Examining Visitor Collected Species Data from Denali National Park and Preserve

HEATHER FISCHER [iD]

ELIZABETH WENTZ [iD]

LEAH GERBER [iD]

*Author affiliations can be found in the back matter of this article

]u[ubiquity press

ABSTRACT

In citizen science programs, the concept of "participant" is often used as a catch-all without considering how different participants (e.g., young, old, wealthy, poor, differently-abled, local inhabitants, and visitors) affect scientific outcomes of citizen science programs (e.g., the data collection, data analysis, publications, etc.). This research advances the understanding of tourist participants' ability to produce data comparable to participants who live near the study area. To examine data collected by tourist participants, we performed a case study on wildlife observation data collected through the Map of Life-Denali program in Denali National Park and Preserve in Alaska, USA. The species observation data collected by tourists and Alaska Residents were compared visually using heatmaps and statistically with the Ripley's K function variation, the L function. Results from the analysis show that the tourist and resident data have similar point patterns for the three species we compared, *Ursus arctos* (Grizzly Bear), *Rangifer tarandus* (Caribou), and *Alces alces* (Moose). Our results indicate that tourists can be effective citizen science participants. And show the potential for leveraging this large pool of untapped participants in popular tourist destinations such as U.S. National Parks.

CORRESPONDING AUTHOR:

Heather Fischer

Oregon State University, US

heather.fischer@oregonstate.edu

INTRODUCTION

Citizen science programs are a potential way for recreational and protected areas, such as U.S. national parks, to engage with their visitors (e.g., through bio-blitzes, bird counts, and picture post stations) (Cohn 2008; Delaney et al. 2007; MacKenzie et al. 2017). The participants for these programs are generally local study area residents, but tourists may also be effective participants. While other studies have compared the scientific outcomes (i.e., data products and other scientific research products) produced by different groups of participants (Crall et al. 2011; Lewandowski and Specht 2015), tourist participants and their effect on scientific outcomes have not been explicitly examined. This paper explores the potential for tourist participants to engage in citizen science programs by comparing species observation data collected by tourist and resident participants in the Map of Life-Denali citizen science program. This comparison allows us to better understand if tourists can be effective participants and produce similar scientific outcomes to resident participants.

Scientific outcomes of citizen science include datasets and data products such as analysis and modeling results, peer-reviewed papers, and reports (Wiggins et al. 2018). For this paper, we focus on the datasets collected by participants, specifically species occurrence data, which is commonly collected in citizen science programs (Dickinson et al. 2012; Tulloch et al. 2013). These data types often include taxonomic information, spatial information about where the observation occurred, temporal information about when the observation occurred, and sometimes species absences (Jordan et al. 2011; Lukyanenko et al. 2011). The geographic aspects of species occurrence data give it properties concerning data quality that are not often encountered in other types of volunteer-generated data (Elwood et al. 2012; Goodchild 2009).

Species observation data is generally collected and depicted as point data and can be aggregated to create species range and habitats maps and in species distribution models (Franklin 2013). While researchers have long been collecting species data through surveys, GPS collar tracking, and other means, citizen science participants can now collect similar point-based data with additional taxonomic information through phone-based applications. Smartphone-based citizen science applications, like Map of Life, iNaturalist, or eBird, are examples of platforms that allow participants to collect these data accurately. These data are often uploaded to online data hubs like GBIF and are free to researchers and the general public (Flemons et al. 2007). Species observation data collected through these apps have been used for species modeling and other conservation biology efforts (Toomey and Domroese 2013).

These data, however, can be prone to observer error, such as false observations, inaccurate location data, or incorrect species identification (Lukyanenko et al. 2014; Ward et al. 2015). Previous studies have examined the quality of citizen science data compared with authoritative datasets and have developed various methods to do so (Comber et al. 2013; Fischer et al. 2021; Haklay 2010; Senaratne et al. 2016; Wiggins et al. 2011). These methods range from an expert review of data records to data quality flagging, statistical analysis, and determining data fitness for use. For example, Comber et al. 2013, used a control data comparison method that involved taking a subset of data collected by the most experienced participants as control data and then comparing the spatial accuracy of the participant data to the control data.

The types and severity of observer error may be influenced by a participant's characteristics, including demographics (age, education, etc.), training, or familiarity with the study area and discipline (Delaney et al. 2008; Ratnieks et al. 2016). Studies have used some of the validation methods described above to compare data collected by different types of participants. These studies have focused mainly on age, education, and level of project-specific training. For example, Delaney et al. (2008) found that more highly educated participants produce more accurate data through an expert review of their data. We aim to add to the understanding of how participant characteristics impact the scientific outcomes of a citizen science project by examining the differences between tourist and resident participants and their effect on data accuracy.

We focus this research on a comparison between tourists and residents because there is a tendency for citizen science programs to use participants who either live near the study area or tourists who can participate in training or commit to longer-term involvement through an ecotourism program (Cooper et al. 2007; Cousins et al. 2009; Delaney et al. 2008; Lee et al. 2006; Kobori et al. 2016). In these programs, tourists intend to devote their trip to participate in a research project, such as Earthwatch trips, gap year trips for students, etc. (Brightsmith, Stronza, and Holle 2008; Brosnan, Filep, and Rock 2015; Cousins et al. 2009; Crabbe 2012; Weiler and Richins 1995). These programs have proven to be impactful for science and the participants. For example, in a study of ecotourist participants by Weaver (2015), participants felt they learned more about the environmental problems of the study area and improved their personal well-being. Ecotourism programs, however, are not accessible to a more general tourist population; they can be costly and time-consuming.

In some geographic settings, such as National Parks, historical sites, and popular urban centers, the typical tourists often visit without the intention of volunteering

as a significant part of their trip or even a part of their trip. Regardless, these typical tourists are eager to learn more about the area they are visiting (Burger 2000; Lück 2003). These typical tourists represent a large pool of potential citizen science participants (Cousins et al. 2009); the National Park Service alone expects an 8–23% increase in visits in the next few years (Fisichelli et al. 2015). To reach a wider tourist audience who may not be seeking out ecotourism opportunities or may not have the money or time to dedicate to an ecotourism trip, these potential tourist participants can become involved in less time-consuming and less financially intensive citizen science programs.

Some hesitancy around recruiting tourists to participate in citizen science programs may be because of the possibility of unique spatial and temporal biases. In particular, with species identification, participants familiar with the study area or subject are shown to produce more accurate data (Kosmala et al. 2016). Additionally, tourist destinations have a time of year that is more popular for tourists, and tourists may not venture away from popular sightseeing areas. However, there are established protocols for assessing some of these issues. For example, prequalifying skills or knowledge assessment is one way to determine participant knowledge/skills and the potential for accurate data collection (Kosmala et al. 2016). If participants do not already possess the skills or knowledge needed for the programs, some programs offer training sessions on using instruments or correctly collecting data. If these training sessions are short and regularly offered (or on-demand), more tourists may be able to participate.

Additionally, smartphone applications and other technologies make it more feasible for tourists to participate in citizen science programs. Mobile device–based citizen science applications, such as eBird, iNaturalist, GLOBE, and Map of Life (Newman et al. 2012), are free and easy to operate. This technology allows for accurate, efficient, and cost-effective data collection for citizen science programs, especially when participants collect spatial data. Many smartphones have built-in location systems that can be used with or without phone service (Bruce et al. 2014; Kim et al. 2011). Many people already know how to use smartphones and therefore do not need instrument training. With simple in-app tutorials, any smartphone user could potentially contribute valuable data in popular tourist areas (Devisch and Veestraeten 2013).

