Three Frontiers for the Future of Biodiversity Research Using Citizen Science Data

COREY T. CALLAGHAN, ALISTAIR G. B. POORE, THOMAS MESAGLIO, ANGELA T. MOLES, SHINICHI NAKAGAWA, CHRISTOPHER ROBERTS, JODI J. L. ROWLEY, ADRIANA VERGÉS, JOHN H. WILSHIRE, AND WILLIAM K. CORNWELL

Citizen science is fundamentally shifting the future of biodiversity research. But although citizen science observations are contributing an increasingly large proportion of biodiversity data, they only feature in a relatively small percentage of research papers on biodiversity. We provide our perspective on three frontiers of citizen science research, areas that we feel to date have had minimal scientific exploration but that we believe deserve greater attention as they present substantial opportunities for the future of biodiversity research: sampling the undersampled, capitalizing on citizen science's unique ability to sample poorly sampled taxa and regions of the world, reducing taxonomic and spatial biases in global biodiversity data sets; estimating abundance and density in space and time, develop techniques to derive taxon-specific densities from presence or absence and presence-only data; and capitalizing on secondary data collection, moving beyond data on the occurrence of single species and gain further understanding of ecological interactions among species or habitats. The contribution of citizen science to understanding the important biodiversity questions of our time should be more fully realized.

Keywords: citizen science, biodiversity, abundance, opportunities in citizen science, GBIF

Citizen science, or community science, is a rapidly advancing field, with an ever-increasing number of projects (Jordan et al. 2015, Welvaert and Caley 2016, Pocock et al. 2017). Many of these projects are focused on biodiversity, generally aiming to put points on a map for a given taxon (Pocock et al. 2017). We are also in the middle of a big data revolution in ecology and conservation (Farley et al. 2018) with increasingly available remote-sensing data (Kwok 2018) and trait databases (Schneider et al. 2019). For example, citizen science is largely responsible for the Global Biodiversity Information Facility (GBIF) having accumulated approximately 1.4 billion biodiversity records globally (Chandler et al. 2017a). As of March 2020, data from GBIF has been used in 4307 research papers. Collectively, these data are expanding the spatial and temporal scale of questions that can be answered in ecology, conservation, and natural resource management (McKinley et al. 2017).

Despite the opportunities, there are still obstacles blocking widespread use of these data within both academic research and development of government policies (Burgess et al. 2017, Troudet et al. 2017, Young et al. 2019). Chief among these, unsurprisingly, are questions surrounding data quality, which have been discussed in depth elsewhere (Kosmala et al. 2016, Aceves-Bueno et al. 2017). In the present article, we do not focus on data quality, but rather focus on a series of opportunities that make use of the particular qualities of citizen science data. These opportunities could allow professional scientists to build tools both to better direct the incredible amount of citizen science effort and to better use the rapidly accumulating data sets in biodiversity research (Tulloch et al. 2013b).

There are many avenues for increasing the utility of citizen science research (Newman et al. 2012, Bonney et al. 2014). In the present article, we focus on three frontiers that we believe present substantial opportunities for progress to advance the field of citizen science and, in particular, to answer fundamental questions important to understanding and conserving biodiversity: (1) using citizen science to increase representation of undersampled regions and taxa, (2) developing pipelines to estimate species' abundance in space and time, and (3) capitalizing on secondary data collection (i.e., data held in the user contributions but not part of the initial aim of the contribution). These perspectives primarily apply to semistructured and unstructured citizen science projects (e.g., iNaturalist, eBird, FrogID, iSpot) that are largely opportunistic in nature (Shirk et al. 2012, Danielsen et al. 2014). We treat each of these perspectives in turn, by summarizing the current state of literature and

BioScience 71: 55–63.
doi:10.1093/biosci/biaa131 Advance Access publication 10 November 2020

providing illustrative examples of how these are being tackled. We conclude by providing a series of key objectives and scientific questions for the future of citizen science within each of these frontiers.

Sampling the undersampled

Most citizen science data—across and within projects—have redundancies and gaps in taxonomic focus, space, and time (Boakes et al. 2010, Bayraktarov et al. 2019). Understanding these limitations is extremely important for appropriate use of citizen science data (Burgess et al. 2017). However, many of the redundancies and gaps in citizen science data sets (i.e., taxonomic, spatial, and temporal biases) also exist in professional science data sets (Boakes et al. 2010). Professional science is often highly constrained by funding, logistics, and time, leading to incomplete biodiversity data sets. In some instances, it is possible that citizen science may do a better job of sampling at least some parts of the world's biodiversity. We suggest that the future of citizen science research should further capitalize on three areas in which citizen science is likely to have invaluable contributions to understanding biodiversity and informing conservation: sampling biodiversity on private land, advancing biodiversity understanding in developing countries and remote areas, and sampling underrepresented taxa.

Biodiversity on private lands may differ from that on public lands (Scott et al. 2001) and most endangered species rely, in part, on habitats on private land—many entirely so (Bean and Wilcove 1997). Given the large proportion of land that is privately owned (e.g., approximately 60% of the United States, Hilty and Merenlender 2003; 63% of Australia, ABS 2002; and higher in other parts of the world, Scott et al. 2001), sampling biodiversity on private lands is essential to monitor trends in species' distributions and population status. Professional science is not effective at sampling private lands (Hilty and Merenlender 2003), because gaining access is often time consuming and difficult. Citizen science, however, is uniquely positioned to sample biodiversity on private lands—leveraging public citizens and community members to collect large amounts of data, including from their own backyards. As an example, FrogID—a national citizen science project focused on recording frogs in Australia (Rowley et al. 2019)—has 92% of their records from private lands (from 2017–2019). Citizen science targeting private land is most likely to benefit from backyard contributions, often in residential areas (e.g., Cooper et al. 2007). For example, project FeederWatch focuses on backyard birdwatching in the United States, using semiautomated filters to help both participants and researchers have confidence in the data being collected (Bonter and Cooper 2012). Whereas Gardenwatch, ran by the British Trust for Ornithology in the United Kingdom, focuses on different missions for individuals to submit data on birds, invertebrates, and mammals in their backyards (BTO 2020). However, for larger tracts of land (e.g., agriculture, resource extraction), citizen scientists are likely to face similar access constraints as professional biodiversity monitoring.

