METHOD

Using iNaturalist in a Coverboard Protocol to Measure Data Quality: Suggestions for Project Design

Julie Wittmann, Derek Girman and Daniel Crocker

We evaluated the accuracy of data records generated by citizen scientist participants using iNaturalist in a coverboard sampling scheme, a common method to detect herpetofauna, involving 17 species of amphibians and reptiles. We trained and observed 10 volunteers working over an eight-month period at a study site in Sonoma County, California, USA. A total number of 1,169 observations were successfully uploaded to the iNaturalist database by volunteers including a new locality for Black Salamander. Volunteers were generally successful in collecting and uploading data for verification but had more difficulties detecting smaller, camouflaged salamanders and photographing quick-moving lizards. Errors associated with uploading smartphone data were reduced with volunteer experience. We evaluated all observations and determined that 82% could be verified to the species level based on the photograph. Forty-five percent of the observations made it to the iNaturalist "research grade" quality using their crowdsourcing tools. Independently we evaluated the crowdsourcing accuracy of the research grade observations and found them to be 100% accurate to the species level. A variety of factors (herpetofauna group, species, photograph quality) played a role in determining whether observations were elevated to research grade. Volunteer screening and training protocols emphasizing smartphones that have appropriate battery and data storage capacity eliminated some issues. Our results suggest that a variety of factors can help scientists and resource managers improve the experiences and quality of the data in citizen science biodiversity programs.

Keywords: Citizen science; iNaturalist; herpetofauna; volunteers; crowdsourcing; data quality; smartphone methodologies

Introduction

Smartphone apps are aiding in conservation of biodiversity and natural resource management by enhancing opportunities for documenting biological occurrences (Kelling 2018; Pimm et al. 2015), including those involving amphibians and reptiles (hereafter herpetofauna). Volunteer efforts have the potential to play a crucial role in the study of changing herpetofauna communities (Cohn 2008). While there are limited apps dedicated specifically to herpetofauna, apps such as iNaturalist allow for the documentation of organisms including herpetofauna and provide data quality assurance strategies (Jacobs 2016).

One technique that is amenable to citizen science (CS) projects that could incorporate the iNaturalist app is the coverboard technique, a common and long-established method used to inventory species of herpetofauna (Stebbins 1954). Coverboards are thin, flat pieces of wood or other hard material (e.g., plywood, corrugated metal) that have a low investment of time and cost to establish and maintain and result in relatively low levels of disturbance to habitat (Hesed 2012; Grant et al. 1992). In comparison to other herpetofauna monitoring methods, the coverboard technique has been considered at least comparable to other techniques (drift fence/pitfall trap arrays, transect searches, and natural cover) in terms of allowing for a similar diversity of species (Bonin and Bachand 1997; Tietje and Vreeland 1997; Harpole and Haas 1999; Houze and Chandler 2002). However, the coverboard technique may have some limitations including the need for coverboards to be made of other materials to target specific herpetofauna groups or species; need for mark-recapture technique in conjunction for certain species to avoid recapture data; skewed size classes for certain species; potential for vandalism or disturbance by livestock; and excessive disturbance involving coverboard lifting by curious naturalists not involved in the project. Nevertheless, the coverboard technique is commonly used by herpetofauna monitoring programs (Lee and Inventory 2012; Price and Dorcas 2011) and requires limited training to implement and monitor (Hesed 2012). Any missed or delayed coverboard checks do not harm organisms unlike many other herpetofauna sampling methods (Tietje and Vreeland 1997). Further, the coverboard technique provides a consistent area to search and document what is underneath and therefore provides a unique opportunity to assess volunteer data quality.

Data quality problems are often the first critique of CS data and have been assessed in numerous studies (Crall et

Sonoma State University, US
Corresponding author: Julie Wittmann (protecthabitat@gmail.com)

al. 2011; Dickinson, Zuckerberg, and Bonter 2010; Kosmala et al. 2016), so assuring data quality is important in the field of CS. The lack of volunteer training and its effect on the reliability or credibility of species identification of animal observation data is an important factor when considering data accuracy (Conrad and Hilchey 2011). A high quantity of data to be processed (i.e., verified) creates a need for methods supporting semi-automatic validation of observation data. It is also important that the specific properties of CS data are considered, because these data are often not produced in a systematic way which may result in, for example, incompleteness (Jacobs 2016). For coverboard studies of herpetofauna, data collection and semi-automatic validation tools need to be evaluated for their efficacy for improving data quality.

We compared herpetofauna occurrence data collected by volunteers using the iNaturalist app, with minimal training and no prior experience, to data collected by herpetologists through all steps of the data collection and upload process over multiple surveys and evaluated the accuracy and effort of the crowdsourcing approach for data quality assurance. Specifically, we measured impacts on data quality by determining (1) whether smartphone technology and/or the app interface affected volunteers' abilities to record data, (2) the level of accuracy of data from volunteers for variables associated with herpetofauna group, abiotic factors, and level of volunteer experience, (3) the efficacy of the crowdsourcing approach for data quality assurance, and (4) whether the crowdsourcing effort was affected by photograph quality, or by herpetofauna group.

Methods

First, we describe the methods for the coverboard sampling of herpetofauna that would be part of any coverboard project and then we describe how we assessed the data quality of the citizen scientists involved in the project.

Coverboard sampling design

A herpetofauna coverboard sampling scheme was established to quantify the diverse herpetofauna found in the Pepperwood and Osborn reserves of the Mayacamas and Sonoma Mountains within the Russian River Watershed of California (**Figure 1**). The coverboards were centered
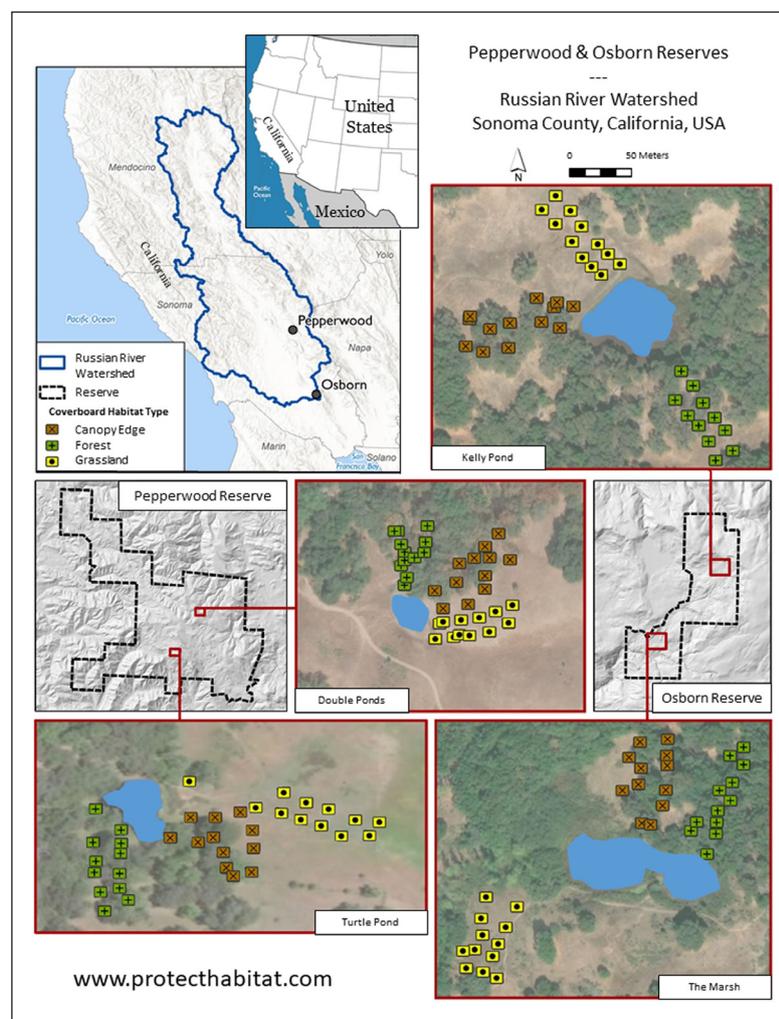

