The risks and rewards of community science for threatened species monitoring

Peter Soroye1 | Brandon P. M. Edwards2 | Rachel T. Buxton2

Jeffrey P. Ethier1 | Acacia Frempong-Manso2 | Hannah E. Keefe1

Albana Berberi2 | Maisy Roach-Krajewski1 | Allison D. Binley2

Jaimie G. Vincent2 | Hsien-Yung Lin2 | Steven J. Cooke2 | Joseph R. Bennett2

1Department of Biology, University of Ottawa, Ottawa, Canada
2Department of Biology, Carleton University, Ottawa, Canada

Abstract
Finding ways of efficiently monitoring threatened species can be critical to effective conservation. The global proliferation of community science (also called citizen science) programs, like iNaturalist, presents a potential alternative or complement to conventional threatened species monitoring. Using a case study of ~700,000 observations of >10,000 IUCN Red List Threatened species within iNaturalist observations, we illustrate the potential risks and rewards of using community science to monitor threatened species. Poor data quality and risks of sending untrained volunteers to sample species that are sensitive to disturbance or harvesting are key barriers to overcome. Yet community science can expand the breadth of monitoring at little extra cost, while indirectly benefiting conservation through outreach and education. We conclude with a list of actionable recommendations to further mitigate the risks and capitalize on the rewards of community science as a threatened species monitoring tool.

KEYWORDS
citizen science, community science, outreach, iNaturalist, participatory monitoring, species assessments, threatened species

1 | INTRODUCTION
Monitoring threatened species often requires collecting rigorous long-term data at large spatial scales, and correspondingly significant financial investment (Parsons et al., 2018; Theobald et al., 2015). Currently, monitoring accounts for an average of half of threatened species conservation budgets (Buxton et al., 2020). Exploring more cost-effective methods to monitor threatened species could enable funds to be redirected towards conservation actions, which are often needed quickly to avoid extinction (Martin et al., 2012).

Community science (CS) data are increasingly used in threatened species monitoring programs (Callaghan et al., 2020; Carpenter et al., 2017; Rosenberg et al., 2017). Harnessing this growing source of information has

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. Conservation Science and Practice published by Wiley Periodicals LLC on behalf of Society for Conservation Biology.
enormous potential to reduce the cost of monitoring while increasing the scope of monitoring and public engagement in conservation (Ellwood et al., 2017; Soroye et al., 2018). However, using CS programs for threatened species monitoring also comes with considerable risks (Cooke et al., 2016; Meeuwig et al., 2015; Ottinger, 2010).

In this data-driven perspective, we outline the risks and rewards of using CS for monitoring threatened species. We inform much of our discussion with a case study that uses data from iNaturalist, one of the most widely-used CS platforms. iNaturalist contains data on a global scale and on a wide range of taxa, and thus provides an excellent illustration of potential utility for threatened species monitoring. We downloaded all research-grade iNaturalist observations from the Global Biodiversity Information Facility (GBIF.org; accessed on September 27, 2020, https://doi.org/10.15468/dl.6m63rx; Ueda 2020), and filtered to all Threatened species from the IUCN Red List (accessed on September 28, 2020; IUCN, 2012). We used these filtered data to illustrate some risks of using CS data by considering the reliability of the data, sampling biases, and increased disturbances to threatened species; we also illustrate some rewards of using CS data by considering the increased scope of threatened species data, the inspiration of conservation advocacy, and the contribution to policy action (see Supplemental Materials for specifics pertaining to data-filtering and data use). We conclude by sharing recommendations for incorporating CS in threatened species monitoring and conservation policy today and in the future (Figure 1).

