Biological Conservation 184 (2015) 439–445

Contents lists available at ScienceDirect

Biological Conservation

journal homepage: www.elsevier.com/locate/biocon

Perspective

Linking place-based citizen science with large-scale conservation research: A case study of bird-building collisions and the role of professional scientists

Scott R. Loss[a,*], Sara S. Loss[a,1], Tom Will[b], Peter P. Marra[a]

[a] Migratory Bird Center, Smithsonian Conservation Biology Institute, National Zoological Park, P.O. Box 37012, MRC 5503, Washington, DC 20013, USA
[b] U.S. Fish and Wildlife Service, Division of Migratory Birds, Midwest Regional Office, 5600 American Blvd. West, Suite 990, Bloomington, MN 55437-1458, USA

ARTICLE INFO

Article history:
Received 10 November 2014
Received in revised form 9 February 2015
Accepted 17 February 2015
Available online 10 March 2015

Keywords:
Bird-building collisions
Citizen science
Data-intensive science
Public participation in scientific research
Window collisions

ABSTRACT

A primary benefit of incorporating public participation in scientific research is the increased ability to use data from multiple localities to address conservation research and management objectives that span national, continental, and even global scales. Although the importance of incorporating data from local citizen science programs into large-scale research has been widely recognized, there has been relatively little discussion of specific steps that will facilitate this bridging of scales. We use the example of bird collisions with buildings in North America—an issue for which the majority of data have been collected by citizen science programs that each operate in a different city—to outline simple study design and data collection steps that will ensure that data can contribute to large-scale research syntheses. We also describe how taking a scientific approach to defining research questions and hypotheses at the beginning of a study will: (1) result in a high level of rigor throughout the scientific cycle, most notably at the critical stage when programs formulate study design and data collection protocols, and (2) produce results that effectively inform local policy and management decisions while also contributing to large-scale science. Given the funding and staffing limitations of citizen science programs, we argue that the responsibility is with professional conservation scientists to reach out to programs and provide feedback that assists them in bridging local and large scales. These collaborations will expand the collective contribution of citizens to conservation science and management.

Published by Elsevier Ltd.

Contents

1. Introduction

The incorporation of public participation in scientific research (PPSR; Bonney et al., 2009a) has become widely recognized as an invaluable way to generate scientific knowledge, motivate public engagement with scientific research and advocacy, and provide

---

* Corresponding author at: Department of Natural Resource Ecology and Management, Oklahoma State University, 008C Ag Hall, Stillwater, OK 74078, USA. Tel.: +1 405 744 4607; fax: +1 405 744 3530.

E-mail addresses: scott.loss@okstate.edu (S.R. Loss), sara.loss@okstate.edu (S.S. Loss), tom_will@fws.gov (T. Will), marrap@si.edu (P.P. Marra).

[1] Current address: Department of English, Oklahoma State University, 205 Morrill Hall, Stillwater, OK 74078, USA.

http://dx.doi.org/10.1016/j.biocon.2015.02.023
0006-3207/Published by Elsevier Ltd.

education about scientific subject matter. A common form of PPSR is the citizen science programs that seek to document and conserve biological diversity or to assess ecological responses to anthropogenic stressors (Bonney et al., 2009b; Dickinson et al., 2010, 2012). Citizen scientists are involved in projects as diverse as documenting the abundance and distribution of taxa ranging from earthworms to elephants and investigating anthropogenic global change issues ranging from infectious diseases to invasive species (Citsci.org, 2015; Cornell University, 2015).

A primary scientific benefit of citizen science—along with the numerous societal and educational benefits—is an increased ability to address research questions that span regional, national, continental, and even global scales (hereafter, "large-scale research;" Bonney et al., 2009a,b; Cooper et al., 2007; Newman et al., 2011). Volunteer-contributed observations from individual localities can be combined to provide perspectives that are national or international in scope (e.g., clarifying patterns of bird abundance and phenology with eBird; Sullivan et al., 2009) or that span decades (e.g., estimating bird population trends with U.S. Breeding Bird Survey data; Sauer et al., 2012). The place-based nature of programs that operate in individual localities (hereafter: "local programs") remains an integral component of conservation and ecology research, drawing public interest and participation (Chandler et al., 2012) and leading to solutions to local problems (Miller-Rushing et al., 2012). However, achieving the full potential of citizen science requires collaboration among programs and professional scientists to bridge local data with-large scale research objectives (Bonney et al., 2009a; Couvet et al., 2008; Newman et al., 2011, 2012).

Not all citizen science programs have a scientific approach to data collection at their core. Rather than basing all program activities on explicitly stated and quantitative research questions (e.g., to what degree is there a relationship between an anthropogenic stressor and an ecological variable of interest?"), some programs may work primarily to address education, outreach, or policy-related objectives for which an ecological phenomenon only needs to be roughly documented (e.g., whether the phenomenon associated with an anthropogenic stressor occurs frequently enough to merit policy and/or management concern). Even within similar areas of research, different citizen science programs can have different objectives that lead to study designs and data collection protocols with varying levels of scientific rigor. Non-science related objectives are a critical part of achieving the societal benefits of citizen science, and the methodological variation among citizen science programs understandably arises from programs' local interests, funding constraints, and ecological and socio-political contexts. Nonetheless, adhering closely to the scientific method and having clearly-stated and measurable objectives at the core of all program decisions and activities is crucial for addressing local problems and generating broad benefits to science, program participants, and socio-ecological systems (Shirk et al., 2012). Additionally, methodological variation among programs complicates large-scale analyses that require rigorously collected data and a standardization of approaches (Dickinson et al., 2012; Kelling et al., 2009; Parfitt, 2013; Silvertown, 2009).

