important component of our framework is relying on the strong relationships we found between land-cover metrics and species richness (Figures 1 and 2), and our SEM model suggested that variables with good predictive power of sampling effort are likely to be strongly linked to species diversity. We presented the results of the two extremes ($q = 0$ and $q = 2$) of the sampling completeness curve, but

this curve is continuous (Chao et al., 2020), providing an interested user the opportunity to fine-tune exactly how much preference they want to weight toward rare or dominant species. We found little difference between $q = 1$ and $q = 2$ (Appendix S1: Figures S6 and S7, Table S1), but this could change when abundance-based calculations are used as opposed to incidence-based calculation—all of which is adaptable from our proposed workflow. Other studies and implementation of this approach could tailor the predictors, such as land use or other land-cover attributes or environmental covariates, to their specific context. Additionally, we only focused on one measure of sampling completeness, and other measures of survey completeness (e.g., Lobo et al., 2018) could be used in implementation of this framework.

Citizen science data will continue to play an important role in biodiversity monitoring in future (Chandler et al., 2017; McKinley et al., 2017). Despite their promise, there remains reluctance to use these data (Burgess et al., 2017), in large part stemming from gaps and redundancies. To increase utility of these data, a key goal is to understand how biodiversity sampling should be conducted in space and time. We find relatively few samples are necessary to meet 95% completeness, thus allowing for relatively robust comparisons of species diversity across the landscape. Hence, our results highlight the potential of citizen science data to make informed comparisons of species diversity in space and/or time. However, the sampling effort inherently depends on the monitoring goal, for example, whether all species or only more common species are targeted. The generalizable workflow presented here allows for the quantification of sampling effort needed to estimate species diversity with citizen science data and shows how citizen science sampling effort might be targeted toward better estimates of biodiversity.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Code and data (Callaghan, 2021) are available from Zenodo: <https://doi.org/10.5281/zenodo.5734200>.