To examine the possibility of tourists participating in citizen science programs, we focused on assessing the accuracy of tourist-collected species data collected through the Map of Life- Denali program. We hypothesized that although tourist participants are likely more unfamiliar with the park wildlife than the resident participants at the start of their participation in the program, using a simple mobile app, the tourists and resident participants would collect comparable datasets. To address this hypothesis, we analyzed the results of a wildlife knowledge quiz embedded into a pre-visit survey to assess the local resident and tourist participants' knowledge about the park and wildlife. We then compared the species occurrence data collected by the two groups through the Map of Life (MOL) mobile application to see if they still produced data similar to knowledgeable local participants despite the tourists' unfamiliarity with the park.

METHODS

STUDY AREA

This study was conducted in Denali National Park and Preserve (Denali NP&P) (Figure 1). The park covers 6 million acres and is home to charismatic megafaunas such as Grizzly bears, moose, caribou, Dall sheep, and wolves. The wildlife viewing opportunities and landscapes of Denali NP&P attract ~500,000 summer visitors each year. Most (90%) of the visitors are tourists (i.e., not from Alaska) who spend, on average, three days in the park (Manni et al. 2012). This large number of tourist visitors in Denali created a potential pool of participants for this study.

Denali NP&P officials were supportive of this research for two main reasons. First, during the Centennial for the National Park Service (2016) and Denali NP&P 100 anniversary (2017) celebrations, park managers welcomed projects that directly engaged the public in park research initiatives. Second, there was potential for this project to support the park's Road Ecology Program (REP). The REP actively monitors how the ecology along the park road is impacted by the transportation system (Kilkus et al. 2011). Data from the park's Ride Observation and Record (ROAR) program are currently used to inform the REP's monitoring and research. ROAR participants are local residents or seasonal workers compensated for riding the park busses and recording wildlife sightings and behavior. Denali NP&P was interested in exploring the use of citizen science to complement the data collection efforts of the ROAR program and wanted to explore how a citizen science program focused on collecting species occurrence data could expand the spatial and temporal extent and scale of the ROAR data.

DATA COLLECTION AND ANALYSIS

The key instrument in the Map of Life-Denali citizen science program is the Map of Life mobile app, designed and developed by the Jetz lab at Yale University (Jetz et al. 2012). (Figure 2). A Denali-specific sub-app (called Map of Life-Denali) was developed to be used without a Wi-Fi or cellular data connection; both are limited in the park.
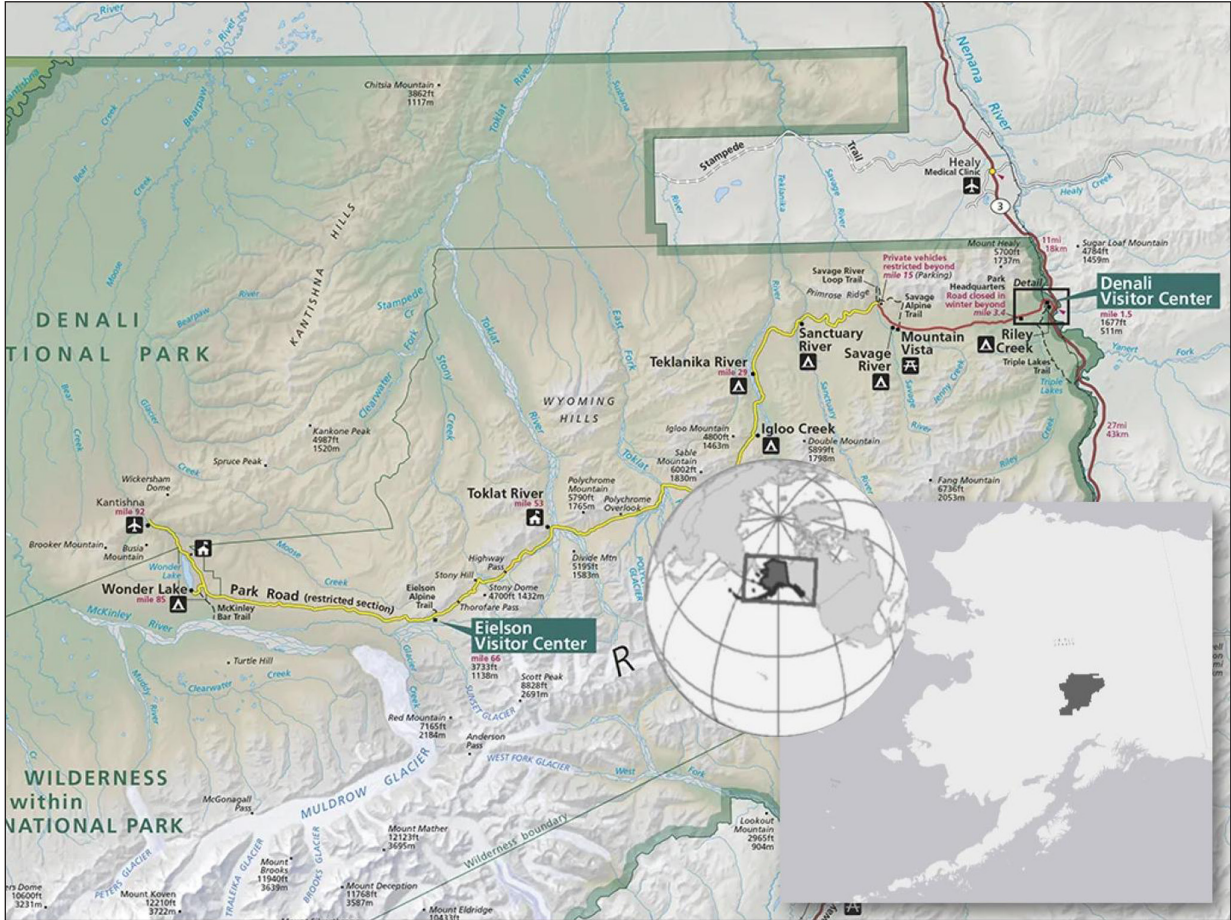
Figure 1 Denali National Park and Preserve. The park road provides the main point of access into the park and meanders through prime habitat for much of Denali's wildlife, and it is only accessible for visitors via bus (Park Map obtained from https://www.nps.gov/carto/).
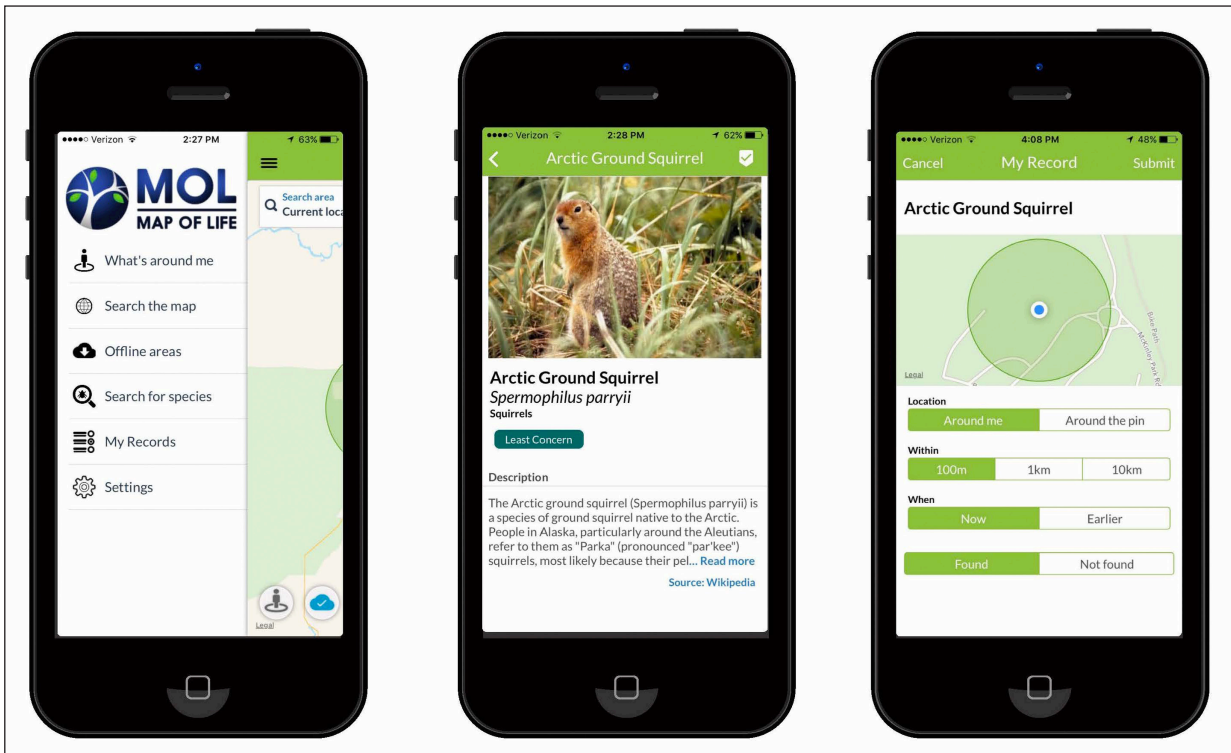