Similar biases for citizen science and professional science also apply at the global scale (Yesson et al. 2007, Boakes et al. 2010), because there are many remote or isolated areas that are sparsely populated and rarely visited by either traditional or citizen scientists. Many parts of the world simply do not have the economic resources to fund a scientific establishment. As a result, global scientific databases often have regions of severe data paucity (figure 1). For example, global plant trait data sets have sparse information for Siberia, Greenland, northern Canada, arid Australia, parts of Saharan and Central Africa, and much of the Amazon (Kattge et al. 2011). Similarly, GBIF has sparse information for Russia, Greenland, northern Canada, Antarctica, parts of Saharan and Central Africa, and for much of the world's oceans. Undersampled areas that are highly diverse, highly endemic, poorly known, or contain highly threatened species or habitats should remain a priority for professional scientists (Tulloch et al. 2013a, Bayraktarov et al. 2019). But in combination with professional science, citizen science can help these developing and remote regions to quantify their biodiversity without necessarily building traditional research institutions (e.g., field stations, museums, and universities) *in situ*. This is parallel to the way that development of mobile phone networks allowed large parts of the world to move from no telecommunications to high connectivity without the establishment of a traditional landline infrastructure (Andrachuk et al. 2019). For example, through a collaborative citizen science project in remote Northern Territory, Australia, the Ngukurr community helped to build knowledge about the local biodiversity, including discovering new species, identifying populations of threatened species, and documenting culturally significant habitats (https://youtu.be/EAnVoA1PB5k). The Custodians of Rare and Endangered Wildflowers (SANBI 2020) program supports citizen scientists working in remote parts of South Africa to survey wildflowers.

In addition to providing local people with the opportunity to document their biodiversity, citizen science is well positioned to make use of records from holidaying biodiversity enthusiasts (Mieras et al. 2017). Ecotourism—especially to remote parts of the world—is a growing industry (Das and Chatterjee 2015) that has the potential to be combined with citizen science data collection. For example, ecotourists have helped monitor cetaceans in Hawaii by photographing cetaceans while on whale-watching tours (Currie et al. 2018), and the CoralWatch citizen science project (https://coralwatch.org) has successfully recruited ecotourists to participate in coral reef surveys, with some in relatively remote regions in the world such as Indonesia (Marshall et al. 2012). This ecotourism, coupled with artificial intelligence, can now analyze millions of social media posts or online photo repositories to glean information about biodiversity (e.g., wildme.org; Menon et al. 2016). Participation with local tourism authorities and managers (e.g., lodge owners, tour guide operators)