ORCID

Corey T. Callaghan  <https://orcid.org/0000-0003-0415-2709>

Shane A. Blowes  <https://orcid.org/0000-0001-6310-3670>

Jonathan M. Chase  <https://orcid.org/0000-0001-5580-4303>

Mitchell B. Lyons  <https://orcid.org/0000-0003-3960-3522>

Henrique M. Pereira  <https://orcid.org/0000-0003-1043-1675>

REFERENCES

- Bakker, V. J., J. K. Baum, J. F. Brodie, A. K. Salomon, B. G. Dickson, H. K. Gibbs, O. P. Jensen, and P. B. McIntyre. 2010. "The Changing Landscape of Conservation Science Funding in the United States." *Conservation Letters* 3: 435–44.
- Bibby, C. J. 2004. "Bird Diversity Survey Methods." In *Bird Ecology and Conservation: A Handbook of Techniques*, Vol 11. Oxford, UK: Oxford Scholarship.
- Boakes, E. H., P. J. McGowan, R. A. Fuller, D. Chang-qing, N. E. Clark, K. O'Connor, and G. M. Mace. 2010. "Distorted Views of Biodiversity: Spatial and Temporal Bias in Species Occurrence Data." *PLoS Biology* 8: e1000385.
- Bonney, R., C. B. Cooper, J. Dickinson, S. Kelling, T. Phillips, K. V. Rosenberg, and J. Shirk. 2009. "Citizen Science: A Developing Tool for Expanding Science Knowledge and Scientific Literacy." *Bioscience* 59: 977–84.
- Borges, P. A., J. Hortal, R. Gabriel, and N. Homem. 2009. "Would Species Richness Estimators Change the Observed Species Area Relationship?" *Acta Oecologica* 35: 149–56.
- Boyd, C., T. M. Brooks, S. H. Butchart, G. J. Edgar, G. A. Da Fonseca, F. Hawkins, M. Hoffmann, W. Sechrest, S. N. Stuart, and P. P. Van Dijk. 2008. "Spatial Scale and the Conservation of Threatened Species." *Conservation Letters* 1: 37–43.
- Brlík, V., E. Šilarová, J. Škorpilová, H. Alonso, M. Anton, A. Aunins, Z. Benkő, et al. 2021. "Long-Term and Large-Scale Multispecies Dataset Tracking Population Changes of Common European Breeding Birds." *Scientific Data* 8: 1–9.
- Buchhorn, M., M. Lesiv, N.-E. Tsendbazar, M. Herold, L. Bertels, and B. Smets. 2020. "Copernicus Global Land Cover Layers—Collection 2." *Remote Sensing* 12: 1044.
- Burgess, H. K., L. DeBey, H. Froehlich, N. Schmidt, E. J. Theobald, A. K. Ettinger, J. HilleRisLambers, J. Tewksbury, and J. K. Parrish. 2017. "The Science of Citizen Science: Exploring Barriers to Use as a Primary Research Tool." *Biological Conservation* 208: 113–20.
- Butchart, S. H., M. Walpole, B. Collen, A. Van Strien, J. P. Scharlemann, R. E. Almond, J. E. Baillie, et al. 2010. "Global Biodiversity: Indicators of Recent Declines." *Science* 328: 1164–8.
- Callaghan, C. 2021. "coreycallaghan/ECS21-0486." Version v1. Data and Code. Zenodo. <https://doi.org/10.5281/zenodo.5734200>
- Callaghan, C. T., A. G. Poore, R. E. Major, J. J. Rowley, and W. K. Cornwell. 2019. "Optimizing Future Biodiversity Sampling by Citizen Scientists." *Proceedings of the Royal Society B: Biological Sciences* 286: 20191487.
- Callaghan, C. T., J. J. Rowley, W. K. Cornwell, A. G. Poore, and R. E. Major. 2019. "Improving Big Citizen Science Data: Moving beyond Haphazard Sampling." *PLoS Biology* 17: e3000357.
- Callaghan, C. T., J. E. Watson, M. B. Lyons, W. K. Cornwell, and R. A. Fuller. 2021. "Conservation Birding: A Quantitative Conceptual Framework for Prioritizing Citizen Science Observations." *Biological Conservation* 253: 108912.
- Chandler, M., L. See, K. Copas, A. M. Bonde, B. C. López, F. Danielsen, J. K. Legind, et al. 2017. "Contribution of Citizen Science towards International Biodiversity Monitoring." *Biological Conservation* 213: 280–94.
- Chao, A., R. K. Colwell, C.-W. Lin, and N. J. Gotelli. 2009. "Sufficient Sampling for Asymptotic Minimum Species Richness Estimators." *Ecology* 90: 1125–33.

