Biological Conservation xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Biological Conservation

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

Lessons from citizen science: Assessing volunteer-collected plant phenology data with Mountain Watch

Caitlin McDonough MacKenzie [a,*], Georgia Murray [b], Richard Primack [a], Doug Weihrauch [b]

[a] Boston University, Department of Biology, 5 Cummington Street, Boston, MA 02215, USA
[b] Research Department, Appalachian Mountain Club, 361 Route 16, Gorham, NH 03581, USA

ARTICLE INFO

Article history:
Received 9 December 2015
Received in revised form 11 July 2016
Accepted 22 July 2016
Available online xxxx

Keywords:
Citizen science
Phenology
Alpine
White Mountain National Forest

ABSTRACT

Citizen science has the potential to expand the scope of data collection, engage the public in research, and answer big scientific questions. But, the quality of volunteer-collected data is often called into question, and citizen science programs must find ways to assess the validity of this concern. Here, we review five years of volunteer-collected data from an alpine flower monitoring citizen science project and present our efforts to investigate the quality of the volunteer-collected data. We found disparity between citizen scientists' self-assessed and actual plant species identification skills, indicating error in either true plant identification or reported location, consequently limiting the use of this dataset. Citizen science programs, including this project, must assess their data, and then make adjustments — in training, data collection methods, or goals — in order to produce quality data consistent with their scientific intentions. Indeed, this project now relies only on observations from seasonal trained staff and a handful of skilled volunteers in light of these findings.

1. Introduction

Citizen science projects with research-oriented goals must develop methods for assessing and improving the quality of their volunteer-collected data. Validating the quality of this volunteer-collected data to uphold the scientific integrity of a project is a common theme among citizen science literature, however no universal rules of data quality have emerged, perhaps because projects vary so much in their scope, scale, and study systems (Bonney et al., 2009; Miller-Rushing et al., 2012). Further, many citizen science programs have education in addition to research goals or have pre-existing audiences with varying skill sets. The high volumes of data from the dispersed data collection model of citizen science can reduce the inherent error in volunteer-collected data (Dickinson et al., 2012), however programs currently engaging in citizen science must still employ a range of Quality Assurance and Quality Control (QA/QC) approaches to fit both the types of data gathered and the audiences that participate. Research-oriented citizen science programs in ecology, climate change biology, or conservation must assess the species identification skills, the field measurements, the qualitative classifications, and the quantitative counts recorded by their citizen scientists. It is important for the citizen science community to share lessons from the fields, the workshops, the classrooms, or the websites where they work to assess and control the quality of volunteer observations. There is a special need for examples of programs that have experienced problems, rather than reporting only on projects that were successful.

Assessing and controlling the quality of volunteer-collected data is often heralded, but practical examples of implementing these measures are missing or folded into larger papers without thorough examination (Cooper et al., 2014; Sullivan et al., 2009). In addition, the largest and most well-known citizen programs have access to resources including infrastructure, experts, and software programming that allow for streamlined QA/QC and adjustments within programs; smaller, local programs often cannot afford these luxuries (Bonter and Cooper, 2012; Wiggins, 2013). A 2010 survey of 128 citizen science programs with a focus on monitoring invasive species — most of which fit this smaller, local category — found that only 39% incorporated quality checks on volunteer-collected data (Crall et al., 2010). Forty percent of the programs in this survey reported that they obtained a majority of their funding from grants; across all types of citizen science programs, short-term funding like this is a common obstacle to efforts to assess volunteer-collected data (Crall et al., 2010).

A recent review of the peer-reviewed literature on the quality of volunteer-collected data in biological monitoring found that most studies assessing citizen science focused on the act of data collection; the most common method reported was comparing volunteers with experts or professionals (Lewandowski and Specht, 2015). In this vein, vegetation surveys have re-sampled permanent transects with professional botanists (Brandon et al., 2003; Galloway et al., 2006), while monitoring programs for pollinators (Kremen et al., 2011), aquatic invertebrates (Delaney et al., 2007), terrestrial invertebrates (Lovell et

* Corresponding author.
E-mail address: mackenz@bu.edu (C. McDonough MacKenzie).

http://dx.doi.org/10.1016/j.biocon.2016.07.027
0006-3207/

al., 2009), and benthic macroinvertebrates (Engel and Voshell, 2002) have compared volunteer observations to data collected by researchers in the same sites. Across these case studies, data comparisons with experts validated the data collection models and improved the associated programs; the volunteer-collected data was rated as high quality, or indistinguishable from the experts, reflecting a "good" citizen science program. However, honest accounts of programs identifying unreliable data and evaluating faults in an underlying data collection model are missing from the literature, and would provide valuable information, especially to smaller, more local citizen science programs with limited resources.

Here, we present a case study of one citizen science program, a project to assess its volunteer-collected data, and the lessons from this QA/QC effort. The Appalachian Mountain Club (AMC), a nonprofit organization dedicated to conservation, education, and recreation in the northeastern United States, launched the Mountain Watch Alpine Flower Watch (Mountain Watch) citizen science program in 2005 to collect long-term alpine plant phenology data in the White Mountain National Forest, New Hampshire. Alpine ecosystems are generally sensitive to changes in climate (Pauli et al., 2014) and phenological timing has implications for the success and long-term persistence of the plants within those systems (Inouye, 2008), but the remote location of alpine habitats makes it a challenging place to obtain observational data with good spatial and temporal resolution.

