Biological Conservation 227 (2018) 141–151

Contents lists available at ScienceDirect

Biological Conservation

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

Assessing the validity of crowdsourced wildlife observations for conservation using public participatory mapping methods

Greg Brown[a,b,*], Clive McAlpine[c,d], Jonathan Rhodes[c], Daniel Lunney[d,e], Ross Goldingay[f], Kelly Fielding[g], Scott Hetherington[h], Marama Hopkins[h], Clare Manning[i], Mathew Wood[j], Angie Brace[k], Lorraine Vass[l]

[a] California Polytechnic State University, San Luis Obispo, CA 93401, United States of America
[b] School of Earth and Environmental Sciences, University of Queensland, Brisbane, Australia
[c] School of Earth and Environmental Sciences, University of Queensland, Brisbane 4072, Australia
[d] School of Life and Environmental Sciences, University of Sydney, NSW 2006, Australia
[e] Office of Environment and Heritage, NSW 2220, Australia
[f] School of Environment, Science and Engineering, Southern Cross University, Lismore, Australia
[g] School of Communication and Arts, University of Queensland, Brisbane 4072, Australia
[h] Tweed Shire Council, Murwillumbah, NSW 2484, Australia
[i] Byron Shire Council, Mullumbimby, NSW 2482, Australia
[j] Ballina Shire Council, Ballina, NSW 2478, Australia
[k] Lismore City Council, Lismore, NSW 2480, Australia
[l] Friends of the Koala, Lismore, NSW 2480, Australia

A R T I C L E I N F O

Keywords:
Koala
Citizen science
PPGIS
VGI
Wildlife observation
Public participation

A B S T R A C T

Public participatory mapping is a method of crowdsourcing where the lay public can contribute spatial information for a range of applications including conservation planning. When used to collect wildlife observation data, participatory mapping becomes a type of "geographic citizen science" that involves collaboration with members of the public. While the potential of crowdsourcing to assist in wildlife conservation appears to be large, the quality and validity of the observational data collected remain a key concern. In this study, we examined the quality and validity of spatial data collected in a public participatory mapping project implemented in northern New South Wales (Australia) in 2018 where the public was asked to identify and map the location and frequency of koala (*Phascolarctos cinereus*) sightings using an internet mapping application. The iconic koala is a nationally-listed threatened species and has wide public recognition, making it an ideal test of our approach to examining the value of citizen science for wildlife. We assessed the validity of koala observation data from two perspectives of *validity-as-accuracy* (positional accuracy and data completeness) and *validity-as-credibility* (characteristics of spatial data contributors). To assess *validity-as-accuracy*, we analysed the distribution of citizen observations of koala sightings compared to an expert-derived probability distribution of koalas (likelihood model). To assess *validity-as-credibility*, we analysed the survey data to determine which participant characteristics increased the credibility of observational data. We found significant spatial association between crowdsourced koala observations and the likelihood model to validate koala locations, but there was under-reporting in more rural, remote areas. Significant variables contributing to accuracy in koala observations included participant knowledge of koalas, age, length of residence, and formal education. We also compared the crowdsourced results to a field-based citizen science koala observation project implemented in the same region and found crowdsourced participatory mapping provided comparable, if not superior results. Crowdsourced koala observations can augment field-based koala research by covering large geographic areas while engaging a broader public in conservation efforts. However, effective geographic citizen science projects require a significant commitment of resources, including the creation of community partnerships, to obtain high quality spatial data.

* Corresponding author.
E-mail address: ggbrown@calpoly.edu (G. Brown).

https://doi.org/10.1016/j.biocon.2018.09.016
Received 22 June 2018; Received in revised form 17 August 2018; Accepted 10 September 2018

1. Introduction

Public participatory mapping and volunteered geographic information (Goodchild, 2007) are methods of crowdsourcing (Howe, 2006) where the lay public can contribute spatial information for a range of environmental applications, including research for conservation planning. Citizen science has been defined as activities in which non-professional scientists participate in data collection, analysis and dissemination of a scientific project (Cohn, 2008). The term "geographic citizen science" refers to a subset of general citizen science where the collection of location information is an integral part of the activity (Haklay, 2013). The potential of crowdsourcing in geographic citizen science to assist in environmental problems, such as species conservation, appears large. However, the quality and validity of the citizen observation data collected remain a key concern (Alabri and Hunter, 2010; Brown et al., 2015). For example, Hunter et al. (2013) describe some of the weaknesses in general citizen science that also apply to geographic citizen science projects including the use of poorly-designed methods of data collection resulting in incomplete or inaccurate data. In participatory mapping, often called public participation GIS (PPGIS), and volunteered geographic information (VGI), a solid framework for assessing the quality of crowdsourced spatial data has yet to be established given these methods are fundamentally different to traditional geospatial assessment. The difference is due to social factors driving public contribution and the variety of types and sources of spatial content (Antoniou and Skopeliti, 2015). Furthermore, comparable authoritative data may not be available for assessing and evaluating citizen contributed data, thus requiring the use of proxy data or modelling estimates of spatial distribution.