- Chao, A., N. J. Gotelli, T. Hsieh, E. L. Sander, K. Ma, R. K. Colwell, and A. M. Ellison. 2014. "Rarefaction and Extrapolation with Hill Numbers: A Framework for Sampling and Estimation in Species Diversity Studies." *Ecological Monographs* 84: 45–67.
- Chao, A., and L. Jost. 2012. "Coverage-Based Rarefaction and Extrapolation: Standardizing Samples by Completeness Rather than Size." *Ecology* 93: 2533–47.
- Chao, A., Y. Kubota, D. Zelený, C.-H. Chiu, C.-F. Li, B. Kusumoto, M. Yasuhara, et al. 2020. "Quantifying Sample Completeness and Comparing Diversities among Assemblages." *Ecological Research* 35: 292–314.
- Chase, J. M., B. J. McGill, P. L. Thompson, L. H. Antão, A. E. Bates, S. A. Blowes, M. Dornelas, et al. 2019. "Species Richness Change across Spatial Scales." *Oikos* 128: 1079–91.
- Colwell, R. K., and J. A. Coddington. 1994. "Estimating Terrestrial Biodiversity through Extrapolation." *Philosophical Transactions of the Royal Society B: Biological Sciences* 345: 101–18.
- Cornwell, W. K., W. D. Pearse, R. L. Dalrymple, and A. E. Zanne. 2019. "What We (Don't) Know about Global Plant Diversity." *Ecography* 42: 1819–31.
- Cosentino, B. J., D. M. Marsh, K. S. Jones, J. J. Apodaca, C. Bates, J. Beach, K. H. Beard, et al. 2014. "Citizen Science Reveals Widespread Negative Effects of Roads on Amphibian Distributions." *Biological Conservation* 180: 31–8.
- Courter, J. R., R. J. Johnson, C. M. Stuyck, B. A. Lang, and E. W. Kaiser. 2013. "Weekend Bias in Citizen Science Data Reporting: Implications for Phenology Studies." *International Journal of Biometeorology* 57: 715–20.
- Essl, F., W. Rabitsch, S. Dullinger, D. Moser, and N. Milasowszky. 2013. "How Well Do We Know Species Richness in a Well-Known Continent? Temporal Patterns of Endemic and Widespread Species Descriptions in the European Fauna." *Global Ecology and Biogeography* 22: 29–39.
- Fleishman, E., R. F. Noss, and B. R. Noon. 2006. "Utility and Limitations of Species Richness Metrics for Conservation Planning." *Ecological Indicators* 6: 543–53.
- Freiwald, J., R. Meyer, J. E. Caselle, C. A. Blanchette, K. Hovel, D. Neilson, J. Dugan, J. Altstatt, K. Nielsen, and J. Bursek. 2018. "Citizen Science Monitoring of Marine Protected Areas: Case Studies and Recommendations for Integration into Monitoring Programs." *Marine Ecology* 39: e12470.
- Gilfedder, M., C. J. Robinson, J. E. Watson, T. G. Campbell, B. L. Sullivan, and H. P. Possingham. 2019. "Brokering Trust in Citizen Science." *Society & Natural Resources* 32: 292–302.
- Gorelick, N., M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore. 2017. "Google Earth Engine: Planetary-Scale Geospatial Analysis for Everyone." *Remote Sensing of Environment* 202: 18–27.
- Gotelli, N. J., and R. K. Colwell. 2001. "Quantifying Biodiversity: Procedures and Pitfalls in the Measurement and Comparison of Species Richness." *Ecology Letters* 4: 379–91.
- Heberling, J. M., J. T. Miller, D. Noesgaard, S. B. Weingart, and D. Schigel. 2021. "Data Integration Enables Global Biodiversity Synthesis." *Proceedings of the National Academy of Sciences* 118(6): e2018093118.