Figure 1: Map of study sites. Pepperwood and Osborn reserves are located in the Russian River Watershed in the San Francisco Bay Area in northern California, USA. Pepperwood reserve: Turtle Pond latitude 38.574918/longitude –122.706904; Double Ponds latitude 38.584573/longitude –122.699525; Osborn reserve: Kelly Pond latitude 38.350616/longitude –122.586928; The Marsh latitude 38.340911/longitude –122.596426. Using: *ArcGIS* [GIS software]. Version 10.6. Redlands, CA: Environmental Systems Research Institute, Inc., 2018.

around two ponds at each reserve which contained water for at least part of the year because herpetofauna diversity is usually greatest in areas with moisture (Brode and Bury 1984). Each pond site contained three coverboard arrays that were placed within estimated average annual water level and spread outward from the pond into grassland, forest canopy edge, and forest habitats (Matlack 1993; Ryan et al. 2002; Block and Morrison 1998). Within each habitat type there were two parallel 25-meter long transects, 15 meters apart, composed of six coverboards placed every five meters (Grant et al.1992; Tietje and Vreeland 1997; Ryan et al. 2002) (6 coverboards per transect, 2 transects per habitat, 3 habitats per pond, 2 ponds per reserve = 72 coverboards per reserve, **Figure 1**). Each 2 by 4 foot coverboard was made of half-inch thick plywood and seasoned in the described environment for three months prior to data collection (Grant et al. 1992; Tietje and Vreeland 1997; Houze and Chandler 2002; Ryan et al. 2002). Duff and vegetation were gently set aside during coverboard installation such that soil and coverboard contact occurred. Coverboards were sampled once every week (Houze and Chandler 2002) through a single wet through dry season (December 2015 through July 2016), thereby allowing for herpetofauna peak abundances (Joppa et al. 2009). To support a structured approach, surveys were conducted at variable times of day to capture variation in environmental conditions.

Weather variables
Precipitation during each survey was recorded as either rain or no rain. Percent cloud cover during each survey was categorized as (1) no or few clouds, (2) moderately cloudy, or (3) very cloudy. The time of day included either the survey half as first half (first pond visited) or second half (second pond visited). Weather stations deployed at the reserves collected a variety of other weather-related data and were obtained at the end of the study to add to appropriate field datasheets. Air temperature and soil moisture data were collected every 15 minutes and averaged for each survey time frame. Both weather stations logged temperature at a height of 1.5 meters.

Volunteer observer training and evaluation of their data
For the purposes of this research, the human subjects who volunteered for the field data collection and upload using the iNaturalist app are referred to as "observers" and the uploaded records that they contributed are referred to as "observations" and consist of images, descriptions, location and time data, and community identifications (Van Horn et al. 2018). Herpetofauna occurrences recorded by herpetologists are referred to as "records." The observers were not informed about their accuracy involving the various survey steps or quantity of surveys conducted (survey experience), therefore, this was a blind study. Individuals who provided data quality assurance through the crowdsourcing approach (i.e., data verification) are referred to as "identifier respondents" and unknowingly contributed to this research. "Crowdsourcing effort" is directly related to the number of observations elevated to "research grade" by identifier respondents through the crowdsourcing

approach (e.g., the greater the number of research grade observations, the greater the crowdsourcing effort). Two formally trained herpetologists evaluated the accuracy of observers and accuracy and effort of the identifier respondents involved in the crowdsourcing approach for data quality assurance using the iNaturalist project named "Herpetofauna Project" (available online http://www.inaturalist.org/projects/herpetofauna-project), which was established for the purposes of this study.

Volunteer observers were sought to use the iNaturalist platform for herpetofauna surveys, each using their personal smartphones and committing to a mandatory three-hour training and five surveys (**Table 1**). The project was advertised to students and both reserve affiliates in 2015 via a variety of sources including email and fliers. The first 10 respondents to the announcement owned smartphones and were accepted into the study. The single training session consisted of (1) a short presentation about the general overview of the project purpose (to document herpetofauna occurrences for conservation and inventory purposes), (2) familiarization of the iNaturalist website using a guided worksheet, (3) time to download the iNaturalist app, and (4) familiarization with the observer survey protocol at a pond site within the reserve (**Table 1**). Finding coverboard locations in the field, lifting coverboards, and capturing organisms were not included in the training or evaluation. Further, a review of common species likely to be found, photograph instructions, or how to use the app and website beyond the use of documenting herpetofauna outlined in **Table 1** were not included in the training.

Using iNaturalist to measure data quality
The evaluation of data quality was conducted for each observation by comparing the results of an observer using the iNaturalist smartphone app with those by a herpetologist using a field datasheet. Only one observer was evaluated at a time to maintain the blind aspects of this study, and one of two herpetologists doubled as a "field partner" for this purpose. The three survey steps for an observation to be classified as an observation with the possibility to become elevated to research grade include that the observer (1) detect herpetofauna specimen, (2) photograph specimen detected, and (3) upload observation datum to the Herpetofauna Project website (**Table 1**). To measure volunteer data accuracy using iNaturalist and the coverboard technique within the reserves, observations were compared to established reserves' list of herpetofauna species observation data.

Photograph quality
Only observations that were able to be identified in the observation photograph(s) to the species level by the herpetologists were included in further analyses. Photograph quality, specifically how much of the specimen was captured in the photograph, was categorized in terms of percent of specimen showing on the first (mostly the only) observation photograph. These categories may not necessarily be associated with the percent of the specimen to the rest of the photograph (e.g., an entire lizard, from head to tail may be shown in a photograph but it may be

Table 1: Summary of the protocols for observers and participating herpetologists.

