International Journal of Zoology

The Role of Citizen Science and Volunteer Data Collection in Zoological Research

Guest Editors: Adam Hart, Richard Stafford, Anne Goodenough, and Simon Morgan

The Role of Citizen Science and Volunteer Data Collection in Zoological Research

International Journal of Zoology

The Role of Citizen Science and Volunteer Data Collection in Zoological Research

Guest Editors: Adam Hart, Richard Stafford,
Anne Goodenough, and Simon Morgan

Editorial Board

Contents

Hindawi Publishing Corporation
International Journal of Zoology
Volume 2012, Article ID 105345, 3 pages
doi:10.1155/2012/105345

Editorial

The Role of Citizen Science and Volunteer Data Collection in Zoological Research

Adam Hart,[1] Richard Stafford,[2] Anne Goodenough,[1] and Simon Morgan[3]

[1] *School of Natural and Social Sciences, University of Gloucestershire, Cheltenham GL50 4AZ, UK*
[2] *School of Applied Sciences, Bournemouth University, Fern Barrow, Poole, Dorset BH12 5BB, UK*
[3] *Wildlife Act Breda Park, 27 Breda Street, Gardens, Cape Town 8001, South Africa*

Correspondence should be addressed to Adam Hart, ahart@glos.ac.uk

Received 6 November 2012; Accepted 6 November 2012

In many ways, science has never been as popular as it is now. With an ever-increasing number of popular science books on everything from astronomy to climate change and evolution and entire TV channels devoted to science output, the public seems spoilt for choice. However, paradoxically, there is also an increasing disconnect between science—and scientists—and society, and this is certainly evident in the life sciences. This disconnect comes in two forms: interest and level of knowledge. Indeed, one has only to look at the 2012 US presidential election campaign to see the lack of scientific knowledge possessed by many of the political elite about topics such as climate change. If high profile scientific topics are still so widely misunderstood by those in the public eye, it is unsurprising that there is such a lack of understanding of, and interest in, scientific topics in the general public. It should, in theory, be the easiest to address this discontent in subjects like zoology, where the evidence is all around us and can be easily seen, appreciated, and studied by the world's citizens.

Citizen science makes use of "citizens" (whether members of the general public of dedicated volunteers) rather than professional scientists to undertake data collection for, often, large-scale scientific studies. We argue that recent technological developments mean that citizen science has never had as much relevance, or potential, as it does right now. Furthermore, we assert that, given a general public interest in animals, citizen science in zoology can dramatically improve public scientific literacy and provide a gateway into the serious consideration of the many complex scientific issues pertinent to the modern world.

Citizen science in zoology is a product of our time. Historically, there was no such thing as a "professional" naturalist, and nonprofessionals readily collected data and published in major scientific journals well into the 20th century. This changed to a great extent in the mid-to-late 20th Century, as research in zoology, related disciplines like ecology, and subdisciplines like ethology, became hypothesis and experiment driven with increasingly sophisticated experimental designs, equipment, and statistical analyses. Such changes were necessary to understand such diverse topics as the neurological basis of behaviour and the ecological mechanisms controlling community structure. However, with these changes came the inevitable decline of the interested amateur in the collection and analysis of scientific data.

This polarisation of the amateur and the professional, as well as increasing compartmentalisation of topics, parallels other scientific disciplines. Indeed, the increasing fragmentation of the traditional sciences of Biology, Chemistry, and Physics into ever-increasing numbers of subdisciplines in recent years has meant that even professional scientists often know little or nothing about other areas. It is hardly surprising that the public often feels disconnected with modern science.

Citizen science in zoology, and other disciplines, has two clear roles. The first is to allow the collection of large amounts of data, which can then be analysed to test scientific hypotheses. The development of the internet, the ubiquity of internet-enabled devices, the rise of social media, and the development of increasingly sophisticated mobile telephones have given the scientific world the potential to connect with the public in easy and exciting ways and to access a huge pool of volunteer data collectors. However, despite clear advantages, citizen science approaches have numerous drawbacks related to, amongst other things, the reliability of data (especially when members of the public require specific skills, such as accurate species identification) and the statistical issues that can arise from the analysis of presence-only data (i.e., because species absences are not usually recorded except in full scientific surveys, such that citizen science data can be prone to false absences). If citizen science is to become a reputable approach, it is necessary to acknowledge these limitations, confront them head-on, and devise ways around them that enhance the "science" but do not diminish the role of the "citizen".