- Hillebrand, H., B. Blasius, E. T. Borer, J. M. Chase, J. A. Downing, B. K. Eriksson, C. T. Filstrup, et al. 2018. "Biodiversity Change Is Uncoupled from Species Richness Trends: Consequences for Conservation and Monitoring." *Journal of Applied Ecology* 55: 169–84.
- Hsieh, T., K. Ma, and A. Chao. 2016. "iNEXT: An r Package for Rarefaction and Extrapolation of Species Diversity (Hill Numbers)." *Methods in Ecology and Evolution* 7: 1451–6.
- Hugo, S., and R. Altwegg. 2017. "The Second Southern African Bird Atlas Project: Causes and Consequences of Geographical Sampling Bias." *Ecology and Evolution* 7: 6839–49.
- Husby, M., K. S. Hoset, and S. Butler. 2021. "Non-Random Sampling Along Rural–Urban Gradients May Reduce Reliability of Multi-Species Farmland Bird Indicators and Their Trends." *Ibis* 163: 579–92.
- IPBES. 2019. In *Global Assessment Report on Biodiversity and Ecosystem Services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services*, edited by E.S. Brondizio, J. Settele, S. Díaz, and H.T. Ngo. Bonn: IPBES Secretariat.
- Isaac, N. J., and M. J. Pocock. 2015. "Bias and Information in Biological Records." *Biological Journal of the Linnean Society* 115: 522–31.
- Kelling, S., D. Fink, F. A. La Sorte, A. Johnston, N. E. Bruns, and W. M. Hochachka. 2015. "Taking a 'Big Data' Approach to Data Quality in a Citizen Science Project." *Ambio* 44: 601–11.
- Kelling, S., A. Johnston, A. Bonn, D. Fink, V. Ruiz-Gutierrez, R. Bonney, M. Fernandez, et al. 2019. "Using Semistructured Surveys to Improve Citizen Science Data for Monitoring Biodiversity." *Bioscience* 69: 170–9.
- Kirchhoff, C., C. T. Callaghan, D. A. Keith, D. Indiarto, G. Taseski, M. K. Ooi, T. D. Le Breton, T. Mesaglio, R. T. Kingsford, and W. K. Cornwell. 2021. "Rapidly Mapping Fire Effects on Biodiversity at a Large-Scale Using Citizen Science." *Science of the Total Environment* 755: 142348.
- Lefcheck, J. S. 2016. "piecewiseSEM: Piecewise Structural Equation Modelling in r for Ecology, Evolution, and Systematics." *Methods in Ecology and Evolution* 7: 573–9.
- Lennon, J. J., C. M. Beale, C. L. Reid, M. Kent, and R. J. Pakeman. 2011. "Are Richness Patterns of Common and Rare Species Equally Well Explained by Environmental Variables?" *Ecography* 34: 529–39.
- Liaw, A., and M. Wiener. 2002. "Classification and Regression by randomForest." *R News* 2: 18–22.
- Lobo, J. M., J. Hortal, J. L. Yela, A. Millán, D. Sánchez-Fernández, E. García-Roselló, J. González-Dacosta, J. Heine, L. González-Vilas, and C. Guisande. 2018. "KnowBR: An Application to Map the Geographical Variation of Survey Effort and Identify Well-Surveyed Areas from Biodiversity Databases." *Ecological Indicators* 91: 241–8.
- Magurran, A. E., and B. J. McGill. 2011. *Biological Diversity: Frontiers in Measurement and Assessment*. Oxford, UK: Oxford University Press.
- McGlenn, D. J., X. Xiao, F. May, N. J. Gotelli, T. Engel, S. A. Blowes, T. M. Knight, O. Purschke, J. M. Chase, and B. J. McGill. 2019. "Measurement of Biodiversity (MoB): A Method to Separate the Scale-Dependent Effects of Species Abundance Distribution, Density, and Aggregation on Diversity Change." *Methods in Ecology and Evolution* 10: 258–69.
- McKinley, D. C., A. J. Miller-Rushing, H. L. Ballard, R. Bonney, H. Brown, S. C. Cook-Patton, D. M. Evans, et al. 2017. "Citizen