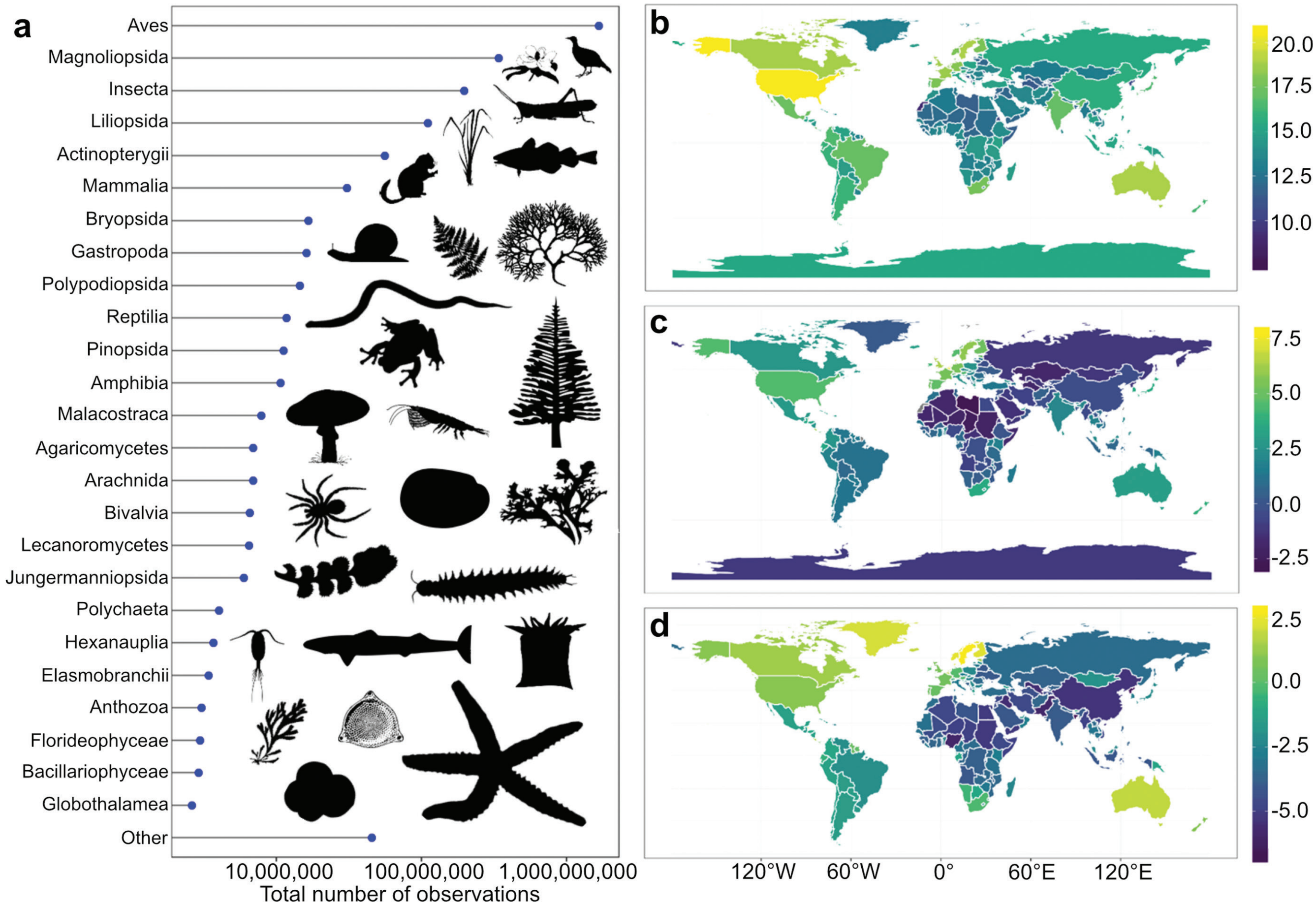

Figure 1. (a) The total number of records for the top 25 classes present in GBIF, demonstrating the potential for citizen science to capitalize on those undersampled taxa. (b) The total number of records in GBIF, by country, on a log scale, showing the bias toward well-sampled areas (e.g., United States, Europe, Australia). (c) The total number of records in GBIF, standardized by the area of the country, again demonstrating those parts of the world that could most benefit from increased biodiversity knowledge gained through citizen science. (d) The total number of records in GBIF per capita, showing that parts of Europe have the most observations per capita in the world.

in these undersampled parts of the world could provide benefits for opportunistic citizen science observations. Furthermore, some companies (e.g., adventurescientists.org) are now targeting dedicated volunteers who are willing to go to remote parts of the world to collect data for specific scientific projects—a unique form of ecotourism blended with citizen science.

Another mechanism by which undersampled regions can be boosted is through global citizen science projects—those that target observations from anywhere in the world, such as iNaturalist (www.inaturalist.org). iNaturalist, for example, collates photos of any living organism in the world, allowing for community validation of these photos. Although not formally quantified, there is likely an increase in citizen science observations in these remote and developing parts of the world—contributed both by local naturalists and by ecotourists (Pocock et al. 2019). We suggest that experts should optimize their time spent identifying opportunistic observations—for example, in iNaturalist (Tulloch et al. 2013a, Callaghan et al. 2019a)—by prioritizing verification of records from poorly sampled regions (e.g., the tropics, pacific islands) rather than only verifying additional records from well-sampled regions (e.g., United States, Europe; Orr et al. 2020).

Taxonomic bias is an inherent feature of organismal research, including biodiversity records, with the representation of taxa in the literature and in biodiversity databases failing to reflect their representation in nature (May 1988, Bonnet et al. 2002, Troudet et al. 2017). In general, invertebrates tend to be underrepresented in biodiversity databases, and within vertebrates, birds and mammals tend to be overrepresented (May 1988, Troudet et al. 2017). Some reasons for such taxonomic bias are obvious—some organisms are more difficult to study than others because they are difficult to locate or identify. We argue that citizen science projects are well suited to minimize some of these biases. First, identification of any given taxa is no longer dependent on local experts and can be globally exported through platforms such as iNaturalist (Orr et al. 2020). For example, there should be increased effort in identifying photographs to a species level by both professional scientists and other experts for given taxa that are not often identified to species (e.g., only 32% of polychaetes on iNaturalist are identified to species; figure 1). Second, societal preferences greatly influence taxonomic biases (Czech et al. 1998, Troudet et al. 2017) but professional scientists can work through citizen science initiatives to help minimize and overcome these biases in our biodiversity knowledge. Citizen science has been used to map species distributions of saproxylic beetles (Zapponi et al. 2017) and to identify introduced species of Hymenoptera in New Zealand (Ward 2014), both examples of uncharismatic invertebrates. Another example is from Jones and colleagues (2019), who highlight that “iNaturalist was instrumental in facilitating the discovery” of a rare crayfish that was able to successfully be given conservation status, and they state that “had it not been for iNaturalist, its presence may have remained undetected.”

One potential avenue for minimizing taxonomic, and other, biases in citizen science projects is the gamification of citizen science—the process by which participants are rewarded for their sightings in a game-like fashion (e.g., by receiving badges). Gamification may lead to increased retention of current participants, but also recruit new participants to a particular citizen science project (Bowser et al. 2013, Chandler et al. 2017b). As an example, to minimize taxonomic biases inherent in citizen science projects, participants could be encouraged to find or identify underrepresented taxa (e.g., invertebrates).

Estimating species’ abundances in space and time

A key benefit of massive citizen science data sets is the ability to monitor biodiversity in space and time at a frequency and geographic extent that has not been possible before (Schmeller et al. 2009, Tulloch et al. 2013b, Chandler et al. 2017a). This is key for both detecting range expansions and contractions and understanding the many ways that individual species are responding to the changing world. Because of conservation implications and basic research importance, understanding biodiversity in space and time has received a tremendous amount of research interest both related (Chandler et al. 2017a) and not related (Gotelli and Colwell 2001) to citizen science. Such understanding has traditionally been based on traditional data sets (e.g., museum collections, intensive survey data), however, the increasingly dense sampling of citizen science data sets in both space and time offer both new opportunities and new statistical challenges.

A key future prospect is to estimate organism abundance—with associated uncertainty—in space and time. For semistructured projects that provide complete snapshots of the biodiversity encountered on a survey (e.g., eBird, www.ebird.org; Reef Life Survey, www.reeflifesurvey.com), it is straightforward to model abundance of one species at one point in time or space relative to itself at another point in time or space, and indeed, this has already been done for many well-sampled North American bird species (e.g., https://ebird.org/science/status- and-trends; Fink et al. 2010). The key information to estimate relative abundance is the existence of true absences in the data set; absences allow modeling of when a species both was and was not encountered in space and time. In this class of data, absences are inferred from complete checklists, in which observers submit lists of all species they were able to identify along with a proxy for effort, allowing for modeling of the probability of presence or absence (e.g., Johnston et al. 2020). But these citizen science data have variations in observer skill and effort, as well as observer bias in when and where to sample—problems often true for professionally collected scientific data too. There are already statistical approaches to minimize these biases, such as hierarchical modeling or spatial and temporal subsampling (Gonsamo and D’Odorico 2014, Johnston et al. 2020), and these are continuously being improved.