Protocol Type	Observer	Herpetologist
General	*Prior to training:* No or minimal experience with field surveys. No use of biodiversity smartphone apps. *Prior to first survey:* Minimal training.	*For each survey:* Use of field datasheet to document reserve, pond site, observer, observer quantity of surveys conducted, survey half (time of day), cloud cover, and precipitation.
Each coverboard (one survey = 72 coverboards)	Stand in front of coverboard. Open iNaturalist smartphone app. Inform herpetologist of any (1) technology issues and cause (battery or data storage capacities) and/or (2) app interface issues.	Stand in front of coverboard and lift coverboard up toward individuals (herpetologist and observer).* Time observer for 20 seconds (excluding the time that observer actively used their app to document any herpetofauna specimens).
Detect herpeto-fauna specimen	No use of capture tools. No handling specimens. No moving debris, soil, etc.	Count numbers of specimens for each species. Observe and record specimens detected by observer.
Photograph specimen detected	Use of iNaturalist smartphone app to photograph each specimen observed as a separate observation. Make finest taxonomic guess within app for each observation. Use "project name" field to add to "Herpetofauna Project." Inform herpetologist if any specimen(s) escaped.	Continue to time coverboard lift. Continue to observe observer.
Immediately post coverboard replacement	Summarize number of specimens detected if asked to clarify by herpetologist.	Confidential use of field datasheet to record the quantity and species of herpetofauna specimens. Confidential use of field datasheet to note if specimen(s) (1) were not detected by observer, based on if the observer did not photograph and/or attempt to photograph specimen(s) or (2) escaped during photograph opportunity, based on whether observer attempted but failed to photograph specimen(s) and/or reported any escape(s).
Upload observations	Upload observations after each survey to the iNaturalist Herpetofauna Project using Internet access. Observations become available to iNaturalist identifier respondents for data quality assurance through the crowdsourcing approach. Data become publically available (including downloadable files).	Reminder (email or text message) sent to each observer who did not upload a batch of observations within the first 24 hours of their survey. Three-months post survey, a query on observers' public data for comparison to herpetologists' records. "Data not uploaded" is the difference between the herpetologists' records and the expected observers' observations.

* Exceptions to lifting the coverboard up toward individuals are those coverboards in grassland habitats in the warmer months, which should be lifted away from individuals in the event that a venomous snake may need to safely escape.

taken from a greater distance thereby potentially showing a smaller percent of the specimen to rest of photograph compared to other photographs). Photograph quality categories included (1) 95–100% of specimen showing, (2) 50–94% of specimen showing, and (3) 0–49% of specimen showing. Photograph quality categories below 95% percent of specimen showing were considered to be of lower quality and were due either to (1) cover (i.e., vegetation, soil, or coverboard) or (2) photograph cut off. Other reasons to categorize photograph quality (e.g., zoom, focus, background contrast, number of specimens per observation) were not assessed in this study.

Crowdsourcing approach for data quality assurance
Observations were made available to the online community for three months' post survey season (August 2016 to October 2016). According to the iNaturalist website (2019), three data quality grades that differ on the required data and/or community effort include (from low- to high-quality) (1) casual, (2) needs identification, and (3) research. The accuracy and effort of the crowdsourcing approach for data quality assurance was determined through queries within the Herpetofauna Project. Subspecies were reduced to species for all evaluations in this study. Herpetologists separately reviewed each

research grade observation to determine the accuracy of the verification. Additionally, each of the remaining (needs identification grade) observations were separately reviewed by herpetologists to determine the quantity of verifiable observations based on the photograph (to both herpetofauna group and species). The determination of verifiable observations was based on 100% agreement among participating herpetologists.

Because it was statistically difficult to determine if the observations verified by the crowdsourcing approach were based on the identifier respondents' interest or timing, the 16 top contributors, all with more than 10 identifications, were contacted using iNaturalist's direct message to inquire about their personal motivations to identify images.

Measuring data quality
A simple measure of data quality is to compare the success of observers detecting specimens, photographing specimens, and uploading observations for identification and comparing records made by participating herpetologists.

Statistical analyses measured data quality using the iNaturalist smartphone app. Results were evaluated using generalized mixed linear models with a binary response distribution and logit link function (SAS 9.4, Proc GLIMMIX), and volunteer observer was included as a random effect for these analyses. Following evidence for a significant effect of herpetofauna group, data were visually inspected to identify likely drivers of these effects and subdivided to represent membership in that species or group. Post-hoc tests were then performed to identify the specific species driving probabilities of the response.

To determine whether observations were affected by the quantity of prior surveys conducted (survey experience) by the observer or by herpetofauna group, the variables quantity of surveys conducted and herpetofauna group were modeled as fixed effects on observations. In additional analyses, quantity of surveys conducted and herpetofauna group were modeled as fixed effects on three response variables including specimen not detected, specimen detected but escaped during photograph opportunity, and datum not uploaded to the Herpetofauna Project. Analysis of each of these response variables did not include either of the other two responses, because these three response variable categories were mutually exclusive. To assess the effect of specimen escape due to additional factors such as the presence of herpetofauna group and weather effects, we included an interaction of herpetofauna group and quantity of surveys conducted in the model. We determined whether the escape of specific herpetofauna groups, for which escape during photograph opportunity was significant, were affected by specific herpetofauna species, quantity of surveys conducted, or weather variables. This model assessed effects of survey half (time of day), precipitation, cloud cover, average temperature and soil moisture during the survey, quantity of surveys conducted, and herpetofauna species on the occurrence of specimens that were detected but escaped during photograph opportunity. Finally, we determined if the quantity of surveys conducted affected whether data were uploaded to the Herpetofauna Project.

To assess the effect of increasing quantity of surveys conducted during the span of the study, we included an interaction of herpetofauna group and quantity of surveys conducted in the model. To assess the potential for multicollinearity in the weather variables, correlation between variables were examined and when two variables were highly correlated, one was considered for removal from the model. To access associations between technology issues or crowdsourcing effort and accuracy associations, we used contingency tables evaluated based on likelihood ratio X^2 in JMP Pro 13. Results were considered statistically significant at $p < 0.05$.

Results
Eight of ten observers completed the mandatory training and five surveys each as part of this study. Two observers dropped after completing only two surveys each. The number of species under one coverboard at any one time ranged from zero to four, and the number of specimens ranged from zero to eleven (higher numbers always involved Sierran Tree Frog close to the water edge). The following results are organized in the order of investigation.

Data collection
iNaturalist smartphone application interface and technology
The iNaturalist smartphone app interface was successful for all 100% of attempts of observers to use their smartphone to document herpetofauna (1,169 observations plus 53 "data not uploaded" records). Observers did not report any app interface issues. All of the technology issues were associated with the two observers who dropped out of the study. Although smartphone models of observers varied, the two observers with technology issues both owned older smartphone models. Of the specimens associated with technology issues, significantly more were associated with battery capacity issues (84.1%) than photograph storage ability issues (15.9%) ($X^2_{1,\,245} = 214.77$, $p < 0.0001$).

Survey data quality overview
Observers documented 17 species of herpetofauna between the two reserves (17 out of 27 from the established Pepperwood herpetofauna list; 14 out of 25 from the established Osborn herpetofauna list of observation data) plus a new record for Black Salamander at Pepperwood using the iNaturalist app with the coverboard technique. This resulted in a match of 63.0% plus a new record for Pepperwood and a match of 56.0% for Osborn, averaging a match of 59.5% for combined reserves (=30 different species total on established reserves' herpetofauna lists of observation data). Herpetofauna Project observations included 805 from Pepperwood and 364 from Osborn.