2 | RISKS

2.1 | Reliability of data

While CS programs have allowed for an increased understanding of topics such as shifts in species distribution, migration routes, disease patterns, and effects of climate change (Dickinson et al., 2010), many researchers and decision-makers remain skeptical of the capacity of unpaid volunteers to generate high-quality datasets needed for rigorous assessments of species population change (Aceves-Bueno et al., 2017; Kosmala et al., 2016). Because the precision of threatened species data can have legal implications (e.g., whether they are listed for protection), accuracy of data and comparison of participants’ skill sets to professionals is particularly important. Generally, the quality of CS data provided by volunteers are more variable than data collected by professionals (e.g., see Aceves-Bueno et al., 2017; Austen et al., 2016; Hoyer et al., 2012; Kosmala et al., 2016), but can often reach an acceptable threshold for scientific research. CS’s reliability for monitoring threatened species, which are often rare, may be lower (Cox et al., 2012) which may result in biases that differ from those concerning other species (Theobald et al., 2015). Understanding whether biases associated with CS for threatened species confirm or contradict these known biases is important to our understanding of how CS can be used to help these species.

Since it was created in 2008, iNaturalist has collected over 700,000 research-grade observations of threatened species data, indicating that each record was identified to species-level or lower by at least two out of three community identifiers in iNaturalist, usually based on a photograph and/or sounds submitted by the observer, and that the observation is not of a captive or cultivated organism. These observations consist of more than 10,000 species appearing as Data Deficient or in various classifications of endangerment on the IUCN Red List. However, there are clear biases towards certain taxa, such as birds, plants, and mammals (Table 1). These taxonomic biases are well-known in large biodiversity datasets (Mair & Ruete, 2016; Troudet et al., 2017) and in the IUCN Red List (Donaldson et al., 2016), limiting understanding of
conservation trends worldwide. However, we note that compared to professionally collected data that generally have a bias towards birds and mammals (Troudet et al., 2017), the iNaturalist data collected here have a high percentage of “less charismatic” species such as plants and invertebrates.

Additionally, we found that threatened species data in iNaturalist show evidence of spatial sampling bias (Figure 2). Numbers of threatened species observed appear highest in the U.S. (1427 species) and Mexico (1231 species) (Figure 2a). The U.S., Canada, Mexico, Russia, and New Zealand contained the most observations of threatened species, making up ~58% of all threatened species sightings in iNaturalist (Figure 2b). CS data for all species also tend to be spatially clustered in locations easily accessible by road and within residential and urban areas (Dickinson et al., 2010). Using broad categories of global land cover types (see Supplemental Materials), we found similar patterns for threatened species data, where observations in urban and cropland/pasture environments were disproportionately represented, and few observations were recorded in land cover types that are more challenging to access (e.g., wetland, tundra, and barren of sparsely vegetated; Table S1).

Threatened species data in iNaturalist show evidence of temporal sampling bias. These temporal biases are similar to biases seen across other large CS programs such as eBird (Zhang, 2020). Total observations of threatened species peak dramatically in late-spring and summer in the Northern Hemisphere, and in fall and spring in the Tropics and Southern Hemisphere (Figure 3). Given that the number of observers submitting observations also peaks during these periods, this could be due at least in part to biased effort (Knape et al., 2022). However, it could also be due to annual patterns in presence and detectability of species (Sólymos et al., 2018; Zuckerberg et al., 2016).

2.2 | Increased risk to threatened species

Encouraging the public to seek out threatened species can have risks. We found that human disturbance or poaching and harvesting are listed as major threats for 57.9% of threatened species reported in iNaturalist. Only 38% of all Red List threatened species are at risk from these threats, suggesting that the threatened species reported to iNaturalist disproportionately tend to be those threatened by disturbance and harvesting. Even for species not at high risk of poaching or illegal harvesting, incentivizing non-professional monitoring creates the potential for negative impacts from disturbance (Quinn, 2021). Recreation was specifically listed as a threat to 907 (8.5%) threatened species in iNaturalist. Disturbance analogous to that caused by recreation increases the risk that CS will unintentionally facilitate poaching by habituating wildlife and modifying behavior (Baral, 2013; Larson et al., 2016).