In contrast, modeling abundance is more difficult if starting from presence-only data, such as those traditionally generated by museum or herbarium records, and more recently many opportunistic citizen science projects (e.g., iNaturalist, FrogID, iSpot, and many others). The lack of absences in these data requires additional analysis steps, and methods to use this class of data more fully are being rapidly developed (Fithian et al. 2015, Meyer et al. 2015, Roberts et al. 2017). The most powerful of these new approaches is informing the inference from presence-only data sets with high-quality information from another source including plot, distance sampling, or remote-sensing data (He et al. 2015). In general, this approach works by statistically combining multiple data sources with different characteristics, such as low quality presence-only citizen science data in combination with high quality professional survey data (Pacifici et al. 2017). For example, Fithian and colleagues (2015) showed that by pooling presence only and presence or absence data together in a complex statistical model, many of the biases in the presence-only data can be minimized. In another example, Pacifici and colleagues (2017) showed that models of brown-headed nuthatch distributions are improved when incorporating both citizen science data in addition to structured survey data into a synthetic understanding of the species’ range. This promising area of exchange between professional field ecologists, citizen scientists, and statisticians shows how professional scientists could maximize the impact of their limited time in the field by generating data sets specifically designed to unlock aspects of the massive potential of citizen science data. This will require both a full understanding of the statistical approaches used to integrate data (Fithian et al. 2015, Pacifici et al. 2017) and forward-looking statistical models that can dynamically predict where the most valuable data should come from for increased confidence around specific scientific objectives (e.g., Callaghan et al. 2019a, 2019b). If models can be continuously updated with data from both citizen scientists and professionals, then data gaps in space and time can be identified and filled (Callaghan et al. 2019b).

Citizen science observations are continuously and globally contributed, in a near real-time fashion. For example, in May 2019, eBird received 7.5 observations *per second* throughout the entire month. This is a rate of data collection never before seen in ecology and biodiversity research. Species distribution models and abundance models, however, are often treated as static objects in the current scientific literature. As statistical power increases and researchers are able to estimate biodiversity changes in space and time with greater certainty, automated or semiautomated pipelines are being developed that are dynamically updated as data are contributed to the data set (Callaghan et al. 2019a). This near real-time approach will have the added benefit of detecting sudden declines more quickly (Inger et al. 2015), and the temporal scale of updated trends and status will be relevant to a given taxon and likely dependent on the rate of the data being submitted. For example, GreenMaps used broadscale data from GBIF on plant occurrences to develop modeled range maps for more than 190,000 species, which can then be validated by citizen scientists on the basis of the occurrence of each species in the field. This approach can be automated to continuously update the range maps, providing increased confidence surrounding a given species modeled range map. This is similar to global aggregation of all biodiversity records from museum, herbarium, government, and citizen science sources into GBIF. Another example is an automated method developed by the US National Park Service in combination with iNaturalist, which uses citizen science observations integrated with species lists for National Parks to detect species' responses to climate change (Boydston et al. 2017). Pipelines should be prepared on cloud computing platforms as it is becoming increasingly difficult to download, let al.one analyze, the large data sets created by citizen science projects on a personal computer.

Capitalizing on secondary data collection

Biodiversity research using citizen science has to date largely been focused on recording taxa in time and space—that is, putting biodiversity points on a map (e.g., Adesh et al. 2019, Humphreys et al. 2019). In addition to this main objective of points on a map, many citizen science records, particularly those relying on physical evidence (e.g., photographs or video or audio recordings), contain valuable secondary data such as information about habitat associations or species interactions. We define species observations submitted to citizen science platforms (e.g., iNaturalist, Macaulay Library) with the intention of putting a point on the map as the *primary data*. *Secondary data* is any additional information incidentally captured with that primary observation. Image-based records potentially contain a vast amount of information about species interactions with the natural and human environment additional to the primary observation.

Behavior, interspecific interactions, condition (e.g., breeding or health status), traits of an individual (e.g., phenotypes), microhabitat information, or the presence of additional species (e.g., co-occurrence) are examples of secondary data found in citizen science observations (figure 2). For example, automatic identification of individual animals has been used to understand the biology, habitat use, and population dynamics of whale sharks (Diamant et al. 2018, Norman et al. 2017, McCoy et al. 2018), identify individual cetaceans (Weideman et al. 2017), and reveal site fidelity in tiger sharks (Paxton et al. 2019). Internet images have been used to study commensal relationships between birds and herbivorous mammals (Mikula et al. 2018), bird–bird associations (Mikula and Tryjanowski 2016), and associations between plant species and pollinating insects (Bahlai and Landis 2016, Gazdic and Groom 2019). Leighton and colleagues (2016) used Internet images to study the distribution of white morphs of black bears, the distribution of color variants of black sparrowhawks and barn owls, and the hybridization coloration of carrion and hooded crows. Photographs uploaded to iNaturalist have been used to study variation in the wing patterns of damselflies across the species' geographic extent (Drury et al. 2019). The concept of harvesting secondary data from citizen science photographs is similar to the varied unanticipated uses of traditional museum collections (e.g., DNA, understanding DDT prevalence in egg shells) that have been fundamental for ecology and conservation (e.g., Suarez and Tsutsui 2004, Heberling and Isaac 2017). Even without DNA technology fully developed when many museum specimens were originally collected, the technological revolution in DNA analyses have found historical museum specimens instrumental. Photographs contributed by citizen scientists will likely yield similar results, although they are currently difficult to automatically process.