Herpetofauna records (N = 1,549) were significantly affected by herpetofauna group ($F_{3,\,1538} = 17.64$, $p < 0.0001$). The lizard group was the least likely to be uploaded compared to the other three groups (anuran, salamander, snake) ($F_{1,\,1540} = 49.29$, $p < 0.0001$). There was no effect from observer or observer quantity of surveys conducted. Responses associated with observations (specimen not detected, specimen detected but escaped during

photograph opportunity, and datum not uploaded to the Herpetofauna Project) were further investigated. Out of the 1,549 herpetofauna records, 75.5% (=1,169 records) had the potential to become research grade observations, 9.4% (=145 records) were specimens not detected, 11.7% (=182 records) were specimens detected but escaped during photograph opportunity, and 3.4% (=53 records) were not uploaded to the Herpetofauna Project (**Table 2**).

There was no significant effect of observer as a random variable, however, one of the eight volunteers was responsible for 44.1% of the specimen escape during photograph opportunity, another volunteer was responsible for 29.7% of the data not uploaded to the Herpetofauna Project, and another volunteer failed to regularly use the camera zoom function when photographing herpetofauna and several of these observations were not verifiable by either herpetologist. The three responses relating to survey data quality are described below.

Detecting herpetofauna specimens

Although observers detected all species found by the herpetologists, some herpetologists' individual records were not recorded by the observers. Observers missed some data because some individual specimens were not detected, which was significantly affected by herpetofauna group ($F_{3, 1538} = 9.76$, $p < 0.0001$). The salamander group was the least likely to be detected by observers ($F_{1, 1540} = 25.24$, $p < 0.0001$) and the lizard group was the most likely to be detected by observers ($F_{1, 1540} = 19.03$, $p < 0.0001$, **Figure 2**). The salamander group was subdi-vided into three species' groups for statistical purposes: California Slender Salamander, both newt (*Taricha)* species, and all other salamanders (Ensatina, Black Salamander, and Arboreal Salamander). Salamander specimens detected by observers were significantly affected by species ($F_{1, 634} = 7.66$, $p = 0.0058$, **Table 2**). California Slender Salamander was the least likely to be detected, and all other salamanders were more likely to be detected ($F_{1, 634} = 4.67$, $p = 0.0310$).

Photographing specimens

Specimen escape during photograph opportunity was another source of data missed by observers and was significantly affected by herpetofauna group ($F_{3, 1537} = 23.83$, $p < 0.0001$). The lizard group was the most likely to escape ($F_{1, 1539} = 91.63$, $p < 0.0001$), and the salamander group was the least likely to escape ($F_{1, 1539} = 29.72$, $p < 0.0001$, **Figure 2**). Lizards that escaped were significantly affected by herpetofauna species ($F_{2, 458} = 8.95$, $p = 0.0002$, **Table 2**). Western Fence Lizard was the most likely to escape ($F_{1, 459} = 10.30$, $p = 0.0014$) and Southern Alligator Lizard was the least likely to escape ($F_{1, 459} = 17.20$, $p < 0.0001$). Nearly zero escapes occurred from salamanders; therefore, no analyses were necessary for this group. Specimen escape was significantly affected by survey ($F_{1, 1537} = 9.74$, $p = 0.0018$). However, the interaction of group and survey was not significant ($p = 0.1685$). Therefore, we can attribute specimen escape during photograph opportunity due to additional factors such as the presence of lizards and weather effects.

Table 2: Summary by species of herpetofauna specimens not detected by observers and detected but escaped during photograph opportunity. Note that the California Slender Salamander was the least likely to be detected, and the Western Fence Lizard was the most likely to escape. Scientific names adopted from iNaturalist.

Species	No. of records by herpetologists	% not detected by observers	% escaped from observers
Sierran Tree Frog (*Pseudacris sierra*)	329	9.1	6.4[+]
Western Toad (*Anaxyrus boreas*)	6	0.0	0.0
California Slender Salamander (*Batrachoseps attenuatus*)	472	16.1*	0.2
Ensatina (*Ensatina ensatina*)	79	6.3	0.0
Other Salamanders: Arboreal Salamander (*Aneides lugubris*) and Black Salamander (*Aneides flavipunctatus*)	6	0.0	0.0
California Newt (*Taricha torosa*)	76	9.2	1.3
Rough-skinned Newt (*Taricha granulosa*)	10	0.0	0.0
Western Fence Lizard (*Sceleoporus occidentalis*)	264	4.2	39.0*
Western Skink (*Plestiodon skiltonianus*)	136	4.4	28.7
Southern Alligator Lizard (*Elgaria multicarinata*)	69	4.3	4.3
Ringneck Snake (*Diadophis punctatus*)	66	10.6	19.7
Other Snakes: Gopher Snake (*Pituophis catenifer*), Western Rattlesnake (*Crotalus oreganus*), North American Racer (*Coluber constrictor*), California King Snake (*Lampropeltis californiae*), and Common Garter Snake (*Thamnophis sirtalis*)	36	0.0	2.8
All species combined	1,549 (100%)	9.4%	11.7%

* Significant at $p < 0.05$.
+ All escapes were recorded when three or more specimens occurred at the time of the coverboard check.
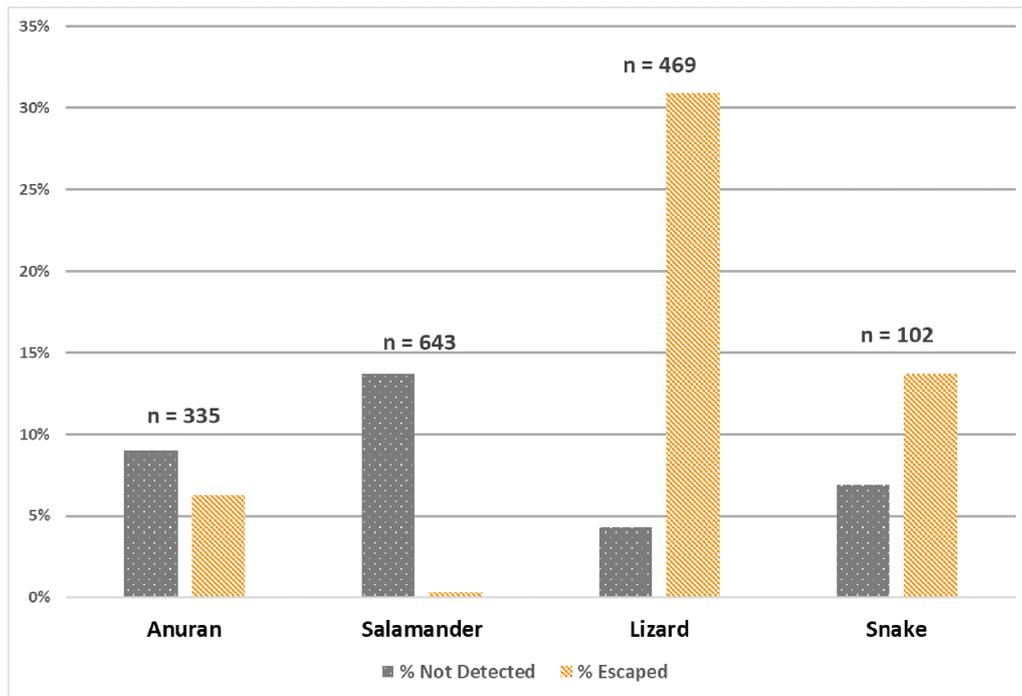
Figure 2: Percent of herpetofauna not detected or escaped by herpetofauna group, N = 1,549. The salamander group was the least likely to be detected by observers and the lizard group was the most likely to escape from observers.