Publishing location information for threatened species, which are often rare and economically valuable, could enable poaching and harvesting (Lindenmayer &

<table>
<thead>
<tr>
<th>Red List category</th>
<th>Number of species (%)</th>
<th>Total observations (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data deficient</td>
<td>1818 (17.09)</td>
<td>130,988 (18.66)</td>
</tr>
<tr>
<td>Near threatened</td>
<td>2681 (25.2)</td>
<td>258,724 (36.85)</td>
</tr>
<tr>
<td>Vulnerable</td>
<td>2981 (28.02)</td>
<td>197,337 (28.11)</td>
</tr>
<tr>
<td>Endangered</td>
<td>2257 (21.22)</td>
<td>85,493 (12.18)</td>
</tr>
<tr>
<td>Critically endangered</td>
<td>855 (8.04)</td>
<td>25,885 (3.69)</td>
</tr>
<tr>
<td>Regionally extinct</td>
<td>17 (0.16)</td>
<td>3107 (0.44)</td>
</tr>
<tr>
<td>Extinct in the wild</td>
<td>20 (0.19)</td>
<td>468 (0.07)</td>
</tr>
<tr>
<td>Extinct</td>
<td>9 (0.08)</td>
<td>12 (0)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Taxa</th>
<th>Number of species (%)</th>
<th>Total observations (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Birds</td>
<td>1626 (15.28)</td>
<td>228,446 (32.54)</td>
</tr>
<tr>
<td>Plants</td>
<td>3284 (30.87)</td>
<td>153,541 (21.87)</td>
</tr>
<tr>
<td>Invertebrates</td>
<td>1565 (14.71)</td>
<td>139,273 (19.84)</td>
</tr>
<tr>
<td>Mammals</td>
<td>837 (7.87)</td>
<td>75,539 (10.76)</td>
</tr>
<tr>
<td>Reptiles and amphibians</td>
<td>1931 (18.15)</td>
<td>66,662 (9.5)</td>
</tr>
<tr>
<td>Fish</td>
<td>1315 (12.36)</td>
<td>36,852 (5.25)</td>
</tr>
<tr>
<td>Fungi</td>
<td>80 (0.75)</td>
<td>1701 (0.24)</td>
</tr>
</tbody>
</table>
Scheele, 2017; McMillan et al., 2020). In recent years, the potential for online data repositories to be misused or breached by hackers has created legitimate concern that data on threatened species’ locations may be abused (Lennox et al., 2020). In Australia for example, biotelemetry tracking data used for Endangered white shark research was misused to locate and kill tagged animals who got close to bathing beaches (Meeuwig et al., 2015). Additionally, the use of the internet has amplified the exotic pet trade industry (Morgan & Chng, 2017), which is a major driver of threatened species poaching (Herrera & Hennessey, 2007; McMillan et al., 2020). To mitigate some of these threats, CS programs like iNaturalist and eBird have taken steps to prevent the access to exact locations of rare or threatened species, by obscuring or completely hiding the locations.

Finally, sharing CS data could also pose a privacy risk to users, since the management of CS programs sometimes requires gathering private information about volunteers (e.g. age, gender, background information) or privately-owned land (Anhalt-Depies et al., 2019). These are risks for CS data in general, however making location data for threatened species publicly accessible can disrupt the relationship between researchers and landowners if it
enables bad actors to use these data to trespass in search of the organisms (Lindenmayer & Scheele, 2017).

3 | REWARDS

3.1 | Increasing the scope of threatened species monitoring

Well-vetted CS data can complement conventional sampling methods, expanding the scope of sampling and even addressing gaps in conventional threatened species monitoring (Binley et al., 2021; Dickinson et al., 2010; Donaldson et al., 2016). We found that while iNaturalist data generally indicate fewer threatened species in a given region than previous observations (see Supplementary Materials), it can match, or even exceed, the previously known vertebrate threatened species richness in 4.5% of sites globally, sometimes finding four times more species than was previously known (Figure S4). When focusing on threatened birds, mammals, and amphibians, iNaturalist matches or exceeds threatened species richness in 9.7%, 29.9%, and 46.8% of sites respectively. CS programs have already been able to generate previously unknown ecological, distributional, and phenological information for threatened species around the globe (Soroye et al., 2018) and across several threatened taxa (e.g., turtles [Cross et al., 2021], plants [Garcia et al., 2021], primates [Ang et al., 2021], birds [Domínguez et al., 2020], Squires et al., 2021], butterflies [Ries & Oberhauser, 2015, Sanderson et al., 2021]).