Figure 2 Map of Life- Denali Mobile Application: home page, species information page, and record observation page (mol.org).

The customized home page includes links to the surveys used in this research. An information page with park-specific information, such as animal safety warnings, and a description of the citizen science project, is featured in the Map of Life-Denali application. When a participant searches for a specific animal, identifying photos and detailed species information appear in the application. The application allows participants to record the precise location of their wildlife observations while touring the park with their phone's internal GPS to capture spatial data. The participants were encouraged to collect data while riding the buses into the park, going on hikes, or camping there. Thus most of the data collected for the project is concentrated around the park road corridor and does not cover the vast background areas of the park. This particular spatial bias was welcome because the park was interested in seeing how the data compares with the ROAR program data and may be helpful to the Road Ecology Program.

Participants were recruited through signage advertising the citizen science program posted at the park visitor centers, the bus depot, campgrounds, the community library, the park employee information board, the local coffee shop, and hotels. Additionally, the authors personally recruited participants while visitors waited to board buses to enter the park. Participation in the program and research study was open to anyone over 18. The Arizona State University Internal Review Board (STUDY00003874) and the National Park Service (DENA-2016-SCI-0002) approved this study and the data collection methods.

The pre-visit survey was linked in the MOL app, and participants were instructed to take the pre-visit survey right after downloading the app. Through the pre-visit survey, participants self-identified as either tourist (non-Alaska residents) or residents (Alaska residents, including seasonal workers). The survey also included a short wildlife quiz that asked participants multiple-choice and true/false questions about the park's wildlife and ecosystems to determine how much the participant knew about the park (Table 1). The authors and the park social scientist developed the quiz. The quiz format was influenced by the quizzes given to participants in the Portland Urban Coyote Project (Rasmussen 2015). The quiz section of the pre-survey was analyzed using descriptive statistics (i.e., mean). The pre-visit survey also included questions that do not pertain to this particular paper's scope, a comprehensive survey analysis, and a full copy of the survey published in Fischer and Wentz 2020.

The species occurrence data used in this study comes from data collected with the MOL app in the Summer of 2016. These data contain the wildlife observation's geographic coordinates, taxonomic information for the species, a time stamp, and a unique observer I.D. These data were retrieved from the Map of Life server on September 30th, 2016, and cover the time period between June 25th to September 25th, 2016.

QUESTION	QUESTION CHOICES	
What should you do if you are confronted by a moose?	• Stand your ground • Throw rocks • Run	• Play dead • I don't know
True or False, both male and female Caribou have antlers.	• True • False	• I don't know
Which bird turns all white in the winter?	• Ptarmigan • Common Loon • Mountain Chickadee	• Gyrfalcon • I don't know
What is an indication of climate change?	• Tree line moving higher in elevation. • Glacier melt	• Wildlife behavior changes • All of the above • I don't know
What are the two major ecosystems in the park?	• Tundra and Taiga (Boreal Forest) • Tundra and Rainforest	• Taiga and Temperate Forest • I don't know
How far must you stay away from a bear?	• 100 Yards (meters) • 300 Yards (meters) • 25 yards (meters)	• 700 yards (meters) • I don't know
True or False, Denali's wolf population has lost nearly ⅔ of its previous population levels.	• True • False	• I don't know

Table 1 Wildlife quiz questions from the pre-visit survey.

There are various methods to compare spatial datasets; we chose to employ relatively simple spatial analysis methods that can be done with open-source software and replicated by other citizen science practitioners with potentially limited experience analyzing spatial data. First, we performed exploratory data analysis using heat map overlays; then, we used Ripley's L function to explore the relationship between the datasets further. We used these methods to compare resident and tourist data for three megafauna in Denali: *Ursus Arctos* (Grizzly bear), *Rangifer tarandus* (caribou), and *Alces alces* (moose).

The heat maps show, in raster layers, the density of wildlife observations made throughout the study period for residents and tourists. The value of each cell in the raster layer represents the number of species observations made in that cell. The raster layers were then used in a spatial overlay analysis to identify where the species observations from these two groups overlap. This analysis shows if the residents' and tourists' species observations for bear, caribou, and moose are in similar geographic locations, and allowed us to visually compare the species observations collected by the tourists with the resident-collected data.

To create the comparative heat maps, the species occurrence point data was converted to raster format through the Kernel Density tools in ESRI's ArcMap (this method can also be done with Q-GIS an open-sourced spatial analysis software). A raster consists of rows and columns of cells where each cell contains a value; in this case, the cell values represent the number of species observations made within that cell. This Kernel Density tool calculates the density of points around each output raster cell; thus, a smooth surface is created. The chosen cell size was 1,000 meters because of the level of error in the Map of Life-Denali application (participants can choose how far they are away from the wildlife; 1,000 meters is the farthest option). The cells with the highest values contain the points, cell values decrease farther away from the point, and cells with zero values are at the limit of the search radius distance from the point. Each cell's values in the two raster layers are added together via ESRI's ArcMap Overlay tool to create a new raster layer, reflecting where these raster layers correspond.

This resulting overlay raster area was calculated to determine how much the tourists' species observations coincide with the residents' species observations. This area is determined by counting the number of cells whose value does not equal zero (if the cell equals zero, these underlying raster layers do not overlap at all in that cell). To calculate how much the tourist's species observations coincide with the resident's species observations, the tourist raster layers' total area was divided by the overlap raster.