In this special issue, we evaluate the current role, and future potential, of citizen science and volunteer-based projects for the collection of publishable data. With many different groups taking on citizen science projects with a varying level of success, many lessons are being learned locally; but the difficulty of publishing negative results, or studies that are more informative about the approach than the results, means that such lessons are not being shared. The papers assembled in this issue are a clear testament to the success of citizen science approaches to collect scientifically useful data. For example, J. Zelt et al. [1] show how phenology patterns of birds have changed from a comparison with historic records and R. L. Davis et al. [2] indicate how citizen science data can be used to determine butterfly habitat use. By exploring the limitations of citizen science approaches, R. L. Williams et al. [3] provide a valuable lesson in how volunteers need careful management to collect accurate data. In this case, the public was asked to collect behavioural data but the limitations of this approach highlighted by R. L. Williams et al. arose not because the public cannot reliably recognise and record different behaviours, but because they did not understand scientific concepts such as "fixed effort sampling." Continuing this theme, L. K. Higby et al. [4] describe how some of the problems of zero inflated datasets can be overcome using citizen science data, in this case using one of the most charismatic macro vertebrates of all, whales. While much of this analysis may be lost in the general public, simplified findings can be communicated back to demonstrate the usefulness of what they are doing, providing an all-important positive feedback loop and encouraging further participation.

However, the second, and in many cases not less important, role of citizen science is in engaging, and educating, the "citizens." Affection for the natural world is normally fostered through a love of a particular environment (e.g., coral reefs) or taxonomic group (e.g., birds). Taking part in citizen science surveys, and the online publishing of data from these surveys, can be an active, and first time, immersion into the scientific process. In education, the concept of "active learning" has been well developed and has been demonstrated to create a deep understanding of the issues being studied [5]. The same is likely to be true for partaking in many citizen science projects. For example, it is likely to be members of the public that initially observe the sprouting of leaves on trees at an earlier date than in previous years, and these observations are likely to create far more public belief in the biological effects of climate change than the results of traditional scientific study that filters down to a disengaged public. Citizen science projects have also provided evidence of evolution in banded snails [6] and although the scientific analysis is complex, presentation of the data in a simple way, especially to the participants of the project, surely increases their understanding of evolution. For example, the results are used as an example of selection in action in schools in the UK, through a lesson plan by the Nuffield Foundation. R. L. Williams et al. [3] may highlight limitations of public data collection in some circumstances, but crucially they provide guidance on how to improve data collected by the public and this project provided an excellent medium for environmental education and generating enthusiasm. Whilst the primary role, and goal, of citizen science projects is to collect scientifically useful data, raising awareness, sharing knowledge, and generating enthusiasm are useful and valid secondary aims.

Within this issue, the need for "citizen" engagement is made clear. C. L. Catlin-Groves [7] illustrates many new novel technologies to further engage the public in scientific research (and improve data quality along the way) and H. R. Cunningham et al. [8] illustrate the role of citizen science projects in educating the public about some less charismatic species—in this case amphibians and reptiles—as well as the role of these species in general ecology. J. A. Oldekop et al. [9] show how initial training from scientists can be passed on through native communities in the Amazon. Finally, M. J. C. Crabbe [10] explores how citizen scientists from both developed and developing countries can combine knowledge and expertise, to collect data that is useful for the community and can be used to formulate policy in developing countries.

In short, citizen science, as amply demonstrated by this special issue, can provide useful scientific data, engage the public, and become a powerful tool in the public understanding of science. In a world where scientific "impact" is becoming more and more important, we cannot afford to overlook the multiple roles that citizen science can play.