Although the importance of linking data from local citizen science programs with large-scale research has been widely acknowledged (Bonney et al., 2009a,b; Chandler et al., 2012; Cooper et al., 2007; Newman et al., 2011, 2012; Sullivan et al., 2014), there has been relatively little discussion of the specific steps that facilitate this bridging of scales. Our goal is to illustrate how a consideration of scales beyond the scope of local programs will enhance contributions to large-scale conservation research while also benefiting efforts to address local conservation policy and management issues. We provide an example for a conservation issue—bird collisions with buildings in North America—for which the majority of data have been collected by several local programs that have a variety of objectives. Based on our experience with this data and our interactions with program staff, we outline simple and low-cost study design and data collection steps that citizen science programs can take to produce data that is better able to address both local and large-scale research objectives. We also discuss how programs should consider explicitly stating quantitative research questions and/or hypotheses that can guide all subsequent steps of study design and data collection. Given funding and staffing limitations faced by most citizen science programs, we argue that the responsibility is with professional conservation scientists to seek out programs in their fields of expertise and to contribute to bridging local and large scales. Notably, although the steps and approaches that we describe here are drawn from our specific experience with bird-building collisions, the recommendations generally apply to any field of conservation or ecology research with a strong representation from citizen scientists.

2. Citizen scientist monitoring of bird-building collisions

Birds are killed by a variety of direct human-caused mortality sources, including collisions with man-made structures and vehicles, predation by free-roaming pets, and poisoning by toxins (Calvert et al., 2013; Loss et al., 2012). However, there has been little research to estimate the amount of mortality caused by different threats, to assess factors governing spatiotemporal variation in mortality, or to quantify patterns of species-specific vulnerability. Therefore, the U.S. Fish and Wildlife Service and Smithsonian Conservation Biology Institute initiated a study that had the following specific objectives: (1) generation of data-driven estimates of annual U.S. mortality, (2) quantification of species-specific vulnerability, and (3) systematic assessment of major outstanding research needs.

As part of this study, we conducted a systematic review to estimate annual U.S. bird mortality from building collisions (Loss et al., 2014a). During our search to locate building collision data, we queried databases of peer-reviewed literature (e.g., Google Scholar and Web of Science), as well as broader search engines (e.g., Google). During these initial searches, we incidentally noted that a large proportion of data, especially for buildings in major metropolitan areas, had been collected by numerous volunteer-driven citizen science programs (Table 1) and had not been published in the peer-reviewed literature. When we became aware of these programs, we adapted our search strategy beyond traditional databases of peer-reviewed studies to also search for websites of citizen science programs that study bird-building collisions (Loss et al., 2014a). Once we were in contact with program coordinators, we asked if they knew of similar North American monitoring programs. Through this informal network of contacts we were made aware of a comprehensive list of U.S. building collision monitoring programs (C. Sheppard, *personal communication*).

After contacting program coordinators, we observed that local programs varied greatly with regards to their primary objectives and study design and data collection methods. Objectives included at least one of the following: (1) rehabilitation and release of birds that experience sub-lethal effects of collisions, (2) documentation that mortality occurs frequently enough to merit policy intervention (without quantifying a specific mortality amount), (3) quantification of the amount of mortality occurring in the study area, (4) documentation of which species are most frequently killed, and (5) identification of building characteristics that cause disproportionately high mortality. The degree to which local programs collaborated with professional scientists from agencies or academic institutions was also variable and covered the entire range of previously defined citizen science project types, including

Table 1
Citizen science programs in the United States and Canada that provided data for the analysis of bird-building collisions in the U.S. (Loss, 2014a).

Project name	Location	Amount of data	
		Years	Number of fatality records
Bird Safe Portland	Portland, Oregon	3	89
Chicago Bird Collision Monitors	Chicago, Illinois	11	21,549
Calgary Fatal Light Awareness Program	Calgary, Alberta (Canada)	3	388
Fatal Light Awareness Program – Canada	Toronto, Ontario (Canada)	11	25,136
Lights Out Baltimore	Baltimore, Maryland	5	1276
Lights Out Columbus	Columbus, Ohio	1	28
Lights Out DC	Washington, District of Columbia	3	316
Lights Out Indy	Indianapolis, Indiana	2	158
Lights Out Winston-Salem	Winston-Salem, North Carolina	2	107
Pennsylvania Audubon	Philadelphia, Pennsylvania	4	527
Project Birdsafe Minnesota	Minneapolis/St. Paul, Minnesota	4	2129
Project Safe Flight New York	New York, New York	3	387
Wisconsin Night Guardians	Milwaukee, Wisconsin	5	419

contributory projects, collaborative projects, and co-created projects (Bonney et al., 2009a), as well projects developed and carried out without any input from professional scientists (collegial contributions; Shirk et al., 2012).