- Science Can Improve Conservation Science, Natural Resource Management, and Environmental Protection.” *Biological Conservation* 208: 15–28.
- Pacifici, K., B. J. Reich, R. M. Dorazio, and M. J. Conroy. 2016. “Occupancy Estimation for Rare Species Using a Spatially-Adaptive Sampling Design.” *Methods in Ecology and Evolution* 7: 285–93.
- Pavlacky, D. C., Jr., P. M. Lukacs, J. A. Blakesley, R. C. Skorkowsky, D. S. Klute, B. A. Hahn, V. J. Dreitz, T. L. George, and D. J. Hanni. 2017. “A Statistically Rigorous Sampling Design to Integrate Avian Monitoring and Management within Bird Conservation Regions.” *PLoS One* 12: e0185924.
- Pereira, H. M., J. Belnap, M. Böhm, N. Brummitt, J. Garcia-Moreno, R. Gregory, L. Martin, et al. 2017. “Monitoring Essential Biodiversity Variables at the Species Level.” In *The GEO Handbook on Biodiversity Observation Networks* 79–105. Cham: Springer.
- Pocock, M. J., H. E. Roy, T. August, A. Kuria, F. Barasa, J. Bett, M. Githiru, et al. 2019. “Developing the Global Potential of Citizen Science: Assessing Opportunities that Benefit People, Society and the Environment in East Africa.” *Journal of Applied Ecology* 56: 274–81.
- Pocock, M. J., J. C. Tweddle, J. Savage, L. D. Robinson, and H. E. Roy. 2017. “The Diversity and Evolution of Ecological and Environmental Citizen Science.” *PLoS One* 12: e0172579.
- Proença, V. M., H. M. Pereira, J. Guilherme, and L. Vicente. 2010. “Plant and Bird Diversity in Natural Forests and in Native and Exotic Plantations in NW Portugal.” *Acta Oecologica* 36: 219–26.
- R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. <https://www.R-project.org/>
- Roman, L. A., B. C. Scharenbroch, J. P. Östberg, L. S. Mueller, J. G. Henning, A. K. Koeser, J. R. Sanders, D. R. Betz, and R. C. Jordan. 2017. “Data Quality in Citizen Science Urban Tree Inventories.” *Urban Forestry & Urban Greening* 22: 124–35.
- Rosenzweig, M. L., et al. 1995. *Species Diversity in Space and Time*. Cambridge, UK: Cambridge University Press.
- Sanderlin, J. S., W. M. Block, and J. L. Ganey. 2014. “Optimizing Study Design for Multi-Species Avian Monitoring Programmes.” *Journal of Applied Ecology* 51: 860–70.
- Scheffers, B. R., L. N. Joppa, S. L. Pimm, and W. F. Laurance. 2012. “What We Know and Don’t Know about Earth’s Missing Biodiversity.” *Trends in Ecology & Evolution* 27: 501–10.
- Schuster, R., S. Wilson, A. D. Rodewald, P. Arcese, D. Fink, T. Auer, and J. R. Bennett. 2019. “Optimizing the Conservation of Migratory Species over Their Full Annual Cycle.” *Nature Communications* 10: 1–8.
- Sgarbi, L. F., L. M. Bini, J. Heino, J. Jyrkänkallio-Mikkola, V. L. Landeiro, E. P. Santos, F. Schneck, et al. 2020. “Sampling Effort and Information Quality Provided by Rare and Common Species in Estimating Assemblage Structure.” *Ecological Indicators* 110: 105937.
- Specht, H. M., H. T. Reich, F. Iannarilli, M. R. Edwards, S. P. Stapleton, M. D. Weegman, M. K. Johnson, B. J. Yohannes, and T. W. Arnold. 2017. “Occupancy Surveys with Conditional Replicates: An Alternative Sampling Design for Rare Species.” *Methods in Ecology and Evolution* 8: 1725–34.
- Stork, N. E. 1993. “How Many Species Are There?” *Biodiversity and Conservation* 2: 215–32.
- Sullivan, B. L., J. L. Aycrigg, J. H. Barry, R. E. Bonney, N. Bruns, C. B. Cooper, T. Damoulas, et al. 2014. “The eBird Enterprise: An Integrated Approach to Development and Application of Citizen Science.” *Biological Conservation* 169: 31–40.
- Sullivan, B. L., C. L. Wood, M. J. Iliff, R. E. Bonney, D. Fink, and S. Kelling. 2009. “eBird: A Citizen-Based Bird Observation Network in the Biological Sciences.” *Biological Conservation* 142: 2282–92.
- Tiago, P., A. Ceia-Hasse, T. A. Marques, C. Capinha, and H. M. Pereira. 2017. “Spatial Distribution of Citizen Science Casuistic Observations for Different Taxonomic Groups.” *Scientific Reports* 7: 1–9.
- Tiago, P., M. J. Gouveia, C. Capinha, M. Santos-Reis, and H. M. Pereira. 2017. “The Influence of Motivational Factors on the Frequency of Participation in Citizen Science Activities.” *Nature Conservation* 18: 61.
- Tulloch, A. I. T., K. Mustin, H. P. Possingham, J. K. Szabo, and K. A. Wilson. 2013. “To Boldly Go Where No Volunteer Has Gone Before: Predicting Volunteer Activity to Prioritize Surveys at the Landscape Scale.” *Diversity and Distributions* 19: 465–80.
- Van der Wal, R., H. Anderson, A. Robinson, N. Sharma, C. Mellish, S. Roberts, B. Darvill, and A. Siddharthan. 2015. “Mapping Species Distributions: A Comparison of Skilled Naturalist and Lay Citizen Science Recording.” *Ambio* 44: 584–600.
- Van Wilgenburg, S. L., C. L. Mahon, G. Campbell, L. McLeod, M. Campbell, D. Evans, W. Easton, et al. 2020. “A Cost Efficient Spatially Balanced Hierarchical Sampling Design for Monitoring Boreal Birds Incorporating Access Costs and Habitat Stratification.” *PLoS One* 15: e0234494.
- Vardi, R., O. Berger-Tal, and U. Roll. 2021. “iNaturalist Insights Illuminate COVID-19 Effects on Large Mammals in Urban Centers.” *Biological Conservation* 254: 108953.
- Welvaert, M., and P. Caley. 2016. “Citizen Surveillance for Environmental Monitoring: Combining the Efforts of Citizen Science and Crowdsourcing in a Quantitative Data Framework.” *Springerplus* 5: 1–14.
- Wickham, H., M. Averick, J. Bryan, W. Chang, L. D. McGowan, R. François, G. Grolemund, et al. 2019. “Welcome to the tidyverse.” *Journal of Open Source Software* 4: 1686.
- Xu, H., Y. Cao, D. Yu, M. Cao, Y. He, M. Gill, and H. M. Pereira. 2021. “Ensuring Effective Implementation of the Post-2020 Global Biodiversity Targets.” *Nature Ecology and Evolution* 5: 411–8.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher’s website.

How to cite this article: Callaghan, Corey T., Diana E. Bowler, Shane A. Blowes, Jonathan M. Chase, Mitchell B. Lyons, and Henrique M. Pereira. 2022. “Quantifying Effort Needed to Estimate Species Diversity from Citizen Science Data.” *Ecosphere* 13(4): e3966. <https://doi.org/10.1002/ecs2.3966>