As quantities of biodiversity data continue to grow exponentially (Farley et al. 2018), it is important that robust, open-access infrastructure is implemented to allow appropriate filtering and management of these data (Bayraktarov et al. 2019). For example, tools should be implemented to allow identifications to be either shared among citizen scientists (i.e., "crowdsourcing"), reducing the effective workload, or fully automated using machine learning techniques. One such filtering tool is the "Project" feature on iNaturalist. Projects allow the collation of data, allowing grouping by location, taxon or a combination of both. This collation can occur automatically using the observation's metadata (e.g., GPS coordinates), or manually by individual users. The latter represents a form of crowdsourcing as the onus of filtering is on the many observers themselves instead of a single researcher. A recent study in North America, for example, identified bird collision hotspots and informed decisions on mortality prevention through building retrofitting (Winton et al. 2018). The data for this study were collated through an iNaturalist project (www.inaturalist.org/projects/bird-window-collisions), a framework without which data collation

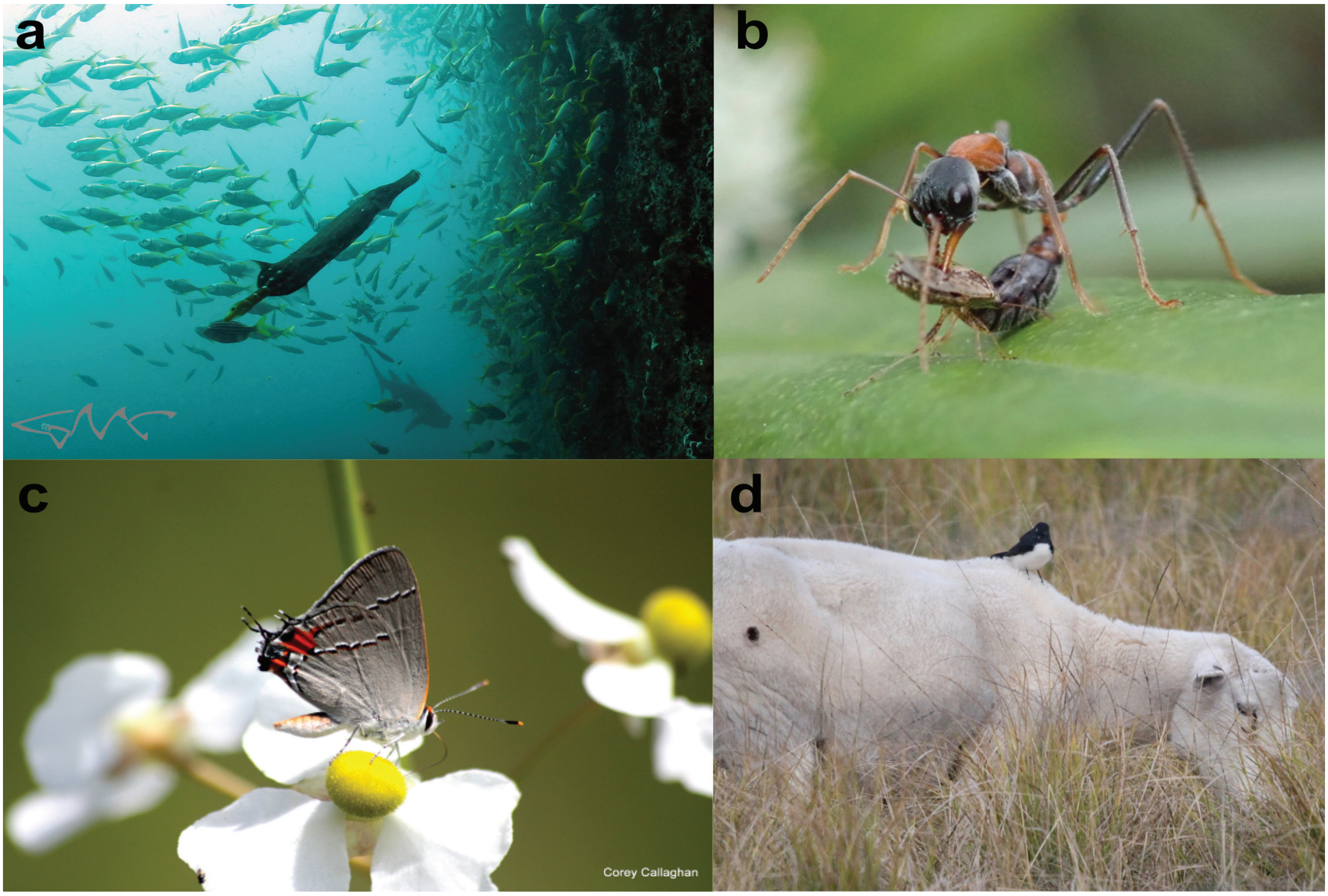

*Figure 2. Examples of the diversity of secondary data that can be extracted from biodiversity observations. (a) Community composition: A single image shows the presence of eastern pomfred (***Schuettea scalaripinis***), yellowtail scad (***Trachurus novaezelandiae***), grey nurse shark (***Carcharias taurus***) and painted trumpetfish (***Aulostomus chinensis***). (b) Species interactions: a jumping jack ant (***Myrmecia nigrocincta***) is preying on a seed bug (***Nysius** sp.). (c) Species interaction and phenology: gray hairstreak (***Strymon melinus***) feeding on a flowering bulltongue arrowhead (***Sagittaria lancifolia***). (d) Commensalism: Willie Wagtail (***Rhipidura leucophrys***) foraging for insects from the back of a domestic sheep (***Ovis aries***) and collecting nesting material.*

would have been impractical, given there are around 5 million observations of birds on iNaturalist (Van Horn et al. 2018). Automated image classification is being implemented across a wide range of projects to extract ecologically valuable information from imagery efficiently and cost effectively (Weinstein 2018). Wildbook is an example open-source platform designed to identify individual organisms on the basis of natural markings using deep convolutional neural network machine learning (http://wildbook.org). Expanding and replicating automated image classification tools such as those developed by Wildbook is therefore a priority for expediting the collation and analysis of citizen science biodiversity data.

Conclusions

In the present article, we highlight three important future directions for citizen science—among many possible directions—that will help to increase the utility of citizen science data for biodiversity research in the future. Some limitations facing citizen science that we address in the present article include strong societal preferences toward charismatic flora and fauna, a lack of taxonomic expertise in specific taxa to identify images, a lack of funding and technical expertise for citizen science practitioners to develop cloud computing pipelines, and a substantial cost and investment by government and other funding sources to develop automated image recognition technology to harvest secondary data. We believe that the examples presented above help to illustrate that focused collaborations between citizen science participants and professional scientists can overcome these limitations and truly maximize the potential of citizen science data.

As these three frontiers continue to be developed, there are a myriad of scientific questions that can be better addressed. In the present article, we highlight six such objectives—two pertaining to each of the three frontiers—that we believe will benefit from advances in each of the respective frontiers.