To further compare the tourist and resident data, we performed Ripley L Function analysis (a variation of Ripley's K function, also referred to as Besag's L Function). Ripley's K function and its variations—e.g., the L function-compares spatial data made up of points. Specifically, the function determines if point datasets have a similar spatial distribution or if the distribution of points is spatially random at a given distance from each other (Ripley 1977). The function determines this by putting a buffer around a random point in the dataset and computing the proximity of other points, and it repeats this at multiple distances. If multiple points are found in the proximity, then the spatial pattern is likely clustered and not random. The Bivariate L Function is a preferred modification of the K function because it can produce more interpretable results (Besag 1977).

The function output includes a graph showing if the included datasets are clustered together or dispersed at given distances. If the graph shows the data is clustered, it indicates a similar point pattern and distribution between the datasets. If the graph shows that the data is distributed, it indicates the datasets do not have a similar point pattern and are not spatially similar. A mathematical formulation of the L function can be found in Bailey and Gatrell (1995, p. 94). In ecology, the Ripley L function—or variations of it—has been used to compare human-wildlife relations, predator-prey interactions, etc. (Hasse 1995; Schaible et al. 2020; Teixeira-Neto et al. 2014; Zhang et al. 2017). This analysis allowed us to quantify how the tourists' observations were distributed relative to the observations made by residents, and to determine if the tourists were making observations that were spatially similar to those of residents.

To account for significant differences in sample size between the residents and tourists, a random sample was taken from the more plentiful tourist data to match the smaller sample size of the resident data. The random sample was derived from the Biogeography add-on sampling tool in Esri's ArcMap Version 10.3, which selects a random sample from the existing point data. Ripley's L function was run in R 4.0.2 using the spatstats package and L-cross command. The function's confidence envelopes (the range of the best outcomes of the distributions) were calculated with 999 simulations of the model of complete spatial randomness.

In addition to analyzing the tourist and resident data with Ripley's L function, we compared the Map of Life–collected data with ROAR program species data, which park biologists actively use for monitoring. This analysis helped us understand the potential for using species data collected by tourist participants to enhance an authoritative dataset like the ROAR program data. The ROAR data used in this

analysis were also collected during the summer of 2016 and included taxonomic information, location information, and a time stamp. This additional analysis is an initial assessment of the data accuracy of the Map of Life data. A comprehensive data fitness-for-use analysis of the full participant collected dataset is presented in Fischer et al. (2021).

RESULTS

We first analyzed the pre-visit quiz results to assess the local resident (n = 22) and tourist (n = 117) participants' knowledge about the park and wildlife. The residents scored an average of 89.79% on the quiz (meaning they correctly answered 89.79% of the quiz questions), and the tourists scored 59.84% (a difference of 30.25% between the quiz scores). Many of the resident participants answered every

question correctly. The pre-visit quiz shows us that the resident participants are knowledgeable about the wildlife, so we expect to get accurate data from this group and thus treated the resident data like control data (similar to Comber et al. 2013). This quiz also indicates that data from tourist participants may contain some errors (likely species identification errors).

The tourist and resident Grizzly bear observations reflect similar hot spots determined by the overlay layers (Figure 3a). These raster layers were derived from 21-point observations from the residents and a random sample of 21-point observations from the tourists. Darker colors represent hotspot areas, that is, areas with a high density of observation points. The total area of overlap is 502 km2. The tourists' observations overlap 66.93% with the residents' observations; thus, most tourist observations overlap with the resident observations. The darker red area represents the overlay raster layer, where there is a
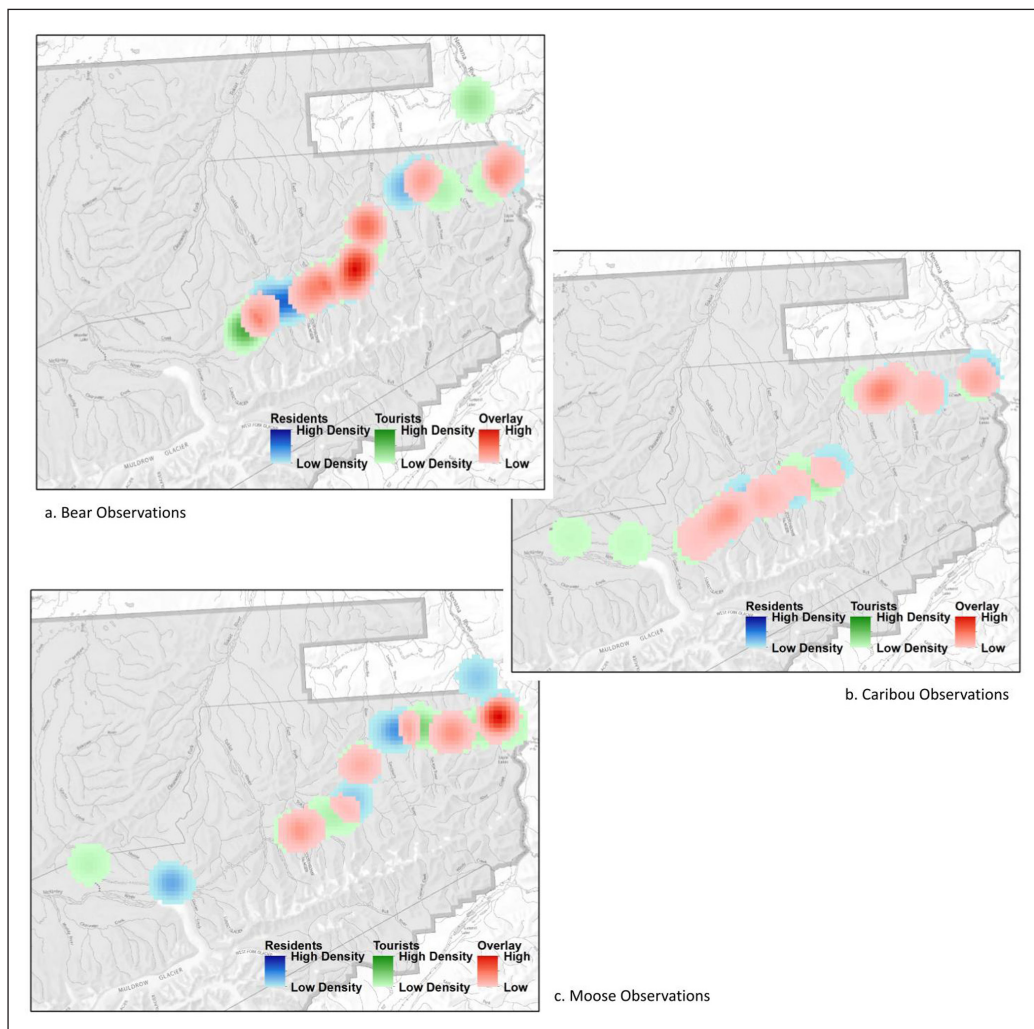

Figure 3 Heatmaps of each of the study species showing overlay between the tourist and resident collected data. **(a)** Grizzly bear observations from tourist (*green*) and resident (*blue*) volunteers (with resampled tourist data due to differences in sample size-resampling methods described in the Methods section). The overlay is shown in red. **(b)** Caribou Observations from tourist and resident volunteers. **(c)** Moose observations from tourist and resident volunteers.

greater consistency amongst these data. Some hot spots in the tourist data are inconsistent with the resident data and vice versa. However, the overlay shows consistent hot spots in known areas of Grizzly bear habitat in the park.