Adam Hart
Richard Stafford
Anne Goodenough
Simon Morgan

References

[1] J. Zelt, J. Courter, A. Arab, R. Johnson, and S. Droege, "Reviving a legacy citizen science project to illuminate shifts in bird phenology," *International Journal of Zoology*, vol. 2012, Article ID 710710, 6 pages, 2012.

[2] A. K. Davis, N. P. Nibbelink, and E. Howard, "Identifying large- and small-scale habitat characteristics of monarch butterfly migratory roost sites with citizen science observations," *International Journal of Zoology*, vol. 2012, Article ID 149026, 9 pages, 2012.

[3] R. L. Williams, S. K. Porter, A. G. Hart, and A. E. Goodenough, "The accuracy of behavioural data collected by visitors in a zoo environment: can visitors collect meaningful data?"

International Journal of Zoology, vol. 2012, Article ID 724835, 13 pages, 2012.

[4] L. K. Higby, R. Stafford, and C. G. Bertulli, "An evaluation of ad hoc presence-only data in explaining patterns of distribution: cetacean sightings from whale-watching vessels," *International Journal of Zoology*, vol. 2012, Article ID 428752, 5 pages, 2012.

[5] K. Exley and R. Dennick, *Giving a Lecture: From Presenting to Teaching*, Routledge, Abingdon, UK, 2nd edition, 2009.

[6] J. Silvertown, L. Cook, R. Cameron et al., "Citizen science reveals unexpected continental-scale evolutionary change in a model organism," *PLoS ONE*, vol. 6, no. 4, Article ID e18927, 2011.

[7] C. L. Catlin-Groves, "The citizen science landscape: from volunteers to citizen sensors and beyond," *International Journal of Zoology*, vol. 2012, Article ID 349630, 2012.

[8] H. R. Cunningham, C. A. Davis, C. W. Swarth, and G. D. Therres, "The Maryland amphibian and reptile atlas: a volunteer-based distributional survey," *International Journal of Zoology*, vol. 2012, Article ID 348653, 11 pages, 2012.

[9] J. A. Oldekop, N. K. Truelove, S. Villamarn, and R. F. Preziosi, "Information flows in community-based monitoring exercises in the ecuadorian amazon," *International Journal of Zoology*, vol. 2012, Article ID 980520, 4 pages, 2012.

[10] M. J. C. Crabbe, "From citizen science to policy development on the coral reefs of Jamaica," *International Journal of Zoology*, vol. 2012, Article ID 102350, 6 pages, 2012.

Hindawi Publishing Corporation
International Journal of Zoology
Volume 2012, Article ID 349630, 14 pages
doi:10.1155/2012/349630

Review Article

The Citizen Science Landscape: From Volunteers to Citizen Sensors and Beyond

Christina L. Catlin-Groves

Department of Natural and Social Sciences, University of Gloucestershire, Cheltenham GL50 4AZ, UK

Correspondence should be addressed to Christina L. Catlin-Groves, c.catlin.groves@googlemail.com

Received 1 March 2012; Accepted 22 June 2012

Academic Editor: Simon Morgan

Within conservation and ecology, volunteer participation has always been an important component of research. Within the past two decades, this use of volunteers in research has proliferated and evolved into "citizen science." Technologies are evolving rapidly. Mobile phone technologies and the emergence and uptake of high-speed Web-capable smart phones with GPS and data upload capabilities can allow instant collection and transmission of data. This is frequently used within everyday life particularly on social networking sites. Embedded sensors allow researchers to validate GPS and image data and are now affordable and regularly used by citizens. With the "perfect storm" of technology, data upload, and social networks, citizen science represents a powerful tool. This paper establishes the current state of citizen science within scientific literature, examines underlying themes, explores further possibilities for utilising citizen science within ecology, biodiversity, and biology, and identifies possible directions for further research. The paper highlights (1) lack of trust in the scientific community about the reliability of citizen science data, (2) the move from standardised data collection methods to data mining available datasets, and (3) the blurring of the line between citizen science and citizen sensors and the need to further explore online social networks for data collection.