Sampling the undersampled. If citizen scientist can sample enough private land, then we will gain an increased understanding of the role private lands play in biodiversity conservation (e.g., Bean and Wilcove 1997). If citizen science and ecotourism can be better linked, then ecotourism projects can both gain valuable data for citizen science projects and bring ecotourism to remote areas with flow-on effects for conservation (e.g., Orams 1995).

Estimating species' abundances in space and time. If professional scientists can build data sets that complement citizen science data, we can identify high priority sites and species that can be used to identify species trends more quickly (Bayraktarov et al. 2019). If there are further developments of semiautomated or automated pipelines to interact with citizen scientists in near real time then the collective effort of citizen scientists will be able to reduce redundancies and gaps in the data collected (Callaghan et al. 2019a, 2019b).

Capitalizing on secondary data collection. If species interactions can be quantified from citizen science photographs at scale, then we can start to better understand the co-occurrence of species in time and space, highlighting key taxa for conservation (e.g., pollination ecology; Domroese and Johnson 2017). If a formal review of the secondary data that has to date been harvested from citizen science photographs is conducted, then we can begin to fully understand the potential these data hold for ecology and conservation across taxa and projects.

Citizen science is currently seeing a rapid increase in contributions from volunteers with the number of citizen science projects, and therefore, biodiversity observations are growing exponentially (Pocock et al. 2017). We believe that it is time to move past the focus on the limitations of these data (after all, no data are perfect), and begin to take advantage of the extraordinary opportunities these data present (Burgess et al. 2017, Tulloch et al. 2013b). If professional scientists develop the right tools—presented in the article around three frontiers—citizen science data can be an important part of future advances in ecology, conservation, and biogeography (McKinley et al. 2017), dramatically advancing our understanding of global biodiversity.

Acknowledgments

We thank the thousands of participants of citizen science projects who tirelessly contribute to citizen science projects. Three anonymous reviewers and the editor provided thoughtful comments that substantially improved the manuscript.

References cited

[ABS] Australian Bureau of Statistics. 2002. Year Book Australia, 2002. https://www.abs.gov.au/Ausstats/abs%40.nsf/94713ad445ff1425ca25682000192af2/c4f37e4488d32591ca256b35007ace07!OpenDocument.

Aceves-Bueno E, Adeleye AS, Feraud M, Huang Y, Tao M, Yang Y, Anderson SE. 2017. The accuracy of citizen science data: A quantitative review. Bulletin of the Ecological Society of America 98: 278–290.

Adesh, K., Ankit S, and Amita K. 2019. Using citizen science in assessing the distribution of Sarus Crane (Grus antigone antigone) in Uttar Pradesh, India. International Journal of Biodiversity and Conservation 11: 58–68.

Andrachuk M, Marchke M, Hings C, Armitage D. 2019. Smartphone technologies supporting community-based environmental monitoring and implementation: A systematic scoping review. Biological Conservation 237: 430–442.

Bahlai CA, Landis DA. 2016. Predicting plant attractiveness to pollinators with passive crowdsourcing. Royal Society Open Science 3: 150677.

Bayraktarov E, Ehmke G, O'Connor J, Burns EL, Nguyen HA, McRae L, Possingham HP, Lindenmayer DB. 2019. Do big unstructured biodiversity data mean more knowledge? Frontiers in Ecology and the Environment 6: 239.

Bean MJ, Wilcove DS. 1997. The private-land problem. Conservation Biology 11: 1–2.

Boakes EH, McGowan PJ, Fuller RA, Chang-qing D, Clark NE, O'Connor K, Mace GM. 2010. Distorted views of biodiversity: Spatial and temporal bias in species occurrence data. PLOS Biology 8: e1000385.

Bonnet X, Shine R, Lourdais O. 2002. Taxonomic chauvinism. Trends in Ecology and Evolution 17: 1–3.

Bonney R, Shirk JL, Phillips TB, Wiggins A, Ballard HL, Miller-Rushing A, Parrish JK. 2014. Nest steps for citizen science. Science 348: 1436–1437.

Bonter DN, Cooper CB. 2012. Data validation in citizen science: A case study from Project FeederWatch. Frontiers in Ecology and the Environment 10: 305–307.

Boydston EE, Barve V, Lee L, Morelli TL, Briggs JS. 2017. Automating the Use of Citizen Scientists' Biodiversity Surveys in iNaturalist to Facilitate Early Detection of Species' Responses to Climate Change. US Geological Survey. www.sciencebase.gov/catalog/item/58b5cd07e4b01ccd54fde37f.

Bowser A, Hansen D, He Y, Boston C, Gunnell L, Reid M, Preece J. 2013. Using gamification to inspire new citizen science volunteers. Pages 18–25 in Nacke LE, Harrigan K, Randall N, eds. Gamification '13: Proceedings of the First International Conference on Gameful Design, Research, and Applications. Association for Computing Machinery.

[BTO] British Trust for Ornithology. 2020. Gardenwatch. https://www.bto.org/our-science/projects/gardenwatch.

Burgess HK, DeBey L, Froehlich H, Schmidt N, Theobald EJ, Ettinger AK, HilleRisLambers J, Tewksbury J, Parrish JK. 2017. The science of citizen science: Exploring barriers to use as a primary research tool. Biological Conservation 208: 113–120.

Callaghan CT, Rowley JJ, Cornwell WK, Poore AG, Major RE. 2019a. Improving big citizen science data: Moving beyond haphazard sampling. PLOS Biology 17: e3000357.

Callaghan CT, Poore AG, Major RE, Rowley JJ, Cornwell WK. 2019b. Optimizing future biodiversity sampling by citizen scientists. Proceedings of the Royal Society B 286: 20191487.

Chandler M, et al. 2017a. Contribution of citizen science towards international biodiversity monitoring. Biological Conservation 213: 280–294.

Chandler M, See L, Buesching CD, Cousins JA, Gillies C, Kays RW, Newman C, Pereira HM, Tiago P. 2017b. Involving citizen scientists in biodiversity observation. Pages 211–237 in Walters M, Scholes RJ, eds. The GEO Handbook on Biodiversity Observation Networks. Springer.