The caribou observations made by tourist and resident participants are represented in Figure 3b. This map shows areas of hot spots from each participant group, and the areas where these data overlap are shown in red. The total area of the overlay with the resampled tourist data is 613 km2. The tourist observations overlap 70.13% with the resident observations; thus, most of the observations overlap. The hot spot areas are not as pronounced as they were with the Grizzly bear observations. The area with greater consistency among the two samples (darker red areas) is in the park's known areas of summertime Caribou habitat.

Tourist and resident moose observations are depicted in Figure 3c. Areas of darker blue (resident observations)

or green (tourist observations) represent hot spots with a higher density of observation points. The total area of the overlay is 422 km2. The tourist observations overlap 65.73% with the resident observations. The overlay area (red) shows where the two sets of observations overlap, and the darker red area indicates a higher consistency amongst these data. It is expected to see a high density of moose observations near the park entrance area (upper right-hand side of the map) because this well-known moose habitat. The results from all heat map analyses show an overlap of the datasets. Thus, we expect to see a correlation between the datasets with Ripley's L function analysis.

Results from the L function analysis on the Grizzly bear observations (Figure 4a) show that the tourist (T) and resident (AK) data are significantly clustered at all considered distances (r) (the solid black line representing the species data in the graph is above the upper bound of
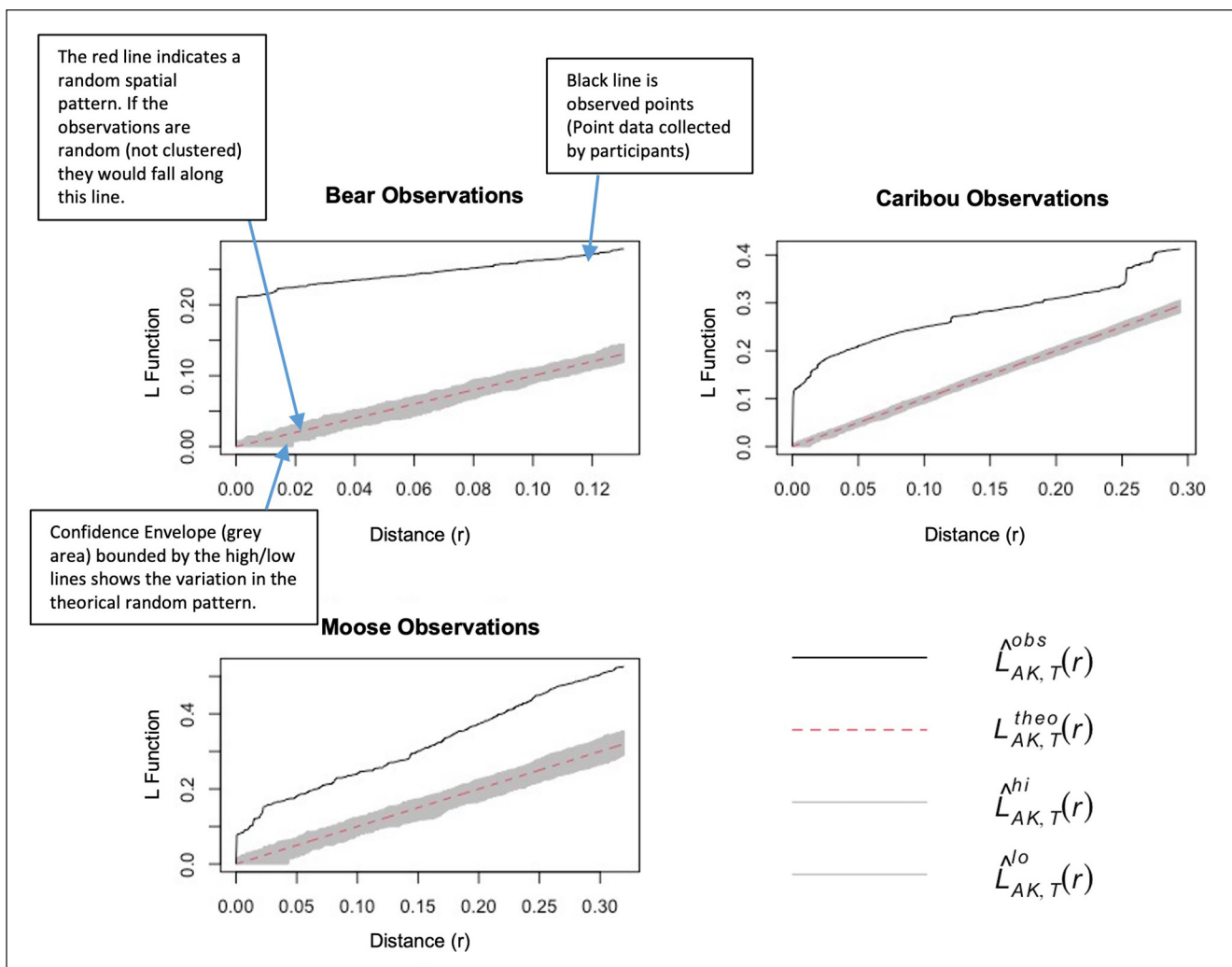

Figure 4 Results from L Function analysis of tourist and resident **(a)** Grizzly bear, **(b)** caribou, and **(c)** moose observations. The solid black line on each graphs shows the observed points. The dotted red line is the expected random pattern, and the grey area bounded by the high/low lines shows the confidence envelope. In each graph, the black line is above the envelope across the considered distances in kilometers (r), indicating the two point datasets are clustered and have similar spatial patterns.

the envelope-expected range of values). L(r) values larger than the confidence envelope indicate that the tourist data pattern more closely followed the pattern of the resident data than expected if it were random points, that is, the tourist data and the resident data are spatially similar. The caribou observations (Figure 4b) and moose observations (Figure 4c) also show clustering between the tourist- and resident-collected data. The clustering of the tourist and resident datasets with all three species indicates that the tourist-collected data is on par with the resident data– at least for these megafaunas.

We used Ripley's L function to compare the point patterns of our three study species (bear, caribou, and moose) collected through Map of Life and the ROAR program to examine the accuracy of the Map of Life data as a whole (with both tourists and resident observations). Park biologists consider the ROAR program data to be accurate and authoritative and are used by park biologists for monitoring reports. Figure 5a shows the bear observations comparison of the Map of Life data and the ROAR program data, and the data are significantly clustered at all considered distances. Figure 5b shows the Map of Life and ROAR Caribou observations, and Figure 5c shows the Map

of Life and ROAR moose data. Both also show clustering. The MOL and ROAR data clustering with all three species indicates that the MOL data is spatially similar to the ROAR data for these species.