1. Introduction

Within conservation and ecology, volunteer participation has always been an important component of research [1–5]. Within the past two decades, use of volunteers in research has begun to proliferate and evolve into the current form of "citizen science" [6, 7]. Citizen science, a term first coined by Irwin [7], is used to describe a form of research collaboration or data gathering that is performed by untrained or "nonexpert" individuals, often involving members of the public, and frequently thought of as a form of crowd-sourcing [1, 8–12].

Citizen science will usually incorporate an element of public education [2, 6, 13–15]. Silvertown [5] described the differentiation between historical and modern forms of citizen science by potential for it to be "available to all, not just a privileged few." This has been recently demonstrated by the rapid development of mobile phone technologies, in particularly the emergence and uptake of high-speed Web-capable smart phones with GPS data collection facilities and data upload capabilities [16]. This allows almost instant collection, transmission, and submission of data and provides researchers with a way to validate data (e.g., to verify the identification of an organism or the location through GPS locators) [10]. The availability of new technologies containing sensors could be argued to move citizen science into a new era whereby citizen scientists also become citizen "sensors." Collection of high-quality data can be made through the sensing capabilities of personal computing and communication technologies, making the user part of a more passive framework for data collection [17–19]. Some of the key strengths of citizen science projects lie in the ease and speed with which data can be gathered by a large number of individuals in a short time. Ordinarily constraints such as money and time would make studies unfeasible or impossible for an individual organisation [10, 15, 20]. Indeed, citizen science programmes are often more resilient to variations in financial support than other programs [19, 21, 22].

With technological connectivity peaking, the ability to select virtual "field assistants," to help gather data is within easy reach; indeed Irwin [7] said that citizen scientists can be considered as the "world's largest research team." A further step is the potential for mining data, for ecological or biological research, from the huge quantities of data which are voluntarily uploaded onto personal social media accounts for the primary reason of storage or sharing with friends. For example, there are over 26,000 images tagged with "manta ray" on Flickr (as of December 12th 2011), a species with stable patterning that can be individually identified [23]. Custom Application Programming Interfaces (API) could theoretically identify these individuals and collect GPS data (where its available). This would create vast quantities of ecological and spatial data that could be utilised in research which tracks individuals. However, despite this, citizen science projects are often limited to: (1) informal education activities or outreach to promote understanding [1, 6, 14, 24, 25]; (2) natural resource monitoring to promote stewardship [26–28]; (3) to promote social activities and action [29, 30]; (4) purely virtual whereby the entire project is ICT-mediated with no physical attribute (e.g., classifying photographs) ([31, 32], see Table 1). Table 2 provides examples of citizen science projects alongside their primary goals.

Few scientific investigative projects exist in ecology or biology using these new technologies for data collection, and where they do, they often encounter difficulties with gaining robust data [5, 6]. Even less take advantage of the rapidly increasing and evolving capabilities of Web 2.0 and social networks such as Facebook (http://www.facebook.com/), Twitter (http://www.twitter.com/) and Flickr (http://www.flickr.com/), through which millions of people upload and share photographs and location data, many citizen science studies also concentrate on data being collected within a very rigid framework, very similar to previous volunteer data collection whereby paper forms are replaced by online submission forms (examples include e-bird, Project Budburst, What's Invasive, and Neighbourhood Nestwatch Program). The possibility of using Web 2.0 and less rigid data collection techniques is relatively underexplored within scientific literature, even less so for biological and ecological applications.

This paper will predominantly cover the uses of citizen science for ecology, biodiversity, and biological insights. However, it may touch on various interdisciplinary citizen science programs or concepts where it is felt that it will be beneficial and may bring together other approaches which may add value. The aim is to establish the current state of citizen science within scientific literature, examine main underlying themes, and explore the possibility of utilising an untapped resource and the benefits that this can hold for the scientific community. It will also attempt to identify possible directions for further research.