Cooper CB, Dickinson J, Phillips T, Bonney R. 2007. Citizen science as a tool for conservation in residential ecosystems. Ecology and Society 12: e11.

Currie JJ, Stack SH, Kaufman GD. 2018. Conservation and education through ecotourism: Using citizen science to monitor cetaceans in the four-island region of Maui, Hawaii. Tourism in Marine Environments 13: 65–71.

Czech B, Krausman PR, Borkhataria R. 1998. Social construction, political power, and the allocation of benefits to endangered species. Conservation Biology 12: 1103–1112.

Danielsen F, Pirhofer-Walzl K, Adrian TP, Kapijimpanga DR, Burgess ND, Jensen PM, Bonney R, Funder M, Landa A, Levermann N, Madsen J. 2014. Linking public participation in scientific research to the indicators and needs of international environmental agreements. Conservation Letters. 7: 12–24.

Das M, Chatterjee B. 2015. Ecotourism: A panacea or a predicament? Tourism Management Perspectives 14: 3–16.

Diamant S, Rohner CA, Kiszka JJ, d Echon AG, d Echon TG, Sourisseau E, Pierce SJ. 2018. Movements and habitat use of satellite-tagged whale sharks off western madagascar. Endangered Species Research 36: 49–58.

Domroese MC, and Johnson EA. 2017. Why watch bees? Motivations of citizen science volunteers in the Great Pollinator Project. Biological Conservation 208: 40–47.

Drury JP, Barnes M, Finneran AE, Harris M, Grether GF. 2019. Continent-scale phenotype mapping using citizen scientists' photographs. Ecography 42: 1436–1445.

Farley SS, Dawson A, Goring SJ, Williams JW. 2018. Situating ecology as a big-data science: Current advances, challenges, and solutions. BioScience 68: 563–576.

Fink, D, Hochachka WM, Zuckerberg B, Winkler D, Shaby B, Munson MA, Hooker G, Riedewald M, Sheldon D, Kelling S. 2010. Spatiotemporal exploratory models for broad-scale survey data. Ecological Applications 20: 2131–2147.

Fithian W, Elith J, Hastie T, Keith DA. 2015. Bias correction in species distribution models: Pooling survey and collection data for multiple species. Methods in Ecology and Evolution 6: 424–438.

Gazdic M, Groom Q. 2019. iNaturalist is an unexploited source of plant–insect interaction data. Biodiversity Information Science and Standards 3: e37303.

Gonsamo A, D'Odorico P. 2014. Citizen science: Best practices to remove observer bias in trend analysis. International journal of biometeorology 58: 2159–2163.

Gotelli NJ, Colwell RK. 2001. Quantifying biodiversity: Procedures and pitfalls in the measurement and comparison of species richness. Ecology Letters 22: 379–391.

He KS, Bradley BA, Cord AF, Rocchini D, Tuanmu MN, Schmidtlein S, Turner W, Wegmann M, Pettorelli N. 2015. Will remote sensing shape the next generation of species distribution models? Remote Sensing in Ecology and Conservation 1: 4–18.

Heberling JM, Isaac BL. 2017. Herbarium specimens as expectations: New uses for old collections. American Journal of Botany 104: 963–965. doi:10.3732/ajb.1700125

Hilty J, Merenlender AM. 2003. Studying biodiversity on private lands. Conservation Biology 17: 132–137.

Humphreys JM, Murrow JL, Sullivan JD, Prosser DJ, and Zurell D. 2019. Seasonal occurrence and abundance of dabbling ducks across the continental United States: Joint spatio-temporal modelling for the genus *Anas*. Diversity and Distributions 25: 1497–1508.

Inger R, Gregory R, Duffy JP, Stott I, Voříšek P, Gaston KJ. 2015. Common European birds are declining rapidly while less abundant species' numbers are rising. Ecology letters 18: 28–36.

Johnston A, Moran N, Musgrove A, Fink D, Baillie SR. 2020. Estimating species distributions from spatially biased citizen science data. Ecological Modelling 422: 108927.

Jones CD, Glon MG, Cedar K, Paiero SM, Pratt PD, Preney TJ. 2019. First record of Paintedhand Mudbug (*Lacunicambarus polychromatus*) in Ontario and Canada and the significance of iNaturalist in making new discoveries. Canadian Field-Naturalist 133: 160–166.

Jordan R, Crall A, Gray S, Phillips T, Mellor D. 2015. Citizen science as a distinct field of inquiry. BioScience 65: 208–211.

Kattge J, et al. 2011. Try: A global database of plant traits. Global Change Biology 17: 2905–2935.

Kosmala M, Wiggins A, Swanson A, Simmons B. 2016. Assessing data quality in citizen science. Frontiers in Ecology and the Environment 14: 551–560.

Kwok, R. 2018. Ecology's remote-sensing revolution. Nature 556: 137–138.

Leighton GR, Hugo PS, Roulin A, Amar A. 2016. Just google it: Assessing the use of google images to describe geographical variation in visible traits of organisms. Methods in Ecology and Evolution 7: 1060–1070.

Marshall NJ, Kleine DA, Dean AJ. 2012. CoralWatch: Education, monitoring, and sustainability through citizen science. Frontiers in Ecology and the Environment 10: 332–334.

May RM. 1988. How many species are there on Earth? Science 241: 1441–1449.

McCoy E, Burce R, David D, Aca EQ, Hardy J, Labaja J, Snow SJ, Ponzo A, Araujo G. 2018. Long-term photo-identification reveals the population dynamics and strong site fidelity of adult whale sharks to the coastal waters of Donsol, Philippines. Frontiers in Marine Science 5: 271.

McKinley DC, et al. 2017. Citizen science can improve conservation science, natural resource management, and environmental protection. Biological Conservation 208: 15–28.

Menon S, Berger-Wolf TY, Kiciman E, Joppa L, Stewart CV, Crall PJ, Holmberg J, Van Oast J. 2016. Animal population estimation using Flickr images. Paper presented at the 2nd International Workshop on the Social Web for Environmental and Ecological Monitoring (SWEEM 2017), June 25, Troy, NY.