Despite allowing rough comparisons of mortality between buildings and species, data collection protocols for many local programs were not designed to generate unbiased local estimates of mortality, to assess factors governing spatiotemporal variation in mortality, or to provide unbiased estimates of species vulnerability. We were still able to incorporate much of the data into our analyses (including ~92,000 fatality records from 12 local programs) after implementing data inclusion criteria, applying statistical adjustments, and qualifying conclusions based on remaining data limitations (Loss et al., 2014a). However, for the U.S. mortality estimate, we excluded >22,000 records for lacking information about sampling effort (e.g., number of buildings surveyed and/or person-hours of sampling) or other limitations. These methodological limitations also prevented us from comparing mortality rates among programs and from identifying mortality rate correlates across sites. Below, we discuss how additional large-scale analyses could have been facilitated by a few simple study design and data collection steps and by implementing all program activities in the context of one or more explicitly stated and quantitative research questions.

3. Steps that facilitate bridging of local and large scales

Foreseeing the future uses of ecological data can be difficult. For example, data from the Christmas Bird Count (Bonney, 2007), a national citizen science program, have been mined extensively to address questions beyond the programs' original population monitoring objectives. Likewise, data from professional scientists is often sought for large-scale systematic reviews and meta-analyses that go beyond the researchers' original objectives. Because local citizen science programs may also be unable to foresee all potential uses for the data they collect, the data may require extensive screening and management prior to use in large-scale data syntheses. Based on our experience working with diverse data types for the synthesis of bird-building collision mortality, we have identified several straightforward steps that can be taken to reduce the need for data screening and qualification, and therefore, to increase the utility of locally collected data for large-scale research. These steps have previously been introduced in the context of bird-building collisions (Loss et al., 2014a,b); however, here we provide a general discussion of these and additional steps that will apply across many types of citizen science programs and many areas of conservation and ecology research. The steps fall into two categories: (1) study design steps, and (2) data recording and management steps. Examples of large-scale objectives that can be addressed more easily as a result of these steps are in Table 2.

3.1. Study design steps

Study design steps include two related recommendations, the randomized selection of sampling units (i.e., study sites, sampling points, sampling periods) and the avoidance of sampling only at times and locations known to experience a phenomenon (e.g., a harmful effect or cause-and-effect relationship). Randomization is important for generating data that is more representative of the location and time period studied, less biased to particular locations and time periods, and more useful for extrapolation across space and time. Depending on program objectives, sampling can be completely random (i.e., random selection of sampling units from all potential units, such as selection of buildings without regard to height) or based on a stratified random approach (i.e., random selection of sampling units within groups, such as selection of buildings from several height classes). Completely randomized sampling along with detailed annotation of sampling units (Section 3.2) can allow researchers addressing large scale questions to conduct post hoc stratification of sampling units. A focus of sampling on units already known to experience a phenomenon may be necessary to address some types of objectives (e.g., identifying "hotspots" at which to focus management attention for an anthropogenic stressor, Loss et al., 2014c,d). However; even in these cases, inclusion of a subset of randomly selected sampling units will ensure that a portion of the data can contribute to large-scale questions that require unbiased data. This randomization approach will also ensure that results are unbiased with regard to questions about the severity of effects of ecological stressors.

Despite the critical importance of random sampling, in practice, several complications may limit the ability of local programs to implement randomized sampling schemes. First, random sampling may be more likely to result in "negative data" (i.e., zero counts). Surveys with negative data may be viewed as unsuccessful by program participants, and participants may under-report negative data or reduce project participation if a large proportion of surveys return negative data. Approaches to ensure that all data are equally reported and to maintain participation rates have been reviewed in detail (e.g., Dickinson et al., 2012; Rotman et al., 2012), and example approaches include: (1) dissemination of information about the importance of negative data for scientific inquiry (e.g., Bonney et al., 2009b), and (2) incentives for participants to sample in locations and time periods that are traditionally under-sampled or that tend to have many surveys with negative data (e.g., the "challenges of the month" used by eBird; Audubon and Cornell Lab of Ornithology, 2015). Second, some local programs may lack the necessary expertise and technological resources (e.g., geographic information systems) to design and execute rigorous randomized sampling. In these cases, increased collaboration between professional scientists and local programs (Section 5) has the potential to not only provide programs with specific needed products (e.g., randomized sampling schemes designed by professional scientists) but to also facilitate ongoing transfer of expertise and technology that will allow programs to more independently design sampling schemes. Third, some random points are likely to be inaccessible

Table 2
Recommend steps to enhance the utility of locally collected citizen science data for large-scale conservation research and management.