2. The Citizen Science Landscape

The ability for intense monitoring by expert individuals on any subject ranging from individual or species distributions to tracking of invasive species is severely limited by both logistical and financial constraints. There are simply not enough resources, whether this is in the form of time, personnel, or money to establish large scale datasets [6, 10, 33, 34]. Citizen science circumvents many of these problems and has proven effective in a number of research areas that can have difficulty gathering large datasets. The areas in which it has, and most probably will continue to have, the greatest impact and potential are that of monitoring ecology or biodiversity at large geographic scales (see Table 3 for examples). This is particularly prevalent due to the recent proliferation of built in GPS technology and Web-capable features that many handheld devices, such as mobile phones and increasingly cameras, now have in an affordable and widely available format [10, 11].

When monitoring for rare, unusual, or declining phenomena, the scale of a large workforce over a large area will increase rates of detection in comparison to a lone researcher on a strict rotation despite having greater expert knowledge [35]. Indeed in early 2006, the rare nine-spotted ladybird (*Coccinella novemnotata*) was rediscovered during a citizen science programme designed to educate the public in biodiversity and conservation. This nine-spotted ladybird was the first discovered in eastern North America in over fourteen years, and only the sixth in the whole of North America within 10 years [36].

Traditional citizen science or volunteer programs have resulted in some of the longest ecological temporal datasets that we can access, particularly in the field of ornithology. The *Christmas Bird Count* (CBC—http://birds.audubon.org/christmas-bird-count/) was launched 1900 by the Audubon Society (in US and Canada) and provides long-term comprehensive data trends for many species for over 100 years. The British Trust for Ornithology, founded in 1932, also regularly uses data collected by amateur birdwatchers and makes up a very substantial amount of the National Biodiversity Network (http://www.nbn.org.uk/) which contains over 31 million records. The data of these programmes have helped to inform conservation actions, for example, by providing information to target conservation management at particular sites by environmental organisations [37].

Citizen science programmes conducted in the last 10 years have successfully followed the spread of invasive species or diseases, impacts of land use or climate change, and have been instrumental in understanding distributions, ranges, and migration pathways (e.g.,[38, 39]). Researchers at Cornell University, USA, have performed a large range of citizen science projects centred around avian species. Some of these projects have resulted in datasets that track the spread of conjunctivitis (*Mycoplasma gallisepticum*) in wild house finches (*Carpodacus mexicanus*)[40] and the impact of forest fragmentation on tanager populations and nesting success [41]. These efforts have led to a large database called eBird, where amateur birdwatchers can upload sightings. These citizen science data have become the basis of trends discovered through data mining and modelling techniques,

Table 1: Citizen science typologies as described by Wiggins and Crowston [1].

Type	Description	Example
Action	Employ volunteer-initiated participatory action research to encourage participant intervention in local concerns.	Shermans Creek Conservation Association (http://www.shermanscreek.org/)
Conservation	Address natural resource management goals, involving participants in stewardship for outreach and increased scope.	Missouri Stream Team Project (http://www.mostreamteam.org/)
Investigation	Focus of scientific research goals focussed on collecting data from the physical environment, usually underpinned by an hypothesis or research goal.	BirdTrack (http://www.bto.org/volunteer-surveys/birdtrack)
Virtual	Similar goals to the investigation project, but are entirely mediated by ICT having no physical element.	Whale FM (http://whale.fm/)
Education	Education and outreach are their primary goals, often data is not collected in a meaningful way that might be useful to other researchers. Often provides formal and informal learning resources.	Bird Sleuth (http://www.birds.cornell.edu/birdsleuth)

Table 2: Primary goals of citizen science projects (adapted and modified from Wiggins and Crowston [1]).