DISCUSSION

Using both a map overlay technique and spatial statistics, this study shows that despite tourist participants being at first unfamiliar with the park wildlife (as indicated by the pre-visit quiz), tourist and resident participants produce similar species occurrence data. We recognize that many factors can affect the ability of tourists to be effective citizen science participants. Our study focused on a simple-to-use mobile app and analyzed data on easy-to-identify megafauna. If projects wish to engage tourists in their program, they need to design the program to be done by participants who may not be familiar with the study and may have limited time to spend on the project but likely have a lot of enthusiasm to learn more about the place they are visiting. We suspect we would not see such high clustering and overlap in the data if we performed an
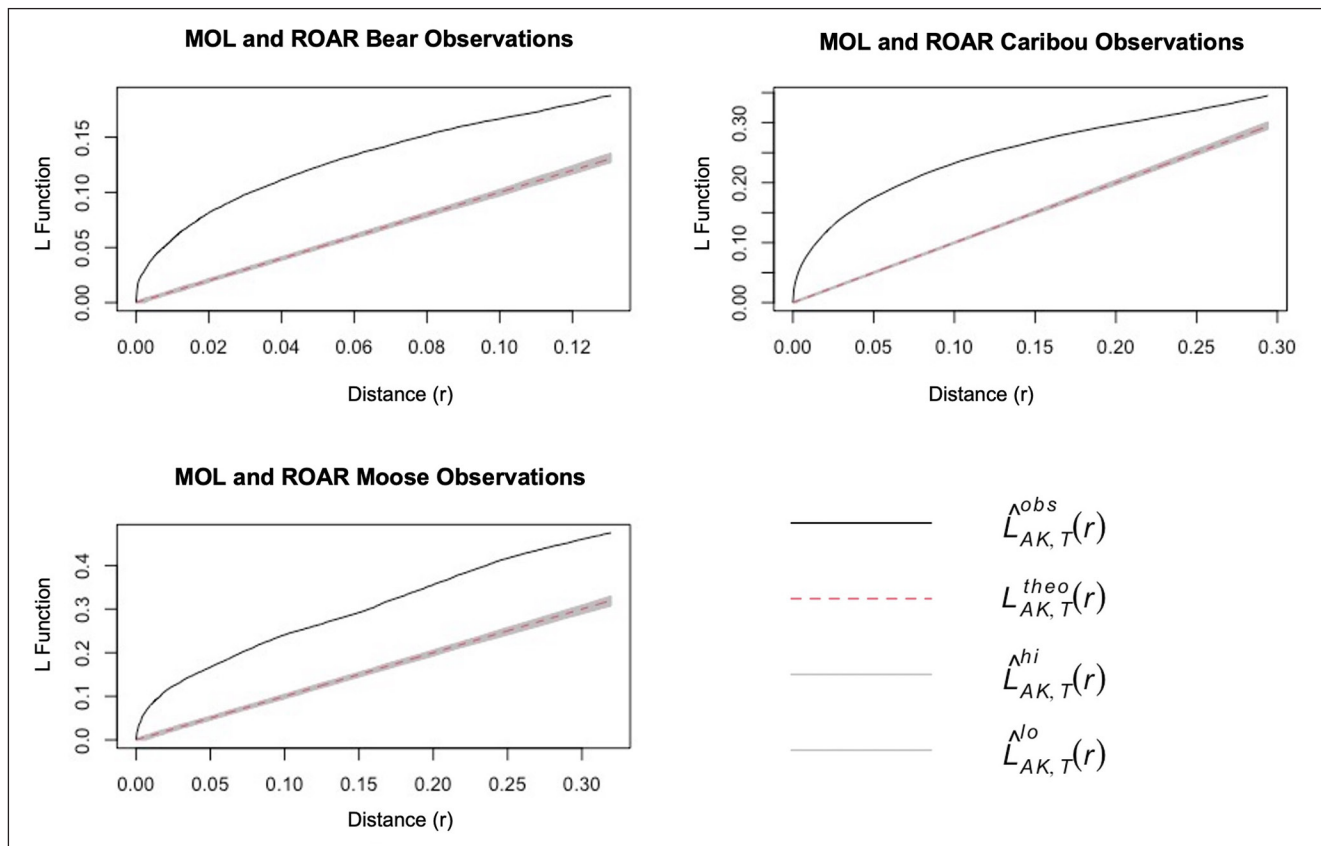

Figure 5 Results from L function analysis of the Map of Life and ROAR program data comparison of Grizzly bear, caribou, and moose observations. Panel *a* compares the Map of Life and ROAR Bear observations, and the two datasets are clustered. Similarly, in panels *b* and *c*, the caribou and moose observation data are also clustered.

analysis comparing tourist and resident data on lesser-known species, such as ground squirrels or some bird species. More studies should be done to understand how to engage tourists in citizen science effectively and how projects can be designed with these participants in mind.

The two methods we used to compare the datasets were reasonably simple to implement and use in this data exploration stage. We hope that citizen science project practitioners will utilize these or similar methods to perform data quality checks on their data and compare different groups of participants to ensure data accuracy across the project. For example, these simple checks can help program facilitators determine whether to retrain participants or adjust training and recruitment methods. These checks can also be used to compare citizen science data with authoritative datasets, as we did with the MOL and ROAR data. The heat maps visually represented the tourists and resident data for each species. The heatmaps also help determine if the tourists are misidentifying species. For example, if a tourist consistently misidentifies a caribou as a Grizzly bear, this sighting may be an outlier in the Grizzly bear overlay analysis. The L function was helpful in statically verifying what is visualized in the heatmaps.

Our analysis had some limitations and challenges. The limited timeframe in which these data were collected created limitations regarding the small sample size for the surveys and the species data. The sample sizes captured in this case study reflect the proportion of tourists versus residents who visit the park (every year, on average, 15% of park visitors are Alaska residents (Manni et al. 2012), and 15% of our participants were residents). However, for spatial data analysis purposes, the difference in sample size between the two groups was addressed by choosing random samples from the tourist dataset to equal the same number of observations in the resident dataset. The random samples were then averaged to create a final assessment of, on average, how much the tourist samples overlay with the resident samples. This method should be further refined and possibly replaced by an automated bootstrapping approach.

CONCLUSION

Researching the capabilities of participants and characteristics of the data they collect, as well as measuring outcomes of citizen science programs, is a prerequisite to ensuring successful programs for both the participants and the scientists. While citizen science data are being more widely accepted as an accurate and valuable source of data, many variables can affect the quality of a particular citizen science dataset, including participant

characteristics. However, a properly designed program that considers participants' characteristics can diminish the potential effects on scientific outcomes. The program we studied was designed to be simple and accessible to many participants, particularly those unfamiliar with the area in which they are collecting data. This research shows the potential for tourists to be effective participants and support research and monitoring in national parks and other tourist areas. This paper also shows that the participant-collected dataset (Map of Life) is similar to the authoritative data (ROAR program) for the species we focused on. This result shows the potential for citizen science data to be used in conjunction with authoritative data to create more spatially and temporally robust datasets.

By researching an untapped pool of participants and showing the potential for engaging them in citizen science, we hope that more programs will consider partnering with recreation areas like national parks and including tourists in their programs. However, we caveat this call to action with caution. Practitioners must consider the program's design if they wish to engage tourists. A citizen science program that includes intensive training or a complicated task is likely not appropriate for tourists. Designing for your participants is key to the success of the project both in terms of participant outcomes and scientific outcomes (Edelson et al. 2018; NASEM 2018). Connecting with tourists, particularly in national parks, can enhance the tourist experience and improve the program's scientific outcomes. Tourists can also help projects collect data on a larger spatial and temporal scale. Additionally, the experience of participating in a citizen science program can be transformative and provide people with a sense of stewardship and intensify their connection to the study area (Haywood, Parrish, and Dolliver 2016; Fischer and Wentz 2020), thus encouraging more people to be engaged, global citizens.

DATA ACCESSIBILITY STATEMENT

Species occurrence data made through the mobile application data are available from the authors upon request. Data from Denali National Park's ROAR program or the pre-visit quiz survey is not publicly available.

ETHICS AND CONSENT

The Arizona State University Internal Review Board (STUDY00003874) and the National Park Service (DENA-2016-SCI-0002) approved this study and the data collection methods.