Weather variables and escape

There was collinearity with precipitation and cloud cover (r = 0.70) and precipitation was removed from the model. Of the detected lizards, escape during photograph opportunity decreased with increasing cloud cover ($F_{1, 438}$ = 24.61, p < 0.0001). Lizard escape was also confirmed again to be significantly affected by species ($F_{2, 438}$ = 7.71, p = 0.0005). There was no effect from observer quantity of surveys conducted, soil moisture, temperature, or time of day (survey half).

Uploading observations

As stated, of the 1,549 records, 75.5% (=1,169 observations) had the potential to become research grade observations. Of these, 45.3% (=530 observations) were elevated to research grade through the crowdsourcing approach, and 54.7% (=639 observations) remained in the needs identification grade. Further, 91.2% (=1,066 observations) could be verified to herpetofauna group by the herpetologists based on the photograph and 81.8% (=956 observations) could be verified to herpetofauna species. The observations not identified to the species level by the herpetologists included many low-quality photographs (including those categorized with less than 95% specimen showing), many of which included newt (*Taricha* spp.), California Slender Salamander, and observations with more than one specimen per observation photograph.

Of the 1,549 records that had the potential to become research grade observations, 3.4% (=53 records) were not uploaded to the iNaturalist Herpetofauna Project (0.6% due to smartphone screen reflection issues; 1.4% were uploaded to iNaturalist but not assigned to the Herpetofauna Project; and 1.4% were casual grade with missing data, mostly missing GPS information). These expected observations not uploaded after

photograph opportunity were affected by survey ($F_{1, 1540}$ = 15.81, p < 0.0001) with a decrease in expected observations not uploaded with observer quantity of surveys conducted. Herpetofauna group had no significant effect (p = 0.0795). When the group and survey interaction term was included, it was found to be significant (p = 0.0205), therefore the increase in expected uploaded data through surveys cannot be attributed to other factors such as the weather effects from the changing season.

Crowdsourcing approach for data quality assurance

All research grade observations uploaded to the Herpetofauna Project and verified by identifier respondents through the crowdsourcing approach (=530 observations) were accurate. Identifier respondent preferences are described below.

Identifier respondent preferences for herpetofauna group

Observations that had the potential to become elevated to research grade through the crowdsourcing approach (and verifiable by herpetologists to the coarser taxonomic level of herpetofauna group, =1,066 observations) were affected by herpetofauna group, which appears to have been driven by a strong preference for identifying the snake group ($X^2_{3, 1062}$ = 91.29, p < 0.0001). This resulted in the greatest crowdsourcing effort involved in the snake group; most of the 68 snake observations were verified by identifier respondents to species (97.1%). Crowdsourcing effort involved in the 267 observations in the anuran group and the 274 observations in the lizard group were moderate (52.2% and 50.7% of observations were elevated to research grade, respectively), and the least crowdsourcing effort was involved in the 453 observations in the salamander group (41.1% of observations were elevated to research grade). Of the 12 top identifier

respondents, 50% had a preference for identifying certain herpetofauna groups whereas 50% did not. Of those 6 identifier respondents who had a preference for identifying certain herpetofauna groups, as supported by the crowdsourced effort results, more (50.0%) had a preference for identifying snakes (whereas 33.3% had a preference for salamanders; 16.7% had a preference for lizards; and 0.0% had a preference for anurans). Of the uploaded observations that could be verified to herpetofauna species by the herpetologists based on the photograph (=956 observations), the fewer the total number of observations for each species, the greater the percent of observations crowdsourced to research grade ($R^2 = 0.45$, $F_{1, 1} = 8.27$, $p = 0.0165$, **Figure 3**).

Identifier respondent preference for photograph quality

There was a significant effect of observation photograph quality on crowdsourcing effort; the greater the percent specimen showing, the more likely an observation was to become elevated to research grade ($X^2_{2, 2} = 14.52$, $p = 0.0007$). Most (84.2%, = 984 observations) of the 1,169 uploaded observations that had the potential to become elevated to research grade were showing 95–100% of the specimen and those in this category accounted for the greatest crowdsourcing effort (47.4% of observations elevated to research grade). Fewer observations (15.8%, = 185 observations) were showing less than 95–100% of the specimen and resulted in reduced crowdsourcing effort (38.7% of observations elevated to research grade

when 50–94% of the specimen was showing; 22.9% of observations elevated to research grade when 0–49% of the specimen was showing). Nearly all photographs showed at least a small portion of the specimen. The reasons for reduced photograph quality (less than 95% of specimen showing) occurred due to cover from vegetation, soil and/or coverboard (93%), or specimen in the photograph being cut off (7%).

Discussion
Data collection
Data quality overview

We have concluded, as have others, that CS projects provide useful data on herpetofauna (Pittman and Dorcas 2006; Smith et al. 2015). The observers documented a new locality for Black Salamander thereby providing important data on this Near Threatened species' severely fragmented population, which is in decline because of threats including habitat destruction (IUCN 2019). The findings in the present study suggest the volunteer implementation of the iNaturalist app and coverboard technique improve data collection efforts for Black Salamander. Complementary to our findings, the Smith et al (2015) study that also used iNaturalist data from members of the public concluded that participants can improve data deficiency for a different salamander species (Green Salamander, Virginia, USA). Additionally, our study supports others in concluding that there is a close relationship between iNaturalist observation numbers and data completeness (Jacobs and
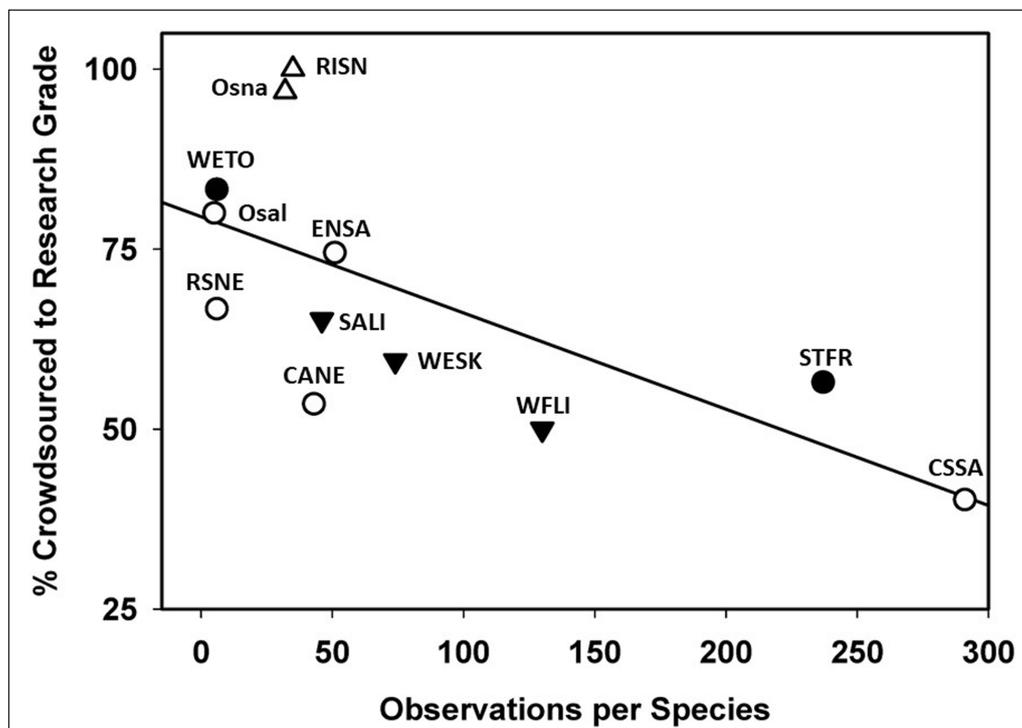