Recommended steps	Example(s) from building collision study	General benefits of implementing step
Study design steps		
Define quantitative research questions and hypotheses instead of/in addition to non-quantitative objectives	How many birds are killed annually at each building? What factors cause particularly high collision rates (as opposed to rough documentation that collisions occur frequently)	Leads to more scientifically rigorous approach to all activities, which should result in less biased and more accurate data that is more useful for both local and large-scale applications
Randomly select data sampling units; do not sample solely at locations known to experience a phenomenon	Randomly select buildings; do not focus only on known "problem" buildings that kill many birds	Data less-biased to particular locations and periods, more representative of study area, and useful for extrapolation to large scales
Take steps to encourage program participants to sample at selected random locations	For any field of conservation/ecology, e.g., disseminate information about scientific benefit of random sampling; provide incentives to sample at under-sampled/random locations	Prevents perception that negative data (zero counts) are less valuable, thus should prevent under-reporting of negative data (under-reporting causes data sets to be biased and less useful for both local and large-scale research)
Data recording/management steps		
Record number of surveyors and hours spent by each surveyor on each survey	Person-hours of effort expended on each building collision survey	Allows correction of response variable raw data by amount of effort expended making data more comparable among locations and time periods; also allows more accurate assessment of response variable correlates
Record effort spent on "negative" surveys (i.e., zero counts and/or surveys where unit of interest not recorded)	Person-hours of effort expended on building surveys that result in no dead birds being found	Allows correction of response variable raw data by amount of effort expended making data more comparable among locations and time periods; also allows more accurate assessment of response variable correlates
Record incidental records separately from those found on formal surveys	Counts of dead birds found on formal surveys recorded separately from counts of birds found incidentally by surveyors and building personnel	Incidental data cannot be standardized by effort; unstandardized data contributes bias to comparisons of response variable classes that do not have equal survey effort and to assessment of response variable correlates
Record number of units sampled	Number of buildings monitored in each survey, each year, etc	Information allows another way to account for sampling effort; allows analyses that require response variable to be in per unit format (e.g., collision mortality per building)
Record detailed description of sample location	Street address; GPS coordinates	Allows researchers to return to sampling location or remotely access it (e.g., using GIS)
Use data entry/management interfaces for data submission	Google Forms submission interface and data management spreadsheet developed for programs that study bird-building collisions (http://tinyurl.com/m3bvxel)	Increases ease of data entry and management for program staff; increases standardization of data within and among programs
Deposit data in online repositories	For any field of conservation/ecology, e.g., Knowledge Network for Biocomplexity (https://knb.ecoinformatics.org/) and Global Biodiversity Information Facility (http://www.gbif.org/)	Ensures open access to data in perpetuity

to local programs as a result of the points falling on private property or far from roads or trails. If project objectives do not require truly randomized sampling across all locations, then pre-stratification of potential sampling areas could exclude these inaccessible locations. If truly randomized sampling is necessary, then programs could select a larger number of points than needed, thus ensuring a sufficient sub-sample of accessible points.

3.2. Data recording and management steps

Data recording and management steps relate largely to the issue of documenting sampling effort and sampling locations. Varying levels of effort can strongly influence results when analyzing data collected under other objectives (Lepage and Francis, 2002; Miller-Rushing et al., 2008, 2012). To minimize this problem, programs should, at minimum, record the number of person-hours expended on each survey, including surveys with negative data. For many studies—such as those that document occurrence or density of rare plant species—the amount of area sampled is an important component of effort that must also be carefully tracked and reported. In addition, just as data collected from non-randomly selected sampling locations may be useful for achieving some objectives, data collected incidentally to formal survey periods may also have uses. However, because incidental data cannot be tied to an amount of sampling time or sampling area, it should be recorded and managed separately so that bias is not added to analyses that require effort-standardized data.

Particular care should be given to clearly documenting locations of sampling units. Clear annotation of locations allows for project participants and other researchers to return to sites at a later date to collect additional data. Annotation of locations also allows sampling data to be remotely linked to a large amount of ancillary data (e.g., landscape characteristics around buildings calculated using geographical information systems) that could be used to conduct analyses beyond the original program objectives. In our study of bird-building collisions, some programs provided detailed street address information for the buildings sampled, and in these cases, we could use Google Maps to confirm the total number of buildings sampled. However, for programs that did not provide this detailed location information, we had to define a range of uncertainty for the number of buildings sampled, and this contributed additional uncertainty to our mortality estimates.

For some local programs, data curation (i.e., data management activities related to long-term preservation and re-use of data) may be insufficient or inconsistent as a result of limited funding and rapid turnover of personnel. Over time, limited attention to formal data curation and a reliance on informally transferred institutional knowledge can result in a loss of the data itself, or equally important, in a loss of the meta-data that is crucial to understanding and using the data. To ensure that data can be useful at a later date—both for program-specific objectives and for analyses by external researchers—citizen science programs should, at minimum, consider entering and managing data through open-access web interfaces that connect to a centralized database hub (e.g., Google Forms; example of a basic interface for bird-building collisions at http://tinyurl.com/m3bvxel). Additionally, to ensure that data is easily accessible in perpetuity and therefore useful for a variety of unforeseen objectives, data and meta-data can be made openly accessible on the internet (Sullivan et al., 2014). Several online data repositories, including the Knowledge Network for

Biocomplexity (https://knb.ecoinformatics.org/) and the Global Biodiversity Information Facility (http://www.gbif.org/), facilitate long-term data curation and open access to data collected by both professional and citizen scientists.