ACKNOWLEDGEMENTS

FUNDING INFORMATION

AUTHOR CONTRIBUTIONS

Drs. Fischer, Wentz, and Gerber contributed to the study and program's conception and design. Dr. Fischer performed data collection and analysis and wrote the manuscript. Drs. Wentz and Geber reviewed the manuscript and provided intellectual support.

COMPETING INTERESTS

The authors have no competing interests to declare.

AUTHOR AFFILIATIONS

Heather Fischer orcid.org/0000-0003-2530-3529
Oregon State University, US

Elizabeth Wentz orcid.org/0000-0002-2881-0668
Arizona State University, US

Leah Gerber orcid.org/0000-0002-6763-6842
Arizona State University, US

REFERENCES

Bailey, TC and **Gatrell, AC.** 1995. Interactive spatial data analysis. Essex: Longman Scientific & Technical.

Besag, J. 1977. Contribution to the discussion on Dr Ripley's paper. *Journal of the Royal Statistical Society*, 39: 193–195.

Brightsmith, D, Stronza, A and **Holle, K.** 2008. Ecotourism, conservation biology, and volunteer tourism: A mutually beneficial triumvirate. *Biological Conservation*, 141(11). DOI: https://doi.org/10.1016/j.biocon.2008.08.020

Brosnan, T, Filep, S and **Rock, J.** 2015. Exploring synergies: Hopeful tourism and citizen science. *Annals of Tourism Research*, 53(0). DOI: https://doi.org/10.1016/j.annals.2015.05.002

Bruce, E, Albright, L, Sheehan, S and **Blewitt, M.** 2014. Distribution patterns of migrating humpback whales (Megaptera novaeangliae) in Jervis Bay, Australia: A spatial analysis using geographical citizen science data. *Applied Geography*, 54: 83–95. DOI: https://doi.org/10.1016/j.apgeog.2014.06.014

Burger, J. 2000. Landscapes, tourism, and conservation. *Science of the Total Environment*, 249: 39–49. DOI: https://doi.org/10.1016/S0048-9697(99)00509-4

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

Comber, A, See, L, Fritz, S, van der Velde, M, Perger, C and **Foody, G.** 2013. Using control data to determine the reliability of volunteered geographic information about land cover. *International Journal of Applied Earth Observation and Geoinformation*, 23: 37–48. DOI: https://doi.org/10.1016/j.jag.2012.11.002

Cooper, CB, Dickinson, J, Phillips, T and **Bonney, R.** 2007. Citizen science as a tool for conservation in residential ecosystems. *Ecology and Society*, 12(2): 11. DOI: https://doi.org/10.5751/ES-02197-120211

Cousins, J, Evans, J and **Sadler, J.** 2009. Selling Conservation? Scientific Legitimacy and the Commodification of Conservation Tourism. *Ecology and Society*, 14(1). DOI: https://doi.org/10.5751/ES-02804-140132

Crabbe, M. 2012. From citizen science to policy development on the coral reefs of Jamaica. *International Journal of Zoology*. DOI: https://doi.org/10.1155/2012/102350

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

Delaney, D, Sperling, C, Adams, C and **Leung, B.** 2008. Marine invasive species: validation of citizen science and implications for national monitoring networks. *Biological Invasions,* 10: 117–128. http://link.springer.com/article/10.1007/s10530-007-9114-0.

Delaney, DG, Sperling, CD, Adams, CS and **Leung, B.** 2007. Marine Invasive Species: Validation of Citizen Science and Implications for National Monitoring Networks. *Biological Invasions,* 10: 117–28. DOI: https://doi.org/10.1007/s10530-007-9114-0

Devisch, O and **Veestraeten, D.** 2013. From Sharing to Experimenting: How Mobile Technologies Are Helping Ordinary Citizens Regain Their Positions as Scientists. *Journal of Urban Technology*, 20(2): 63–76. DOI: https://doi.org/10.1080/10630732.2013.769313

Dickinson, JL, Shirk, J, Bonter, D, Bonney, R, Crain, RL, Martin, J, Phillips, T and **Purcell, K.** 2012. The current state of citizen science as a tool for ecological research and public engagement In a nutshell. *Frontiers in Ecology*

and the Environment, 10: 291–297. DOI: https://doi.org/10.1890/110236

Edelson, DC, Kirn, SL, Baukus, A, Blair, R, Ford, M and **Grow, S.** 2018. Designing citizen science for both science and education. In *A Workshop Report, Colorado Springs*, CO: BSCS Science Learning. Retrieved from https://media.bscs.org/tech-report/2018-1/2018-1.pdf.

Elwood, S, Goodchild, MF and **Sui, DZ.** 2012. Researching Volunteered Geographic Information: Spatial Data, Geographic Research, and New Social Practice. *Annals of the Association of American Geographers*, 102(3): 571–590. DOI: https://doi.org/10.1080/00045608.2011.595657

Fischer, HA, Gerber LR and **Wentz, EA.** 2021. Evaluating the quality of citizen science data for wildlife monitoring in the face of climate change. *Frontiers in Environmental Science*, 9 DOI: https://doi.org/10.3389/fevo.2021.620850

Fischer, HA and **Wentz, EA.** 2020. Place attachment and learning outcomes among tourists who volunteer for a U.S. National Park science volunteer program. *Applied Environmental Education & Communication*, 20(2): 123–138. DOI: https://doi.org/10.1080/1533015X.2020.1726840

Fisichelli, NA, Schuurman, GW, Monahan, WB and **Ziesler, PS.** 2015. Protected Area Tourism in a Changing Climate: Will Visitation at US National Parks Warm Up or Overheat? *Plos One*, 10(6). DOI: https://doi.org/10.1371/journal.pone.0128226

Flemons, P, Guralnick, R, Krieger, J, Ranipeta, A and **Neufeld, D.** 2007. A web-based GIS tool for exploring the world's biodiversity: The Global Biodiversity Information Facility Mapping and Analysis Portal Application (GBIF-MAPA). *Ecological Informatics,* 2: 49–60. DOI: https://doi.org/10.1016/j.ecoinf.2007.03.004

Franklin, J. 2013. Species Distribution Models in Conservation Biogeography: Developments and Challenges. *Diversity and Distributions,* 19(10): 1217–23. DOI: https://doi.org/10.1111/ddi.12125

Goodchild, MF. 2009. The Quality of Geospatial Context. In: Rothermel, K, Fritsch, D, Blochinger, W and Dürr, F (eds.) *Quality of Context*. Berlin: Springer.

Haklay, M. 2010. How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. *Environment and Planning B-Planning & Design,* 37(4): 682–703. DOI: https://doi.org/10.1068/b35097

Haywood, B, Parrish, J and **Dolliver, J.** 2016. Place-based and data-rich citizen science as a precursor for conservation action. *Conservation Biology*, 30(3): 476–486. https://scholar.google.com/citations?view_op=view_citation&continue=/scholar%3Fhl%3Den%26start%3D20%26as_sdt%3D0,3%26scilib%3D1&citilm=1&citation_for_view=0APm3KEAAAAJ:ClCfbGk0d_YC&hl=en&oi=p.