By engaging landowners and sampling across private lands, CS could also increase coverage in areas underrepresented in conventional threatened species monitoring programs. Large percentages of the natural environment where threatened species occur are privately owned (Chacon, 2005; Ciuzio et al., 2013) and globally, many parks and reserves contain significant portions of privately owned land within their boundaries (Kamal et al., 2015). In iNaturalist, we found that ~30% of threatened species observations were reported in areas classified as urban, cropland, or pasture (all of which are presumably largely private lands), despite these environments accounting for <5% of global land area (Table S1).

3.2 | Inspiring conservation advocacy

Encouraging CS monitoring of threatened species is an outreach opportunity that could help build a better public understanding of science, promote awareness and engagement in conservation, and facilitate environmental stewardship (Ellwood et al., 2017; Stepenuck & Green, 2015; Trumbull et al., 2000). Participating in CS can also give individuals the confidence to get more involved in conservation advocacy (Lewandowski & Oberhauser, 2017). Community participation in CS can build a desire for conservation and
enable real, and sometimes rapid, improvements in natural resource management (Stepenuck & Green, 2015).

Approximately 21% of all iNaturalist users (87,651 individuals) submitted threatened species data. However, 75% of threatened species observations were made by only 8.5% of users, suggesting a core of dedicated volunteers and an opportunity to reach more CS users with messaging geared towards threatened species monitoring and conservation. While CS volunteers tend to be skewed towards those who already have an interest and trust in science (Trumbull et al., 2000), CS of threatened species provides an opportunity to engage a diversity of people in conservation through schools, museums, zoos and other science hubs (Ballard et al., 2017; Dickinson et al., 2010).

3.3 | Contributing to policy and action

Jurisdictional boundaries often constrict government programs (Ciuzio et al., 2013). CS programs can inform conservation policies and actions across jurisdictions by expanding the scope of threatened species monitoring across these boundaries. One of the most prominent examples of this is the North American Breeding Bird Survey (Hudson et al., 2017). The Breeding Bird Survey is a volunteer program that provides the most reliable annual abundance indices for over 500 bird species and informs policy in the U.S. Endangered Species Act and bird species conservation assessments across North America (e.g., Hudson et al., 2017; Rosenberg et al., 2017). Internationally, CS monitoring data from programs such as the Breeding Bird Survey, eBird, iNaturalist, and Reef Life Survey (among several others; Chandler et al., 2017) contribute to several of the Group on Earth Observations Biodiversity Observation Network’s Essential Biodiversity Variables (Chandler et al., 2017) and five of the UN’s 244 Sustainable Development Goal indicators (Fraisl et al., 2020). Yet there is room for CS to contribute more. Over 30% of Sustainable Development Goal indicators, for example, could be informed by CS (Fraisl et al., 2020), and the use of CS is still limited in governmental species assessments (Lin et al., 2022).