Project	URL	Primary goal	Description
Globe at Night	http://www.globeatnight.org/	Education	Learning about light pollution with use of mobile phone or Web cam and internet connection.
Fossil Finders	http://www.fossilfinders.org/	Education	Learning about Devonian Fossils through authentic inquiry-based investigation.
Bird Sleuth	http://www.birds.cornell.edu/birdsleuth/	Education	Learning about birds through inquiry-based investigation.
Missouri Stream Team Project	http://www.mostreamteam.org	Conservation	Promotes the formation of “stream teams” which monitor streams in their area.
What’s Invasive	http://whatsinvasive.com/	Conservation	Locating invasive plants.
Shermans Creek Conservation Association	http://www.shermanscreek.org/	Action	Started to oppose the building of a power plant on local land, they now monitor the area and have regular talks.
ReClam the Bay	http://www.reclamthebay.org/	Action	Promotes environmental involvement by growing and maintaining baby clams and oysters to stock their local bay.
*Whale FM	http://whale.fm/	Virtual	Asks participants to listen to and classify whale song.
*Galaxy Zoo	http://www.galaxyzoo.org/	Virtual	Invites participants to classify images of galaxies.
Pathfinder	http://www.pathfinderscience.net/	Virtual	Collaborative online environment for citizen scientists.
Foldit	http://www.fold.it/	Virtual	Proving human superiority at protein folding.

*Part of Zooniverse—https://www.zooniverse.org/projects—citizen science hub for virtual citizen science projects exploiting the human ability to spot patterns and classify data where traditional statistical analysis struggles.

which have led to further more focussed studies (visit http://ebird.org/content/ebird/about/ebird-publications for more information and a full list of publications).

Datasets that have been gathered for a specific purpose will often result in unexpected phenomena or patterns emerging, that will then promote further more focussed studies. Many studies are available in scientific literature where data mining and model construction have resulted in the discovery of new patterns and processes being found in ecological systems (e.g.,[42–44]). Howard and Davis [45, 46] have published a number of peer-reviewed papers on data predominantly collected by citizen scientists, gathering useable scientific data on autumn migration flyways of monarch butterflies (*Danaus plexippus*). Citizen scientists record overnight roosts and report their first spring sightings to assess spring recolonisation rates.

Table 3: Citizen science projects and data collection/submission process(s).

Project	URL	Description	Data collection/submission process(s)
Project PigeonWatch	http://www.birds.cornell.edu/pigeon watch	A US program run by Cornell University. Participants count pigeons and record courtship behaviours observed in their neighbourhood pigeon flocks.	Virtual form submission
eBird	http://ebird.org/	Initially US-ebased but moving more into global records. eBird's goal is to maximize the utility and accessibility of the vast numbers of bird observations made each year by recreational and professional bird watchers. Has an online accessible database and visualisation facilities for the participant and other interested parties.	Virtual form submission
Ecocean	http://www.whaleshark.org/	The ECOCEAN Whale Shark Photo-identification Library is a visual database of whale shark (*Rhincodon typus*) encounters and of individually catalogued whale sharks. It asks participants to upload images and sightings of Whale Sharks.	Virtual form submission
Natures notebook	http://www.usanpn.org/participate/observe	A US program run as part of the National Phenology Network, it asks people to report the phenophases of particular species in their local areas.	Virtual form submission
BirdTrack	http://www.bto.org/volunteer-surveys/birdtrack	Partnership working between the British Trust of Ornithology, Royal Society for the protection of Birds, Birdwatch Ireland, and the Scottish Ornithologists' Club, it collects data on migration movements and distributions throughout Britain and Ireland. Has an online accessible database and visualisation facilities for the participant and other interested parties.	Virtual form submission
British Trust for Ornithology	http://www.bto.org/	Nongovernmental organisation dedicated to using volunteers who follow statistically designed sampling strategies in their research into birds.	Virtual form submission
Project Budburst	http://neoninc.org/budburst/	A US project participants observe plant phenophases. Scientists can use the data to learn more about the responsiveness of individual plant species to changes in climate locally, regionally, and nationally.	Virtual form submission and mobile application submission
What's Invasive	http://whatsinvasive.com/	Asks participants to locate invasive species by making geotagged observations and taking photos to map their spread.	Virtual form submission and mobile application submission
Neighbourhood Nestwatch	http://nationalzoo.si.edu/scbi/MigratoryBirds/Research/Neighborhood_Nestwatch/default.cfm	Participants find and monitor bird nests and record and report their observations. Researchers are especially interested in comparing how successful nests are in urban, suburban, and rural backyards.	Virtual form submission
BeeID	http://www.flickr.com/groups/beeid	A completed pilot project which asked participants to upload geo-tagged images of bees and tag them with "beeid2010" tag on the Flickr photography Website, researchers then extract these images, identify and tag with the species id.	Tagging and data mining from existing social network site with integral mobile upload facilities