Jetz, W, McPherson, JM and **Guralnick, RP.** 2012. Integrating biodiversity distribution knowledge: toward a global map of life. *Trends in Ecology & Evolution*, 27(3): 151–159. DOI: https://doi.org/10.1016/j.tree.2011.09.007

Jordan, RC, Gray, SA and **Howe, DV.** 2011. Knowledge Gain and Behavioral Change in Citizen-science Programs. *Conservation Biology*, 25(6): 1148–54. DOI: https://doi.org/10.1111/j.1523-1739.2011.01745.x

Kilkus, KR, Bernatz, GC, Robertson, AG, Drazkowski, BW, Lee, CE, Iverson, EJ and **Knopf, JC.** 2011. Denali National Park and Preserve: Natural Resource Condition Assessment. *Natural Resource Report NPS/NRSS/WRD/NRR—2011/424*. Fort Collins, Colorado.

Kim, S, Robson, C, Zimmerman, T, Pierce, J and **Haber, EM.** 2011. Creek watch. In: *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems - CHI '11*. New York, USA: ACM Press. pp. 2125–2125. DOI: https://doi.org/10.1145/1978942.1979251

Kobori, H, Dickinson, JL, Washitani, I and **Sakurai, R.** 2016. Citizen science: a new approach to advance ecology, education, and conservation. *Ecological Research,* 31(1): 1-19. DOI: https://doi.org/10.1007/s11284-015-1314-y

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, T, Quinn, M and **Duke, D.** 2006. Citizen, science, highways, and wildlife: using a web-based GIS to engage citizens in collecting wildlife information. *Ecology and Society*, 11(1). DOI: https://doi.org/10.5751/ES-01681-110111

Lewandowski, E and **Specht, H.** 2015. Influence of volunteer and project characteristics on data quality of biological surveys. *Conservation Biology,* 29(3): 713–723. DOI: https://doi.org/10.1111/cobi.12481

Lück, M. 2003. Education on Marine Mammal Tours as Agent for Conservation—but Do Tourists Want to Be Educated? *Ocean & Coastal Management*, 46(9): 943–56. DOI: https://doi.org/10.1016/S0964-5691(03)00071-1

Lukyanenko, R, Parsons, J and **Wiersma, Y.** 2011. Citizen science 2.0: Data management principles to harness the power of the crowd. In *International Conference on Design Science Research in Information Systems*. May: 465–473. DOI: https://doi.org/10.1007/978-3-642-20633-7_34

Lukyanenko R, Parsons J and **Wiersma Y.** 2014. The impact of conceptual modeling on dataset completeness: a field experiment. *Proceedings of the International Conference on Information Systems 2014*. Atlanta: Association for Information Systems.

MacKenzie, CM, Murray, G, Primack, R and **Weihrauch, D.** 2017. Lessons from citizen science: assessing volunteer-collected plant phenology data with Mountain Watch. *Biological*

conservation, 208: 121–126. DOI: https://doi.org/10.1016/j.biocon.2016.07.027

Manni, M, Le, Y, Vander Stoep, G and **Hollenhorst, SJ.** 2012. Denali National Park and Preserve Visitor Study: Summer 2011. Fort Collins, Colorado.

National Academies of Sciences Engineering, & Medicine, Committee on Designing Citizen Science to Support Science Learning, Board on Science Education, Division of Behavioral, Sciences, S., Education. 2018. Learning Through Citizen Science: Enhancing Opportunities by Design (R. Pandya & K. A. Dibner, Eds.). National Academies Press. DOI: https://doi.org/10.17226/25183

Newman, G, Wiggins, A, Crall, A, Graham, E, Newman, S and **Crowston, K.** 2012. The future of citizen science: Emerging technologies and shifting paradigms. *Frontiers in Ecology and the Environment*, 10(6): 298–304. DOI: https://doi.org/10.1890/110294

Rasmussen, ZA. 2015. Coyotes on the Web: Understanding Human-Coyote Interaction and Online Education Using Citizen Science. PhD diss., Portland State University.

Ratnieks, FL, Schrell, F, Sheppard, RC, Brown, E, Bristow, OE and **Garbuzov, M.** 2016. Data reliability in citizen science: learning curve and the effects of training method, volunteer background and experience on identification accuracy of insects visiting ivy flowers. *Methods in Ecology and Evolution*, 7(10): 1226–1235. DOI: https://doi.org/10.1111/2041-210X.12581

Ripley, BD. 1977. Modeling spatial patterns (with discussion). *Journal of the Royal Statistical Society*. B 39: 172–212. DOI: https://doi.org/10.1111/j.2517-6161.1977.tb01615.x

Schaible, B, Frank A, Seybold, D and **Calhoun, BC.** 2020. Ecological Study of Effects of Industrial Watershed on Müllerian Anomalies in an Obstetric Population. *Ecotoxicology and Environmental Safety,* 202: 10819. DOI: https://doi.org/10.1016/j.ecoenv.2020.110819

Senaratne, H, Mobasheri, A, Ali, AL, Capineri, C and **Haklay, M.** 2016. A review of volunteered geographic information quality assessment methods. *International Journal of Geographical Information Science*, 1–29. DOI: https://doi.org/10.1080/13658816.2016.1189556

Teixeira-Neto, RG, da Silva, ES, Nascimento, RA, Belo, VS, de Oliveira, CDL, Pinheiro, LC and **Gontijo, CMF.** 2014. Canine visceral leishmaniasis in an urban setting of Southeastern Brazil: an ecological study involving spatial analysis. *Parasites & vectors*, 7(1): 485. DOI: https://doi.org/10.1186/s13071-014-0485-7

Toomey, AH and **Domroese, MC.** 2013. Can citizen science lead to positive conservation attitudes and behaviors? *Human Ecology Review*, 20(1): 50–62.

Tulloch, AIT, Possingham, HP, Joseph, LN, Szabo, J and **Martin, TG.** 2013. Realising the full potential of citizen science monitoring programs. *Biological Conservation*, 165: 128–138. DOI: https://doi.org/10.1016/j.biocon.2013.05.025

Ward, EJ, Marshall, KN, Ross, T, Sedgley, A, Hass, T, Pearson, SF, Joyce, G, Hamel, NJ, Hodum, PJ and **Faucett, R.** 2015. Using Citizen-Science Data to Identify Local Hotspots of Seabird Occurrence. *PeerJ*, 3(1). DOI: https://doi.org/10.7717/peerj.704

Weaver, D. 2015. Volunteer tourism and beyond: motivations and barriers to participation in protected area enhancement. *Journal of Sustainable Tourism*, 23(5): 683–705. DOI: https://doi.org/10.1080/09669582.2014.992901

Weiler, B and **Richins, H.** 1995. Extreme, Extravagant and Elite. *Tourism Recreation Research*, 20(1): 29–36. DOI: https://doi.org/10.1080/02508281.1995.11014730

Wiggins, A, Bonney, R, LeBuhn, G, Parrish, JK and **Weltzin, JF.** 2018. A Science Products Inventory for Citizen-Science Planning and Evaluation. *BioScience*, 68(6): 436–444. DOI: https://doi.org/10.1093/bioscience/biy028

Wiggins, A, Newman, G, Stevenson, RD and **Crowston, K.** 2011. Mechanisms for data quality and validation in citizen science. *Proceedings - 7th IEEE International Conference on e-Science Workshops. EScienceW 2011*, 14–19. DOI: https://doi.org/10.1109/eScienceW.2011.27

Zhang, B, Xiaozhen L, Jiang J, DeAngelis, DL, Fu, Z and **Zhang, J.** 2017. Similarity of Plant Functional Traits and Aggregation Pattern in a Subtropical Forest. *Ecology and Evolution*, 7(12): 4086–98. DOI: https://doi.org/10.1002/ece3.2973

]u[🔓