Local citizen science programs face a tradeoff between developing the most rigorous sampling schemes and data collection protocols possible while operating under limitations related to randomized sampling (Section 3.1), funding availability, and participant support. As a result, most programs will likely have to focus on collecting the minimum amount of data that will be useful for scaling up to larger scales. The use of standardized study design, data collection, and data management protocols can aid in optimizing this tradeoff and will also increase the comparability of data among different programs. Newman et al. (2011) provide an approach to increase data standardization with their cyber-infrastructure support tool, CitSci.org, which allows fledgling citizen science programs to use data entry forms that: (1) include a common core of required location and attribute information to ensure a minimum comparability of data, and (2) can be customized according to the needs of individual programs. For data that has already been collected, recent modeling efforts are beginning to allow researchers to account for sampling and data collection biases (e.g., spatiotemporal exploratory models to account for geographic biases, Sullivan et al., 2014; modeling approaches to handle opportunistically collected data, Kery et al., 2010; Snäll et al., 2011).

4. Explicit definition of a quantitative research question

At the heart of the scientific process is the definition of specific and quantitative research questions and associated hypotheses and predictions. Once a question has been clearly defined, all decisions about study design and data collection, management, and analysis should follow relatively intuitively. Defining quantitative research questions is crucial for professional scientists to conduct research worthy of peer-reviewed publication. However, even among professional scientists, research and monitoring does not always adhere to a question-based approach to scientific inquiry (Nichols and Williams, 2006). Just as taking a question-based approach is crucial for professional scientists, the explicit definition of quantitative research questions should be an integral part of formulating science-related objectives for citizen science programs (Bonney et al., 2009a,b; Newman et al., 2012). This approach is necessary for providing solid evidence to convincingly inform effective policy and management interventions for local issues. Additionally, defining quantitative research questions will result in an increased level of rigor in study design and data collection and therefore increased utility of data for large-scale research. Finally, taking an approach that is defined by a clearly stated question should also minimize the perception among program participants that data is being collected solely for the sake of data collection (i.e., with no stated intent for how it will be used, MacKenzie, 2012; Nichols and Williams, 2006; Sharpe and Conrad, 2006).

Explicit statement of a program objective in the form of a quantitative research question (e.g., "how many birds are killed annually at each low-rise and high-rise building in this city?") instead of in the form of a non-measurable and non-quantitative statement (e.g., "documentation of bird-collision mortality to inform policy decisions") leads to a suite of study design and data collection steps formulated to address the question. For example, the above quantitative research question should intuitively lead to a stratified study design with random sampling of both low-rise and high-rise buildings. The question also implies a comparison of mortality rates between building classes. An equivalent comparison between building classes requires that sampling occur at a similar number of units with a similar amount of effort in each class, or at minimum, that the number of units and sampling effort are recorded, thus allowing standardization of the response variable. In the context of the quantitative research question and the need for equivalent comparisons, survey designers may also be led to consider other sources of variation that could bias estimates (e.g., removal of carcasses by scavengers) or to measure potential correlates of the response variable that could confound comparisons between classes (e.g., vegetation surrounding buildings). In summary, we argue that the approach of clearly defining a quantitative research question will have the positive side-effect of leading to an increased rigor of studies due to many of our above-recommended steps naturally being taken.

Variation in the degree to which citizen science programs define quantitative questions was reflected in our building collision study. In general, programs that based monitoring on explicit questions (e.g., How many birds are killed? What factors elevate collision risk?) used rigorous study design and data collection approaches that matched most of our recommendations. These programs often recorded detailed schedules of survey effort (including for surveys with no birds found), presented incidentally collected data separately from data found on surveys, and tracked additional information about potential mortality correlates (e.g., building height). Most programs with non-quantitative objectives (e.g., related to rescue and recovery of birds or rough documentation that mortality was occurring) did not collect information beyond the number, species, and date of fatalities. We recognize that these scaled-back efforts are often unavoidable due to funding and staffing limitations. However, defining quantitative research questions and taking the above-recommended study design and data collection steps that do not automatically follow from this question definition is a low-cost way to generate data that is more useful for tackling both local issues and large-scale research questions. In short, these steps ensure that citizen science is indeed based on a scientific approach to data collection.

5. The role of professional scientists

Our call for local programs to take steps to increase the utility of their data for large-scale research and to define quantitative research questions assumes that programs are aware of the potential for inclusion of their data in large-scale analyses. Citizen science programs may be prevented from implementing our recommendations when this awareness is lacking and when information, staffing, and funding are limited. In many cases, therefore, the first step that will be necessary for our specific recommendations to be followed is for professional scientists to seek out citizen science programs in their areas of expertise. Professional scientists can then provide programs with information about the potential large-scale uses for their data, assist programs in taking the above-described study design and data collection steps, and encourage programs to address targeted quantitative questions that complement programs' current activities and benefit large-scale analyses (e.g., experimental assessment of cause-and-effect relationships, targeted collection of data to validate large-scale patterns; Bonney et al., 2009b; Dickinson et al., 2010, 2012). In our study, we incidentally learned of the numerous citizen science programs that study bird-building collisions while conducting a more formalized search of the peer-reviewed literature. Following up with program coordinators led us to additional projects that lacked a web presence. Researchers that seek to use this type of valuable data resource in the future can use our experience as a template, or they can locate collaborators by referencing websites that provide comprehensive lists of citizen science programs in multiple areas of conservation and ecology research (e.g., Citizen Science Central, Cornell University, 2015; CitSci.org, 2015).