Figure 3: Percent of observations crowdsourced to research grade quality vs. number of observations per species, n = 956. Increasing observations per species decreases the crowdsourcing effort. Anuran group includes Western Toad (WETO), Sierran Tree Frog (STFR); salamander group includes California Slender Salamander (CSSA), California Newt (CANE), Rough-skinned Newt (RSNE), Ensatina (ENSA), other salamanders (Osal); lizard group includes Western Fence Lizard (WFLI), Western Skink (WESK), Southern Alligator Lizard (SALI); and snake group includes Ringneck Snake (RISN), Other snakes (Osna).

Zipf 2017). In the present study, observers using the iNaturalist app uploaded more herpetofauna observations from Pepperwood, which resulted in a higher data completeness match for Pepperwood compared to that of Osborn (X^2 = 18.24, p = 0). Further, these observations resulted in an average match of 60% of the established reserves' list of herpetofauna species observation data in a single season of sampling (2015–2016) in a limited area, plus the new Black Salamander record. The Jacobs and Zipf (2017) study, which examined data completeness of bird species represented from opportunistically collected iNaturalist observation data in USA National Parks (including 8 out of 58, 14% of parks in California, USA), showed just an average match of 8% of the existing parks' list of bird species observation data (up to 2014). Overall, the traditional coverboard technique combined with iNaturalist produced high-quality herpetofauna data. Observers were able to contribute "permanent" photographic records of many iNaturalist observations, and the crowdsourcing features of iNaturalist accurately verified data thereby reducing manager data verification effort.

Detecting herpetofauna specimens

We find, as have others, that variability in animal detection could affect data uploaded to biodiversity apps (Kelling et al. 2015). Our data confirm suggestions from previous studies that the ability of volunteers to detect herpetofauna (through visible or auditor cues) can vary (de Solla et al. 2005; McLaughlin and Hilts 1999). In a study that investigated the baseline knowledge of amateur wetland wildlife monitors and compared results of the observations of taxonomic groups including herpetofauna to those of biologists and naturalists, McLaughlin and Hilts (1999) found that monitors did not detect 44% of herpetofauna species recorded by professionals, whereas the present study did not detect 9% of herpetofauna species. We determined that data accuracy involving volunteers' detection of herpetofauna depends on the herpetofauna group and herpetofauna species. Observers were most accurate at detecting members in the anuran, lizard, and snake groups and most species in the salamander group. However, although salamanders were commonly present only in the wet season (surveys 1–3), even after a few surveys conducted, observers still were not completely accurate in successfully detecting the smaller and more camouflaged members in the salamander group, which should be a consideration for managers. It is important to emphasize that the ability for observers to detect certain species in the salamander group (or any herpetofauna detection) was not directly related to the use of a biodiversity smartphone app. However, the type of inaccuracy associated with the detection step should be a consideration for all herpetofauna monitoring programs involving volunteers.

Photographing specimens

Our study demonstrated that little or no experience was required for observers to photograph the salamander, anuran, and snake groups, because the escape during photograph opportunity was minimal. One exception was escape of certain snakes, which have higher levels of mobility and agility. Our study demonstrated that volunteers were less able to capture photographs of some lizard species in the drier part of the season and that weather variables were an important consideration in the use of a biodiversity mobile app for herpetofauna monitoring. These results concur with other studies that showed that reptile mobility is dependent upon a variety of metabolic and environmental factors (DuRant et al. 2007; Porter et al. 1975). For example, we found that a greater level of cloud cover, which is typically associated with lower temperatures, was associated with lower level of data loss because of reduced escape of certain lizard species during photograph opportunity. However, temperature had no significant effect on escape, which was surprising. It is noteworthy that this covariate was recorded at a height of 1.5 meters off the ground and there was no measurement of the surface temperature. Micro-scale temperature at each coverboard at each site might have had impacts on escape activity among these ectotherms. Observers were provided a limited time for a specimen to reappear during escape, and additional time could have reduced the escape determination.

Uploading observations

Data loss from observers' uploading observations from their smartphone to the website can result in data loss (Adriaens et al. 2015). Several volunteers needed reminders to upload their batch of data after their first survey, nevertheless, data loss associated with uploading observations accounted for a small amount (3%) of the records. However, observers improved their ability with survey experience in uploading the expected data and observer survey experience reduced data loss. This result supports the conclusion that observers can be accurate at uploading smartphone data after a limited number of uploaded batches (post-survey) and stresses the importance of early reminders by managers to inexperienced volunteers to upload data after the initial survey.

Verification

Data quality assurance aimed at identifying, correcting, and eliminating errors through the crowdsourcing approach is based on the assumption that when many users are focused on the same data, errors do not persist (Jacobs 2016). All 530 observations elevated to research grade through the iNaturalist crowdsourcing approach were accurate, which supports other studies that concluded that crowdsourcing is the best method to identify errors in biodiversity data (Goodchild and Li 2012). Other studies have assessed the accuracy of CS-verified photographic data but resulted in similar but slightly lower accuracy for species identification (Swanson et al. 2015; Leibovici et al. 2017). For example, the Swanson et al. study (2015) evaluated the Snapshot Serengeti volunteers' identifications of 48 different species and species groups (n = 4,149) of African wildlife, which resulted in 97% accuracy for species identifications with an accuracy rate that varied by species (Swanson et al. 2015, 2016).