3.3.1 | Examples from Canada

In Canada, the Species at Risk Act (SARA) mandates recovery strategies and action plans for threatened species and their habitats, while the Committee on the Status of Endangered Wildlife in Canada (COSEWIC) assesses extinction risk and recommends species to be included in the SARA. COSEWIC recognizes community knowledge as potentially important in threatened species status assessments (COSEWIC, 2021), and CS is used to inform listing and management for many species protected under SARA. For example, local volunteers’ long-term bird banding and wildlife sightings are used to inform recovery strategies for species at risk of extinction (Environment and Climate Change Canada, 2015). One such instance is in the recovery strategy of the piping plover (Charadrius melodus), in which volunteer monitoring and guardian programs are specifically mentioned as recovery activities (Environment and Climate Change Canada, 2021). In the Pacific Northwest, SARA action plans for protected areas highlight local CS programs promoting outreach and engagement (Parks Canada Agency, 2016). Tour operators in Gwaii Haanas National Park Reserve are encouraged to collect species data, and visitors undergo mandatory orientations on appropriate behaviors when encountering at-risk species and their respective habitats (Parks Canada Agency, 2016). In western Québec, a recently proposed plan for protected areas encourages visitors to use large-scale CS platforms, such as iNaturalist and eBird, to further inform regional species recovery measures (Parks Canada Agency, 2020). In a country with many remaining remote spaces, incorporating CS monitoring into SARA action plans and COSEWIC assessments offers a cost-efficient and community-engaging opportunity to monitor threatened species.

4 | RECOMMENDATIONS

Capitalizing on the rewards and mitigating the risks of using CS for threatened species monitoring is crucial for using CS to inform conservation. Species occurrence data are often a limiting factor in species-at-risk assessment. Therefore, practitioners can take advantage of a broader suite of data for conservation decisions by augmenting professionally collected data with CS data for threatened species.

We provide recommendations for structuring and using CS data for threatened species monitoring. They are not presented in order of importance; instead, we hope that practitioners will be able to use them in combination. These recommendations are summarized in in Figure 1, and Table S3 provides specific actions and examples for each of the recommendations below.

4.1 | Provide training and information to community scientists

Providing volunteers with information or training on species identification or best practices for monitoring could augment the accuracy of CS data while helping protect
threatened species from unnecessary disturbance. For example, providing information about species threats or “best practices” for safe observation following a submission of a sensitive species sighting (e.g., through a digital notification) could help alert well-meaning volunteers about the risks to species and habitats of careless sampling or observation techniques (Cooke et al., 2016). Additionally, CS programs could look to direct participants to under-sampled areas to fill in spatial gaps for threatened species. Training does not need to be intensive to improve data quality (Ellwood et al., 2017), and even automated feedback can help improve identification accuracy (van der Wal et al., 2016). For example, eBird provides simple quizzes for community scientists to practice their identification skills (see https://ebird.org/quiz/; Sullivan et al., 2014). Personalized training or data collection alongside professionals can also help improve data quality, and keeping professionals involved and engaged with volunteers increases opportunities for science outreach (Baker et al., 2014; Stepenuck & Green, 2015).

4.2 | Incorporate rigorous data quality protocols

For all CS programs, a robust data vetting process involving threatened species experts can help ensure high data quality after submission and will be especially important for rare and threatened species. Building in multiple layers of quality control throughout the data submission process that combine automated filters and expert input has shown to be a successful strategy in large-scale CS programs like eBird (Sullivan et al., 2014). Fine-tuning filters based on characteristics of the record (e.g., what species is being submitted, where the submission is located) and the recorder (e.g., recorder experience) can help catch errors while removing much of the burden from experts (Sullivan et al., 2014). Continuing to critically evaluate CS programs to better understand the strengths and flaws of their data will help increase trust and transparency (Aceves-Bueno et al., 2017; Callaghan et al., 2019; Kosmala et al., 2016).

4.3 | Develop plans for use, access, and security of threatened species data

With the increased collection of threatened species data, proactive plans for data storage, sharing, and use in CS programs become more important than ever. Opportunistic CS platforms like iNaturalist and eBird attempt to limit sharing of sensitive information of threatened species to prevent this (for example, by obscuring location information). We recommend that by default any observations of Threatened, Endangered, or commonly-poached species be either withheld from public viewing, or have their locations spatially obscured, so they cannot be used to harvest or exploit these species. Alternatively, CS programs could adopt a decision framework such as the one presented by Lennox et al. (2020) when considering how to deal with data collected of a threatened species.