Collaboration between professional and citizen scientists could also lead iteratively to incremental advancement from incidental data collection to the construction of more statistically rigorous sampling and data collection schemes focused on specific research questions (Crall et al., 2010; Stohlgren and Schnase, 2006). Scientists can seek out programs on an as-needed basis (e.g., when addressing a particular research question that requires additional data from multiple localities), or ideally, in advance of such specific needs, thus developing a collaborative network that can be drawn on as needed. In either case, professional scientists should maintain two-way communications with programs and regularly provide useful feedback (Bonter and Cooper, 2012; Chandler et al., 2012; Cooper et al., 2007). For example, with experience gained from working with citizen scientist-collected data and interacting with program staff, professional scientists can distribute fact sheets that outline best practices in a clear and concise format (e.g., Loss et al., 2014b) and develop reporting systems that increase ease of data management within programs and allow seamless merging of data across programs (see examples in Section 3.2).

A major component of using data collected by citizen science programs is the need to verify data quality (Bonter and Cooper, 2012; Crall et al., 2010; Dickinson et al., 2010; Sullivan et al., 2014). Many major problems that limit the large-scale utility of locally collected data can be prevented in advance by considering the potential future uses of the data and by developing data quality control procedures within individual programs. These steps will be most effective when conducted with insight from both professional scientists and local program staff. Data that has received internal program scrutiny may still require post-processing and verification prior to use in large-scale analyses. For example, upon receipt and first review of the data sets in our study of bird-building collisions, we noted several issues that required follow-up contact with project coordinators. Specifically, we requested further information about: (1) availability of additional detailed information about records (e.g., dates of fatality records), (2) the program's data management approach (e.g., whether incidental fatalities were grouped with those found during surveys), (3) study design (e.g., whether buildings were selected randomly), and (4) data collection protocols (e.g., numbers of surveys and surveyors for different time periods). Even when issues such as these have been clarified, additional time-consuming steps for validating data from citizen science programs may still remain. For example, verifying the number of buildings sampled often required us to cross-check street addresses with satellite images in Google Maps. Despite the numerous data quality and verification issues that need to be considered, professional scientists should not be dissuaded from using citizen scientist-collected data and collaborating with citizen scientists. Indeed, the involvement of professional scientists with the development of citizen science program objectives and protocols should reduce the future need for time-consuming data verification efforts prior to large-scale analyses.

Increased communication between professional scientists and citizen scientists is an effective way to increase public engagement and buy-in for research, to expand the scope of ecological inquiry, and to improve the quality of research results (Bonter and Cooper, 2012; Cooper et al., 2007; Dickinson et al., 2010; Reynolds and Lowman, 2013). Of particular note is the need for scientists to take a positive and encouraging approach to communicating with staff in citizen science programs. Invariably, the program coordinators, data managers, and project participants that we interacted with in our study were excited at the prospect of contributing data to a broader research project with large-scale ramifications. Our recommendations for study design and data collection were well-received and enthusiastically adopted whenever logistically possible. By respectfully and genuinely communicating how useful a program already is, while also highlighting simple steps that will allow programs to make an even greater contribution, professional scientists will ensure that citizen scientists retain this enthusiasm.

6. Conclusion

Layering large-scale, question-driven research onto locally collected data sets will become increasingly important in an era of reduced research funding, increased number and complexity of environmental problems, and increased need for large-scale data sets to address these problems. Our experience in locating a large sample of data from citizen science programs that have a variety of objectives, study designs, and data collection protocols—and that are characterized by varying adherence to the use of scientific research questions and hypotheses—has allowed us to identify simple study design and data collection steps that will facilitate an effective bridging of local and large-scale research. Opportunities to bridge local data to large-scale inquiry are likely to exist in any field of conservation and ecology research characterized by a strong representation from citizen scientists. There will continue to be a need to carefully verify and filter citizen scientist-collected data prior to use in large-scale analyses. However, the amount of effort needed to process data will be substantially reduced if programs base all activities on quantitative research questions and follow the steps outlined here. The responsibility lies with professional conservation scientists—including scientists from academic institutions, government agencies, and non-government organizations—to enhance the large-scale utility of programs' locally collected data. This collaboration between professional and citizen scientists will generate actionable results that inform local policy and management and enhance the collective contribution of citizens to conservation science across multiple scales of inquiry.