We concur with others that photograph quality in CS projects, which allow members of the general public to

view and identify species based on the photographic evidence, can result in data loss if verification is impossible because of low photograph quality (Swanson et al. 2015; Adriaens et al. 2015; Marchante et al. 2017). Overall, of the 1,169 uploaded observations that had the potential to become elevated to research grade, 82% could be verified (to species) based on the photograph, a result similar to another study (involving invasive plants) where 76% of 221 observations could be verified (Adriaens et al. 2015). The present study determined that data loss can be reduced if a manager is willing to verify and accept data at a coarser taxonomic level (e.g., herpetofauna group rather than species would increase data accuracy by 9%). The unverifiable observations in the present study consisting of low-quality photographs emphasizes the importance of establishing a volunteer training protocol for photographing specimens to improve data verification. Other herpetofauna monitoring programs may involve different herpetofauna species than recorded in this study, so the unique identifying features of specific herpetofauna groups and species, and common behavior, should be a consideration by managers during initial photograph training efforts.

Snakes are typically encountered at a lower level among herpetofauna in CS projects, however, there is a strong interest in reporting this group (Pittman and Dorcas 2006). This finding complements our results demonstrating an indirect relationship between the total number of observations for each species and crowdsourcing effort (e.g., the fewer the number of observations for a species, the higher the percent of observations elevated to research grade). iNaturalist identifier respondents had a strong interest in snake verification, although the quantity of the snake group observations was lowest. The preference to identify snakes may result from the combination of factors involving the snake observation photographs including the size of the specimen (less need to adjust zooming and generally less effort on the part of the identifier respondents) and clarity of the unique identifying features of snake species, as well as a general fascination of snakes by identifier respondents. Overall, higher photograph quality led to increased crowdsourcing effort thereby reducing manager verification effort, which supports the recommendation for observer training and simple protocol on photograph instructions. However, handling and capture tools are generally not recommended for inexperienced volunteers because of concerns relating to animal safety (e.g., tail shed by autotomy, desiccation) and human safety (e.g., venomous and/or biting specimens).

Measures to improve data collection and verification
As a result of this study, we are able to recommend several measures to improve observer data collection and crowdsourcing effort in data verification:

- Use of smartphones, which have appropriate battery and data storage capacities and use of portable battery charging devices (especially for surveys expected to have a long duration or high number of specimens).
- Simple app protocol with no or minimal additional data entry (e.g., data fields).

- Field partner system, especially if less-experienced observers are involved.
- Herpetofauna group- or species-specific monitoring (e.g., larger salamander species).
- Photograph training on herpetofauna specimens to aid in species verification including regions where unique identifying features are typically located (e.g., eye line, top head view).
- Photograph training on camera techniques to improve photograph quality (e.g., zoom, focus, background contrast, one specimen per observation).
- Photograph training on specimen photograph capture (e.g., position and camera preparation).
- Best guess for initial species identification entry in app.
- Save and upload all observation photographs (even if observers believe the photograph is low quality).
- Review by manager or experienced observer the initial batch(es) of uploaded observations to provide quick feedback to observers to support that (1) data were properly uploaded and in a timely manner and (2) photographs are verifiable.
- Increase online duration in which identifier respondents can access and verify observations through the crowdsourcing approach.

Conclusion and Future Work
Our study was designed to measure data quality using a biodiversity smartphone app by non-expert volunteers for improving data collection and verification. Our results demonstrated that iNaturalist combined with traditional survey techniques can produce high-quality data by enhancing the data collection and verification of herpetofauna. However, our results highlight key areas related to volunteer behavior and technology, as well as study specimens, which need attention to enhance the likelihood of accurate data collection for manager use. Our study suggests that the potential for biodiversity smartphone apps and the crowdsourcing approach for data quality assurance can be applied in other fields of life science (e.g., entomology, ichthyology, botany) in either short- or long-term projects. Because many managers including conservation biologists and others are challenged with streamlining volunteer data collection and accessibility while maximizing accuracy and effort, the inclusion of such technology could provide a significant benefit.

We recommend further studies investigating the factors that affect volunteer data quality for certain survey steps across multiple seasons and years, as well as further evaluation of mechanisms that enhance volunteer accuracy and crowdsourcing effort. The efficacy of implementing volunteer training protocols to (1) find smaller and/or camouflaged salamanders, (2) photograph reptiles to prevent escape during photograph opportunity, and (3) photograph specimens to maximize the crowdsourcing approach for data quality assurance also should be investigated. Other areas of investigation should focus on effects of abundance, life stage, coverboard type, and aquatic specimens using a different detection technique. Further study on increased online time in which observations are

available to the public for data verification also should be explored. Finally, smartphone methodologies involving iNaturalist and other apps to record and verify other specimens through all steps need to be rigorously explored.

Ethics and Consent

Acknowledgements

Competing Interests
The authors have no competing interests to declare.

References
Adriaens, T, Sutton-Croft, M, Owen, K, Brosens, D, van Valkenburg, J, Kilbey, D, Groom, Q, Ehmig, C, Thürkow, F, Van Hende, P and **Schneider, K.** 2015. Trying to engage the crowd in recording invasive alien species in Europe: experiences from two smartphone applications in northwest Europe. *Management of Biological Invasions*, 6(2): 215–225. DOI: https://doi.org/10.3391/mbi.2015.6.2.12

Block, WM and **Morrison, ML.** 1998. Habitat relationships of amphibians and reptiles in California oak woodlands. *Journal of herpetology*, 51–60. DOI: https://doi.org/10.2307/1565478

Bonin, J and **Bachand, Y.** 1997. The use of artificial covers to survey terrestrial salamanders in Quebec. *Herpetological Conservation*, 1: 175–179.

Brode, JM and **Bury, RB.** 1984. The importance of riparian systems to amphibians and reptiles. *California riparian systems: ecology, conservation, and productive management.* Berkeley: Univ. California Press, 30–36.

Cohn, JP. 2008. Citizen science: Can observation contributors do real research? *BioScience*, 58(3): 192–197. DOI: https://doi.org/10.1641/B580303

Conrad, CC and **Hilchey, KG.** 2011. A review of citizen science and community-based environmental monitoring: issues and opportunities. *Environmental monitoring and assessment*, 176(1–4): 273–291. DOI: https://doi.org/10.1007/s10661-010-1582-5

Crall, AW, Newman, GJ, Stohlgren, TJ, Holfelder, KA, Graham, J and **Waller, DM.** 2011. Assessing citizen science data quality: an invasive species case study. *Conservation Letters*, 4(6): 433–442. DOI: https://doi.org/10.1111/j.1755-263X.2011.00196.x

de Solla, SR, Shirose, LJ, Fernie, KJ, Barrett, GC, Brousseau, CS and **Bishop, CA.** 2005. Effect of sampling effort and species detectability on volunteer based anuran monitoring programs. *Biological Conservation*, 121(4): 585–594. DOI: https://doi.org/10.1016/j.biocon.2004.06.018

Dickinson, JL, Zuckerberg, B and **Bonter, DN.** 2010. Citizen science as an ecological research tool: challenges and benefits. *Annual review of ecology, evolution and systematics*, 41: 149–72. DOI: https://doi.org/10.1146/annurev-ecolsys-102209-144636

DuRant, SE, Hopkins, WA and **Talent, LG.** 2007. Energy acquisition and allocation in an ectothermic predator exposed to a common environmental stressor. *Comparative Biochemistry and Physiology Part C: Toxicology and Pharmacology*, 145(3): 442–448. DOI: https://doi.org/10.1016/j.cbpc.2007.01.014