4.4 | Foster connections with local communities

Encouraging connection to local communities through CS programs can further magnify the direct and indirect benefits of CS for threatened species monitoring. For example, the Plover Lovers organization in Sauble Beach, Canada, engages with the local community each year to allow residents to help the Endangered Piping Plover (Charadrius melodus) that nests on the shores of Sauble Beach (see https://www.ploverlovers.com/). Community scientists provide better data when they are more invested in a project and its potential outcomes (Dickinson et al., 2010), and the success of any conservation or monitoring depends partly on engagement from local communities (Díaz et al., 2018). This can also help mainstream the use of CS in threatened species monitoring. Connections with local communities can feed back into CS programs, driving collaborative or co-created programs around threatened species of particular local interest (Chiaravalloti et al., 2021; Shirk et al., 2012; Skarlatidou & Haklay, 2021).

5 | CONCLUSIONS

CS could lead to more effective solutions to conservation problems by providing cost-effective data at large scales. For threatened species, this is particularly important as failure to act quickly has led to extinctions in the past (Martin et al., 2012). The increasing popularity of CS programs has opened many new possibilities for ecological research and conservation. Through CS programs like iNaturalist, tens of thousands of volunteers engage in the global conservation effort by submitting millions of dated, geo-referenced observations of threatened species annually. These data have the potential to complement or – for some species and regions – even replace conventional monitoring efforts, which could divert large portions of conservation budgets away from monitoring and towards other much-needed conservation actions (Buxton et al., 2020). There can be risks associated with using CS for threatened species monitoring, but there are many ways to mitigate these risks and further
increase the benefits (Figure 1, Table S3). CS programs also have the potential to break down traditional barriers to participation in conservation. Embracing the use of CS data for threatened species monitoring could substantially shift the landscape of global conservation for the better.

**AUTHOR CONTRIBUTIONS**

RTB, JRB, SJC conceived the project idea. PS led analyses. All authors contributed to writing, editing, and revisions.

**ACKNOWLEDGMENTS**

We thank the many volunteers contributing to iNaturalist. We also thank the associate editor and all reviewers for their useful feedback in revising this paper. Funding was provided by the Natural Sciences and Engineering Research Council of Canada, University of Ottawa and Carleton University.

**CONFLICT OF INTEREST**

The authors declare no conflicts of interest.

**DATA AVAILABILITY**

Data and scripts to reproduce analyses are available at https://figshare.com/s/f47091e2c7c537846b3e

**ORCID**

Peter Soroye [https://orcid.org/0000-0001-9810-3363](https://orcid.org/0000-0001-9810-3363)

Brandon P. M. Edwards [https://orcid.org/0000-0003-0865-3076](https://orcid.org/0000-0003-0865-3076)

Rachel T. Buxton [https://orcid.org/0000-0002-2772-8435](https://orcid.org/0000-0002-2772-8435)

Jeffrey P. Ethier [https://orcid.org/0000-0001-8285-6440](https://orcid.org/0000-0001-8285-6440)

Albana Berberi [https://orcid.org/0000-0001-6928-1574](https://orcid.org/0000-0001-6928-1574)

Allison D. Binley [https://orcid.org/0000-0001-8790-9935](https://orcid.org/0000-0001-8790-9935)

Steven J. Cooke [https://orcid.org/0000-0002-5407-0659](https://orcid.org/0000-0002-5407-0659)

Joseph R. Bennett [https://orcid.org/0000-0002-3901-9513](https://orcid.org/0000-0002-3901-9513)

**REFERENCES**


Buxton, R. T., Avery-Gomm, S., Lin, H.-Y., Smith, P. A., Cooke, S. J., & Bennett, J. R. (2020). Half of resources in threatened species conservation plans are allocated to research and monitoring. *Nature Communications*, 11, 4668. [https://doi.org/10.1038/s41467-020-18486-6](https://doi.org/10.1038/s41467-020-18486-6)


**SUPPORTING INFORMATION**

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