Acknowledgments

We thank the building collision monitoring programs that contributed data, responded to inquiries, and provided feedback on analyses and suggested best practices. We specifically thank K. Brand (Lights Out Winston-Salem, Forsyth County Audubon Society & Audubon North Carolina), A. Duren (Lights Out Columbus, Ohio Bird Conservation Initiative & Grange Insurance Audubon Center), M. Coolidge (Bird Safe Portland, Audubon Society of Portland), S. Diehl and C. Sharlow-Schaefer (Wisconsin Night Guardians, Wisconsin Humane Society), D. Collister (Calgary Fatal Light Awareness Program, Calgary Bird Banding Society), J. Eckles, K. Nichols, and R. Zink (Project Bird Safe Minnesota, Audubon Minnesota & University of Minnesota), S. Elbin and A. Palmer (Project Safe Flight New York, New York Audubon), D. Gorney (Lights Out Indy, Amos W. Butler Audubon Society), A. Lewis and L. Fuisz (Lights Out DC, City Wildlife), M. Mesure (Fatal Light Awareness Program – Canada), W. Olson (Lights Out Baltimore, Baltimore Bird Club), A. Prince (Chicago Bird Collision Monitors, Chicago Audubon Society), and K. Russell (Audubon Pennsylvania). C. Sheppard provided a list of building collision monitoring programs. J. Rutter and R. Schneider provided assistance with data validation and analysis, and S. Hager, D. Klem, C. Machtans, T. O'Connell, and C. Wedeles provided perspectives and insight on bird-building collisions. S.R.L. was initially supported by a postdoctoral fellowship funded by the U.S. Fish and Wildlife Service through the Smithsonian Institution's Postdoctoral Fellowship program.

Role of the funding source: A biologist from the funding agency (U.S. Fish and Wildlife Service) co-authored this study; however, the U.S. Fish and Wildlife Service as an agency had no role in study design; collection, analysis, and interpretation of data; in writing the report; and in the decision to submit the paper for publication.

References

Audubon and Cornell Lab of Ornithology, 2015. Zeiss eBirder of the month: November & December. <http://ebird.org/content/ebird/news/zeiss_20141112/> (accessed 1.09.15).

Bonney, R., 2007. Citizen Science at the Cornell Lab of Ornithology. In: Yager, R.E., Falk, J.H. (Eds.), Exemplary Science in Informal Education Settings: Standards-based Success Stories. NSTA Press, Arlington, Virginia, USA, pp. 213–229.

Bonney, R., Ballard, H., Jordan, R., McCallie, E., Phillips, T., Shirk, J., Wilderman, C.C., 2009a. Public Participation in Scientific Research: Defining the Field and Assessing Its Potential for Informal Science Education. CAISE Inquiry Group Report. Center for Advancement of Informal Science Education, Washington, D.C.

Bonney, R., Cooper, C.B., Dickinson, J., Kelling, S., Phillips, T., Rosenberg, K.V., Shirk, J., 2009b. Citizen science: a developing tool for expanding science knowledge and science literacy. Bioscience 59, 977–984.

Bonter, D.N., Cooper, C.B., 2012. Data validation in citizen science: a case study from Project FeederWatch. Front. Ecol. Environ. 10, 305–307.

Calvert, A.M., Bishop, C.A., Elliot, R.D., Krebs, E.A., Kydd, T.M., Machtans, C.S., Robertson, G.J., 2013. A synthesis of human-related mortality in Canada. Avian Conservat. Ecol. 8, 11.

Chandler, M., Bebber, D.P., Castro, S., Lowman, M.D., Muoria, P., Oguge, N., Rubenstein, D.I., 2012. International citizen science: making the local global. Front. Ecol. Environ. 10, 328–331.

CitSci.org., 2015. CitSci.org Project List. CitSci.org, Fort Collins, Colorado. <http://www.citsci.org/cwis438/Browse/Project/Project_List.php?WebSiteID=7> (accessed 1.09.15).

Cooper, C.B., Dickinson, J., Phillips, T., Bonney, R., 2007. Citizen science as a tool for conservation in residential ecosystems. Ecol. Soc. 12, 11.

Cornell University, 2015. Finding Projects – Citizen Science Central. Cornell University – Cornell Laboratory of Ornithology, Ithaca, New York. <http://www.birds.cornell.edu/citscitoolkit/projects/find> (accessed 1.09.14).

Couvet, D., Jiguet, F., Julliard, R., Levrel, H., Teyssedre, A., 2008. Enhancing citizen contributions to biodiversity science and public policy. Interdisc. Sci. Rev. 33, 95–102.

Crall, A.W., Newman, G.J., Jarnevich, C.S., Stohlgren, T.J., Waller, D.M., Graham, J., 2010. Improving and integrating data on invasive species collected by citizen scientists. Biol. Invasions 12, 3419–3428.

Dickinson, J.L., Shirk, J., Bonter, D., Bonney, R., Crain, R.L., Martin, J., Purcell, K., 2012. The current state of citizen science as a tool for ecological research and public engagement. Front. Ecol. Environ. 10, 291–297.

Dickinson, J.L., Zuckerberg, B., Bonter, D.N., 2010. Citizen science as an ecological research tool: challenges and benefits. Annu. Rev. Ecol. Evol. Syst. 41, 149–172.

Kelling, S., Hochachka, W.M., Fink, D., Riedewald, M., Caruana, R., Ballard, G., Hooker, G., 2009. Data-intensive science: a new paradigm for biodiversity studies. Bioscience 59, 613–620.

Kery, M., Royle, J.A., Schmid, H., Schaub, M., Volet, B., Haefliger, G., Zbinden, N., 2010. Site-occupancy distribution modeling to correct population-trend estimates derived from opportunistic observations. Conserv. Biol. 24, 1388–1397.

Lepage, D., Francis, C.M., 2002. Do feeder counts reliably indicate bird population changes? 21 years of winter bird counts in Ontario, Canada. Condor 104, 255–270.