Mountain Watch solicits hikers to become citizen scientists, and asks volunteers to record flowering phenology observations along the trails in New Hampshire's White Mountains. With this data, the AMC planned to track the local ecological effects of climate change on plant communities in the small and fragmented alpine habitats of New Hampshire. The citizen science program also has core educational goals to engage the hiking community in the issue of climate change through hands-on monitoring. The available audiences were the large number of hikers (~500,000 per year) visiting AMC facilities in the White Mountains and a self-selected group of already-active volunteers. In addition to the volunteer-collected data from Mountain Watch, the AMC has utilized research staff, as well as seasonal naturalists and interns, to record phenology data at permanent plots in the White Mountains since 2005. Both the citizen science project and the research staff observations follow the same monitoring protocol, however only the research staff observations have resulted in a scientific publication to date.

In 2014, the AMC's research department used long-term weather records from the Mount Washington Observatory and alpine plant phenology data gathered by research staff to hindcast flowering phenology and assess late-spring/early-summer frost risks for three of the Mountain Watch plant species (Kimball et al., 2014). The volunteer-collected data from the Mountain Watch program could broaden the geographic scope of this research (from the twelve plots proximate to Mount Washington's meteorological station included in this analysis to alpine habitats across the northeastern United States) and provide long-term phenology data (expanding on the four years of data included in this analysis with on-going citizen science efforts). However, the potential of the Mountain Watch dataset is dependent on its quality. To this end, we looked at the first five years of volunteer-collected Mountain Watch data from the perspective of quality assurance and quality control.

We reviewed the volunteer-collected Mountain Watch data from 2005 to 2009, conducted vegetation surveys at locations recorded by volunteers, and assessed the Mountain Watch data collection model. We used chi-square tests to describe the relationships between species identification rates and characteristics including relative abundance, phenophase, and the volunteers' self-assessed certainty of identification. In the process, we identified two main challenges in QA/QC for the Mountain Watch data: 1) our ability to review the data hinged on the precision of the geographic location descriptions provided by volunteers, and 2) for the majority of volunteers, we did not know their plant identification skills or prior knowledge of the alpine habitat aside from their self-assessed certainty of identification on the datasheet. From our review of five years of volunteer-collected data, we were able to identify potential shortcomings in the original Mountain Watch data collection model, adjust the citizen science program, and share lessons in QA/QC methods for a small, local program with limited resources.

2. Study area

Alpine habitat in the northeastern United States, is limited to ~34 km^2 of fragmented ridges and summits above treeline. The largest of these alpine areas comprises ~11.3 km^2 in the Presidential Range of the White Mountain National Forest, New Hampshire (Kimball and Weihrauch, 2000). The Presidential Range includes New England's highest peak, Mt. Washington (1917 m a.s.l.), three AMC backcountry huts catering to backpackers, and some of the most popular hiking trails in the White Mountain National Forest. This case study focuses on data collected here.

Six common and charismatic alpine plant species were chosen as Mountain Watch target species: ericaceous shrubs *Rhododendron groenlandicum, Vaccinium uliginosum* and *Vaccinium vitis-idaea;* herbaceous *Geum peckii* which is endemic to the White Mountains and Nova Scotia; alpine sedge *Carex bigelowii;* and the circumpolar pin-cushion plant *Diapensia lapponica.* Criteria considered in target species choice included ease of identification, limited look-alike species, ease in phenophase observation, and a variety of life histories and phenological timing. All six are slow growing, long-lived perennials; the plant communities and species composition in the Presidential Range has not changed over the duration of this study.

3. Methods — Mountain Watch program

The AMC Mountain Watch program builds on the popularity of the White Mountain National Forest trail system and recruits hikers to become citizen scientists. The only prerequisites for participating in Mountain Watch are interest, a species identification field guide, and a blank datasheet, which are available online or at any AMC lodge or backcountry hut. Mountain Watch training was provided at backcountry huts as an evening nature program, but the frequency of these programs varied and AMC did not track which volunteers had attended a training program over the years examined. The datasheet asks volunteers to identify the six target alpine plant species, and record the current phenophase (i.e.: before flowering, flowering, or after flowering) for each observation (Fig 1). Volunteers record an observation by checking each phenophase present and circling the dominant phenophase.

Volunteers also rank their certainty of identification (CID) for each species on a scale from 1 (uncertain) to 3 (very certain) on the datasheet. The target species and locations are unmarked to protect the integrity of the National Forest and to encourage data collection across the alpine habitats of the White Mountains, and volunteers are asked to record the geographic location of their observations in an open-ended space on the datasheet (Fig 1). A map of the Presidential Range was printed on the reverse of the datasheet to provide guidance for the observation location. Occasionally, volunteers provided GPS coordinates in this space, but most often they simply wrote a description of their location. During the years examined in this study, cellphone service in the Presidential Range was spotty to nonexistent, and GPS-enabled smartphones had not yet become ubiquitous accessories for hikers (Wiggins, 2013).

4. Methods — Mountain Watch QA/QC

In 2009 we surveyed the vegetation at geographic locations in the Presidential Range recorded by volunteers in an effort to assess the