Goodchild, MF and **Li, L.** 2012. Assuring the quality of volunteered geographic information. *Spatial statistics*, 1: 110–120. DOI: https://doi.org/10.1016/j.spasta.2012.03.002

Grant, BW, Tucker, AD, Lovich, JE, Mills, AM, Dixon, PM and **Gibbons, JW.** 1992. The use of coverboards in estimating patterns of reptile and amphibian biodiversity. In *Wildlife 2001: populations*, 379–403. Netherlands: Springer. DOI: https://doi.org/10.1007/978-94-011-2868-1_31

Harpole, DN and **Haas, CA.** 1999. Effects of seven silvicultural treatments on terrestrial salamanders. *Forest Ecology and Management*, 114(2–3): 349–356. DOI: https://doi.org/10.1016/S0378-1127(98)00365-X

Hesed, KM. 2012. Uncovering salamander ecology: a review of coverboard design. *Journal of Herpetology*, 46(4): 442–450. DOI: https://doi.org/10.1670/10-220

Houze, CM, Jr and **Chandler, CR.** 2002. Evaluation of coverboards for sampling terrestrial salamanders in South Georgia. *Journal of Herpetology*, 36(1): 75–81. DOI: https://doi.org/10.1670/0022-1511(2002)036[0075:EOCFST]2.0.CO;2

iNaturalist. 2019. Help, 17 June 2019. Available at http://www.inaturalist.org/pages/help#quality [last accessed 20 June 2019].

Jacobs, C. 2016. Data quality in crowdsourcing for biodiversity research: issues and examples. *European handbook of crowdsourced geographic information*, 75. DOI: https://doi.org/10.5334/bax.f

Jacobs, C and **Zipf, A.** 2017. Completeness of citizen science biodiversity data from a volunteered geographic information perspective. *Geo-spatial Information Science*, 20(1): 3–13. DOI: https://doi.org/10.1080/10095020.2017.1288424

Joppa, L, Williams, C, Temple, S and **Casper, G.** 2009. Environmental Factors Affecting Sampling Success of Artificial Cover Objects. *Herpetological Conservation and Biology*, 5(1): 143–148.

Kelling, S. 2018. Improving Data Quality in eBird- the Expert Reviewer Network. *Biodiversity Information Science and Standards*. DOI: https://doi.org/10.3897/biss.2.25394

Kelling, S, Johnston, A, Hochachka, WM, Iliff, M, Fink, D, Gerbracht, J, Lagoze, C, La Sorte, FA, Moore, T, Wiggins, A and **Wong, WK.** 2015. Can observation skills of citizen scientists be estimated using species accumulation curves? *PloS one*, 10(10): e0139600. DOI: https://doi.org/10.1371/journal.pone.0139600

Kosmala, M, Wiggins, A, Swanson, A and **Simmons, B.** 2016. Assessing data quality in citizen science. *Frontiers in Ecology and the Environment*, 14(10): 551–560. DOI: https://doi.org/10.1002/fee.1436

Lee, YM and **Inventory, MNF.** 2012. Developing a Conceptual Framework of Recommendations for Monitoring Amphibians and Reptiles Using Non-Calling Surveys and Volunteers.

Leibovici, DG, Williams, J, Rosser, JF, Hodges, C, Chapman, C, Higgins, C and **Jackson, MJ.** 2017. Earth observation for citizen science validation, or citizen science for earth observation validation? The role of quality assurance of volunteered observations. *Data*, 2(4): 35. DOI: https://doi.org/10.3390/data2040035

Marchante, H, Morais, MC, Gamela, A and **Marchante, E.** 2017. Using a WebMapping platform to engage volunteers to collect data on invasive plants distribution. *Transactions in GIS*, 21(2): 238–252. DOI: https://doi.org/10.1111/tgis.12198

Matlack, GR. 1993. Microenvironment variation within and among forest edge sites in the eastern United States. *Biological conservation*, 66(3): 185–194. DOI: https://doi.org/10.1016/0006-3207(93)90004-K

McLaughlin, L and **Hilts, S.** 1999, October. Monitoring accuracy and the decomposition of error committed by volunteers in a wetland wildlife monitoring program. In *Leading Edge Conference Proceedings* (pp. 6–9).

Pimm, SL, Alibhai, S, Bergl, R, Dehgan, A, Giri, C, Jewell, Z, Joppa, L, Kays, R and **Loarie, S.** 2015. Emerging technologies to conserve biodiversity. *Trends in ecology and evolution*, 30(11): 685–696. DOI: https://doi.org/10.1016/j.tree.2015.08.008

Pittman, SE and **Dorcas, ME.** 2006. Catawba River corridor coverboard program: a citizen science approach to amphibian and reptile inventory. *Journal of the North Carolina Academy of Sciences*, 122(4): 142–151.

Porter, WP, Mitchell, JW, Beckman, WA and **Tracy, CR.** 1975. Environmental Constraints on Some Predator—Prey Interactions. In *Perspectives of biophysical ecology* 347–364. Berlin, Heidelberg: Springer. DOI: https://doi.org/10.1007/978-3-642-87810-7_20

Price, SJ and **Dorcas, ME.** 2011. The Carolina Herp Atlas: an online, citizen-science approach to document amphibian and reptile occurrences. *Herpetological Conservation and Biology*, 6(2): 287–296.

Ryan, TJ, Philippi, T, Leiden, YA, Dorcas, ME, Wigley, TB and **Gibbons, JW.** 2002. Monitoring herpetofauna in a managed forest landscape: effects of habitat types and census techniques. *Forest Ecology and Management*, 167(1): 83–90. DOI: https://doi.org/10.1016/S0378-1127(01)00692-2

Smith, BJ, Rochford, MR, Brien, M, Cherkiss, MS, Mazzotti, FJ, Snow, S and **Hart, KM.** 2015. CFREPTILES and AMPHIBIANS.

Stebbins, RC. 1954. Natural history of the salamanders of the plethodontid genus Ensatina. *Univ California Publ Zool*, 54: 47–123.

Swanson, A, Kosmala, M, Lintott, C and **Packer, C.** 2016. A generalized approach for producing, quantifying, and validating citizen science data from wildlife images. *Conservation Biology*, 30(3): 520–531. DOI: https://doi.org/10.1111/cobi.12695

Swanson, A, Kosmala, M, Lintott, C, Simpson, R, Smith, A and **Packer, C.** 2015. Snapshot Serengeti, high-frequency annotated camera trap images of 40 mammalian species in an African savanna. *Scientific data*, 2: 150026. DOI: https://doi.org/10.1038/sdata.2015.26

Tietje, WD and **Vreeland, JK.** 1997. The Use of Plywood Coverboards to Sample Herpetofauna in a California Oak Woodland. *Transactions of the Western Section of The Wildlife Society*, 33: 67–74.

Van Horn, G, Mac Aodha, O, Song, Y, Cui, Y, Sun, C, Shepard, A, Adam, H, Perona, P and **Belongie, S.** 2018. The inaturalist species classification and detection dataset. DOI: https://doi.org/10.1109/CVPR.2018.00914