Loss, S.R., Will, T., Loss, S.S., Marra, P.P., 2014a. Bird-building collisions in the United States: estimates of annual mortality and species vulnerability. Condor: Ornithol. Appl. 116, 8–23.

Loss, S.R., Will, T., Loss, S.S., Marra, P.P., 2014b. Best practices for data collection in studies of bird-window collisions. <https://www.dropbox.com/s/smd4gxatdgr1wm3/Best%20Practices%20for%20Window%20Data%20FINAL%20eVersion.docx?dl=0> (accessed 1.09.14).

Loss, S.R., Will, T., Marra, P.P., 2014c. Estimation of annual bird mortality from vehicle collisions on roads in the United States. J. Wildlife Manage. 78, 763–771.

Loss, S.R., Will, T., Marra, P.P., 2014d. Refining estimates of bird collision and electrocution mortality at power lines in the United States. PLoS ONE 9, e101565.

Loss, S.R., Will, T., Marra, P.P., 2012. Direct human-caused mortality of birds: improving quantification of magnitude and assessment of population impacts. Front. Ecol. Environ. 10, 357–364.

MacKenzie, D.I., 2012. Study design and analysis options for demographic and species occurrence dynamics. In: Gitzen, R.A., Millspaugh, J.J., Cooper, A.B., Licht, D.S. (Eds.), Design and Analysis of Long-Term Ecological Monitoring Studies. Cambridge University Press, Cambridge, United Kingdom.

Miller-Rushing, A.J., Inouye, D.W., Primack, R.B., 2008. How well do first flowering dates measure plant responses to climate change? The effects of population size and sampling frequency. J. Ecol. 96, 1289–1296.

Miller-Rushing, A., Primack, R., Bonney, R., 2012. The history of public participation in ecological research. Front. Ecol. Environ. 10, 285–290.

Newman, G., Graham, J., Crall, A., Laituri, M., 2011. The art and science of multi-scale citizen science support. Ecol. Inform. 6, 217–227.

Newman, G., Wiggins, A., Crall, A., Graham, E., Newman, S., Crowston, K., 2012. The future of citizen science: emerging technologies and shifting paradigms. Front. Ecol. Environ. 10, 298–304.

Nichols, J.D., Williams, B.K., 2006. Monitoring for conservation. Trends Ecol. Evol. 21, 668–673.

Parfitt, I., 2013. Citizen science in conservation biology: best practices in the geoweb era. PhD Dissertation. University of British Columbia, Vancouver, British Columbia, Canada.

Reynolds, J.A., Lowman, M.D., 2013. Promoting ecoliteracy through research service-learning and citizen science. Front. Ecol. Environ. 11, 565–566.

Rotman, D., Preece, J., Hammock, J., Procita, K., Hansen, D., Parr, C., Lewis, D., Jacobs, D., 2012. Dynamic changes in motivation in collaborative citizen-science projects. In: Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work. ACM.

Sauer, J.R., Hines, J.E., Fallon, J.E., Pardieck, K.L., Ziolkowski Jr., D.J., Link, W.A., 2012. The North American Breeding Bird Survey, Results and Analysis 1966–2011, Version 07.03.2013. United States Geological Survey, Patuxent Wildlife Research Center, Laurel, Maryland.

Sharpe, A., Conrad, C., 2006. Community based ecological monitoring in Nova Scotia: challenges and opportunities. Environ. Monit. Assess. 113, 395–409.

Shirk, J.L., Ballard, H.L., Wilderman, C.C., Phillips, T., Wiggins, A., Jordan, R., McCallie, E., Minarchek, M., Lewenstein, B.V., Krasny, M.E., Bonney, R., 2012. Public participation in scientific research: a framework for deliberate design. Ecol. Soc. 17 (2), 29.

Silvertown, J., 2009. A new dawn for citizen science. Trends Ecol. Evol. 24, 467–471.

Snäll, T., Kindvall, O., Nilsson, J., Part, T., 2011. Evaluating citizen-based presence data for bird monitoring. Biol. Conserv. 144, 804–810.

Stohlgren, T.J., Schnase, J.L., 2006. Risk analysis for biological hazards: what we need to know about invasive species. Risk Anal. 26, 163–173.

Sullivan, B.L., Aycrigg, J.L., Barry, J.H., Bonney, R.E., Bruns, N., Cooper, C.B., Damoulas, T., Dhondt, A.A., Dietterich, T., Farnsworth, A., Fink, D., Fitzpatrick, J.W., Fredericks, T., Gerbracht, J., Gomes, C., Hochachka, W.M., Iliff, M.J., Lagoze, C., La Sorte, F.A., Merrifield, M., Morris, W., Phillips, T.B., Reynolds, M., Rodewald, A.D., Rosenberg, K.V., Trautmann, N.M., Wiggins, A., Winkler, D.W., Wong, W.-K., Wood, C.L., Yu, J., Kelling, S., 2014. The eBird enterprise: an integrated approach to development and application of citizen science. Biol. Conserv. 169, 31–40.

Sullivan, B.L., Wood, C.L., Iliff, M.J., Bonney, R.E., Fink, D., Kelling, S., 2009. eBird: a citizen-based bird observation network in the biological sciences. Biol. Conserv. 142, 2282–2292.