Citizen science data can be a valuable source of information on changes in species distributions and biodiversity (Schmeller et al., 2009) but data quality may be limited due to the potential for observational bias, reporting bias, and geographical bias (van Strien et al., 2013). According to Bonney et al. (2009), contributions from citizen scientists now provide a significant quantity of data about species occurrence and distribution around the world, and include well-established projects such as *eBird*, a web-enabled community of bird watchers who collect, manage, and store their observations in a globally accessible unified database (Sullivan et al., 2009). The number of citizen science projects has grown significantly with the *SciStarter* website providing a database of > 2700 searchable citizen science projects and events (https://scistarter.com/about). With the large, rapid increase in citizen science projects, there is an increasing need for research that evaluates the quality and validity of citizen data, examines the best approaches for integration of citizen and professional/specialist science, and the design of citizen science programs for their long-term sustainability and adaptability (Paul et al., 2014).

Our focus here is on identifying and elaborating methods to evaluate the quality of crowdsourced, citizen-contributed geospatial knowledge in the specific context of species location information. Given that crowdsourced spatial data include both the social processes used to collect spatial data, and the actual spatial data generated, an assessment of data quality and validity (fitness for purpose) should include both elements. To evaluate the quality of crowdsourced data, we used the two perspectives described by Spielman (2014): *validity-as-accuracy* and *validity-as-credibility*. The *validity-as-accuracy* perspective assesses the contributed spatial data against authoritative data while the *validity-as-credibility* assesses the characteristics of the data contributors such as reputation, motivation, and place familiarity that may influence spatial data quality. Van Exel et al. (2010) used the term c*rowd quality* to describe these data quality perspectives. As a general concept, *crowd quality* seeks to assess the collective intelligence of crowd-generated data.

The *validity-as-accuracy* perspective examines spatial data quality using criteria applied to expert-derived spatial data such as *positional accuracy*, *attribute accuracy*, *logical consistency*, *completeness*, and *lineage* (see Federal Geographic Data Committee www.fgdc.gov/metadata/csdgm). Additional criteria for evaluating volunteered geographic information (VGI) data against authoritative data include *temporal accuracy* and *usability* (Antoniou and Skopeliti, 2015). The *validity-as-accuracy* perspective has been applied to VGI systems, such as the positional accuracy and completeness of public contributions to OpenStreetMap (OSM) (Haklay, 2010; Girres and Touya, 2010; Zielstra and Zipf, 2010). These studies indicated the positional accuracy of OSM data were comparable to geographical data maintained by national mapping agencies and commercial providers. Within the domain of conservation planning, moderately high levels of accuracy have been found from crowdsourced data in the location of native vegetation in New Zealand (Brown, 2012), in identifying habitat for threatened species conservation (Cox et al., 2014), and for mapped values in areas of high conservation importance (Brown et al., 2015).

The *validity-as-credibility* perspective in participatory mapping or VGI seeks to account for data quality based on the characteristics of citizen contributors. There have been relatively few published studies that evaluate participant-related variables of data quality for geographic citizen science data. Potential reasons for the lack of data quality assessment from a *validity-as-credibility* perspective include absence of participant-related data beyond basic demographic information, a predisposition towards finite citizen mapping projects over longer-term continuous projects that provide greater opportunity for data collection, and project emphasis on spatial information over user-related information. A consistent participant variable found to influence spatial data quality is participant familiarity and experience in the geographic study area. For example, Brown (2012) found that participant familiarity with the study area contributed to spatial accuracy in identifying native vegetation. In general, participatory familiarity with the study area contributes to greater mapping effort which can be a proxy for data quality when mapping subjective spatial attributes, such as place values, experiences, and preferences (Brown, 2017).

1.1. Citizen science and koala observations

There have been several field-based, citizen science projects in Australia with a focus on koalas (*Phascolarctos cinereus*). The koala has an advantage for citizen science projects because it is unique and no other animal looks like a koala. At 5–10 kg in size, it is easy recognizable once spotted and remains in people's memories. Sequeira et al. (2014) produced the first citizen science-generated estimates of koala habitat suitability and population size in South Australia based on a citizen observation program called the "Great Koala Count" which generated 1359 observations from over 1000 data contributors. While the spatial accuracy was high (i.e., *validity-as-accuracy*) because koala locations were logged using GPS technology, the limitations of the citizen-science collected data included a limited sampling window (one day observation) and significant geographic bias—most of koala observations were made within conservation parks, along streets, or in suburban backyards in areas proximate to Adelaide, South Australia. The citizen participants were also not representative of the entire South Australian population (Hollow et al., 2015). A second "Great Koala Count II" was conducted in South Australia in 2016 to address some of the limitations of the first project including an expanded sampling timeframe (see https://www.discoverycircle.org.au/projects/koala/) with results yet to be published.

Similar to the South Australian koala citizen-science projects, a field-based koala observation program was conducted in New South Wales (NSW) in 2013 and 2014 by the National Parks Association of New South Wales (www.npansw.org) and the Atlas of Living Australia (www.ala.org.au). This project was also called the "Great Koala Count" and the project area included the north coast of New South Wales, the geographic focus of the study reported herein. Data from the NSW "Great Koala Count" provide an opportunity to compare the results from two different citizen science methods (field-based "Great Koala