Meyer C, Kreft H, Guralnick R, Jetz W. 2015. Global priorities for an effective information basis of biodiversity distributions. Nature communications 6: 8221.

Mieras PA, Harvey-Clark C, Bear M, Hodgin G, Hodgin B. 2017. The economy of shark conservation in the Northeast Pacific: The role of ecotourism and citizen science. Advances in Marine Biology 78: 121–153.

Mikula P, Hadrava J, Albrecht T, Tryjanowski P. 2018. Large-scale assessment of commensalistic–mutualistic associations between African birds and herbivorous mammals using Internet photos. PeerJ 6: e4520.

Mikula P, Tryjanowski P. 2016. Internet searching of bird–bird associations: A case of bee-eaters hitchhiking large African birds. Biodiversity Observations 7: 1–6.

Newman G, Wiggins A, Crall A, Graham E, Newman S, Crowston K. 2012. The future of citizen science: Emerging technologies and shifting paradigms. Frontiers in Ecology and the Environment 10: 298–304.

Norman BM, et al. 2017. Undersea constellations: The global biology of an endangered marine megavertebrate further informed through citizen science. BioScience 67: 1029–1043.

Orams MB. 1995. Towards a more desirable form of ecotourism. Tourism Management 16: 3–8.

Orr MC, Ascher JS, Bai M, Chesters D, Zhu CD. 2020. Three questions: How can taxonomists survive and thrive worldwide? Megataxa 1: 19–27.

Pacifici K, Reich BJ, Miller DA, Gardner B, Stauffer G, Singh S, McKerrow A, Collazo JA. 2017. Integrating multiple data sources in species distribution modeling: A framework for data fusion. Ecology 98: 840–850.

Paxton AB, et al. 2019. Citizen science reveals female sand tiger sharks (*Carcharias taurus*) exhibit signs of site fidelity on shipwrecks. Ecology 100: e02687.

Pocock MJ, Tweddle JC, Savage J, Robinson LD, Roy HE. 2017. The diversity and evolution of ecological and environmental citizen science. PLOS ONE 12: e0172579.

Pocock MJ, 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: 274–281

Roberts DR, et al. 2017. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography 40: 913–929.

Rowley JJ, Callaghan CT, CutajaR T, Portway C, Potter K, Mahony S, Trembath DF, Flemons P, Woods A. 2019. FrogID: Citizen scientists provide validated biodiversity data on frogs of Australia. Herpetological Conservation and Biology 14: 155–170.

[SANBI] South African National Biodiversity Institute. 2020. Custodians of rare and endangered wildflowers (CREW) programme. SANBI. www.sanbi.org/biodiversity/building-knowledge/biodiversity-monitoring-assessment/custodians-of-rare-and-endangered-wildflowers-crew-programme.

Schmeller DS, et al. 2009. Advantages of volunteer-based biodiversity monitoring in europe. Conservation biology 23: 307–316.
Schneider FD, et al. 2019. Towards an ecological trait-data standard. Methods in Ecology and Evolution 10: 2006–2019.
Scott JM, Davis FW, McGhie RG, Wright RG, Groves C, Estes J. 2001. Nature reserves: Do they capture the full range of america's biological diversity? Ecological Applications 11: 999–1007.
Shirk JL, et al. 2012. Public participation in scientific research: A framework for deliberate design. Ecology and Society. 1: 17.
Suarez AV, Tsutsui ND. 2004. The value of museum collections for research and society. BioScience 54: 66–74.
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 AI, Mustin K, Possingham HP, Szabo JK, Wilson KA. 2013a. To boldly go where no volunteer has gone before: Predicting volunteer activity to prioritize surveys at the landscape scale. Diversity and Distributions 19: 465–480.
Tulloch AIT, Possingham HP, Joseph LN, Szabo J, Martin TG. 2013b. Realising the full potential of citizen science monitoring programs. Biological Conservation 165: 128–138.
Van Horn G, Mac Aodha O, Song Y, Cui Y, Sun C, Shepard A, Adam H, Perona P, Belongie S. 2018. The iNaturalist species classification and detection dataset. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 8769–8778.
Ward DF. 2014. Understanding sampling and taxonomic biases recorded by citizen scientists. Journal of Insect Conservation 18: 753–756.
Weideman HJ, et al. 2017. Integral curvature representation and matching algorithms for identification of dolphins and whales. In Proceedings of the IEEE International Conference on Computer Vision. IEEE. doi:10.1109/ICCVW.2017.334
Weinstein BG. 2018. A computer vision for animal ecology. Journal of Animal Ecology 87: 533–545.
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: 1890.
Winton RS, Ocampo-Peñuela N, Cagle N. 2018. Geo-referencing bird–window collisions for targeted mitigation. PeerJ 6: e4215.
Yesson C, et al. 2007. How global is the global biodiversity information facility? PLOS ONE 2: e1124.
Young BE, Dodge N, Hunt PD, Ormes M, Schlesinger MD, Shaw HY. 2019. Using citizen science data to support conservation in environmental regulatory contexts. Biological Conservation 237: 57–62.
Zapponi L, et al. 2017. Citizen science data as an efficient tool for mapping protected saproxylic beetles. Biological Conservation 208: 139–145.

Corey T. Callaghan (c.callaghan@unsw.edu.au), Thomas Mesaglio, Jodi J. L. Rowley, John H. Wilshire, and William K. Cornwell are affiliated with the Centre for Ecosystem Science, in the School of Biological, Earth, and Environmental Sciences at the University of New South Wales, in Sydney, New South Wales, Australia. Corey T. Callaghan, Alistair G. B. Poore, Angela T. Moles, Shinichi Nakagawa, Christopher Roberts, Adriana Vergés, and William K. Cornwell are affiliated with the Ecology and Evolution Research Centre, in the School of Biological, Earth, and Environmental Sciences, also at the University of New South Wales. Jodi J. L. Rowley is affiliated with the Australian Museum Research Institute, part of the Australian Museum, in Sydney, New South Wales, Australia.