One of the common features of traditional and many current projects is the formal submission process which occurs on a stand-alone Website or through one-to-one communication between researcher and citizen. The submission is often closed or inaccessible until a result is published, and even in citizen science programmes where data is shared: it is very difficult for the ordinary citizen to visualize; this is shown to have an impact on participation [6, 47]. e-Bird has gone through some lengths to overcome this. By creating an online database system which has many portals and visualisation techniques, citizen scientists and researchers alike can explore the e-Bird database [12]. In April 2006, when this newly improved Website was upgraded allowing participants to explore their own and others data, the number of individuals submitting data nearly tripled [47]. Resources such as this require the citizen scientist to make an active effort to discover the project, find the Website and input, and retrieve data. By integrating data collection into social media and fully exploiting Web 2.0, the quality, geographical range, and quantity of data collected could potentially be significantly increased, and this is something that requires further research. However, despite the lack of financial cost that social media and Web 2.0 present, it is possible that the time and effort cost might not make the process worthwhile when considering the amount of additional data gained.

3. Social Networks and Web 2.0

Web 2.0 is an ambiguous term with almost as many facets and conflicting opinions and definitions as the term citizen science, some even argue against the existence of Web 2.0 as a concept. However, for the purposes of this paper Web 2.0 can be regarded as the socially connected and interactive internet which facilitates participatory data sharing and encourages user-generated content. This medium consists of blogs, podcasts, social networking sites, wikis, crowd-sourcing tools, and "cloud-based" group working environments. Web 2.0 has been expanded to a mobile computing context with the proliferation of new technologies such as smart phones, laptops, and tablet computers [48].

The most obvious purpose for exploiting Web 2.0, which is beginning to be used by researchers, is the power of marketing and advertising, expressing branding, recruiting, retaining, and sharing, and collecting data with the citizen scientist [5, 49]. Delaney et al. [50] advocated the use of Web 2.0 capabilities for the ease of collecting and sharing data via new cloud technologies. Delaney suggests that dynamic linked databases that use online mapping technology such as Google Earth (free and familiar to citizen scientists) would prove ideal for creating a complete graphical "global" database of species. This would likely increase engagement and retention of individuals as they watch their contributions become part of the "bigger picture." In essence, social media is being adopted as part of the communication strategy for engaging individuals who collect data or participate in virtual citizen science programs; this adoption is seemingly in line with that of organisations at large to promote products or engage audiences. This paper will not examine these factors in depth, as they are far too large to be able to cover appropriately (for more further information on this topic, (see [51–53])). Society at large is beginning to understand the increased power of "the social network effect" behind Web 2.0, which increases value to existing users in a feedback loop (e.g., more and more users begin to embrace a service, increasing its popularity, and resulting in rapidly increasing adoption) [54–56].

Figure 1 shows a brief diagram of citizen science. In addition to running programs of research that encourage users to engage in a more traditional data submission process, there is also the underexplored option of mining data from social networks and taking a more opportunistic approach. Indeed, many images, especially those taken on mobile phones, contain GPS information and can readily be searched and mapped via the integrated search facilities on Websites. The mobile interface allows the mobile phone to become a people-centric sensor which is capable of aggregating inputs from local surroundings, enabling data to be collected at a higher resolution [57]. This may be useful in plotting distributions and migration patterns or movements, both of individuals or species. Indeed, large charismatic species with stable patterning such as whales, sharks, rays, and big cats are photographed regularly by tourists and shared online, and the ability to collate and analyse these images could prove valuable to the study of their movement, social grouping, and ultimately conservation.

An emerging and particularly promising but under developed area of citizen science is that of using online social networking sites such as Facebook, Twitter, and mobile social networks such as Foursquare. Many of these have integrated image and location data upload facilities. Indeed, throughout 2011 there has been a proliferation of these facilities throughout popular social networking Websites. These features have been incorporated into basic interfaces, enabling users to simultaneously capture images; GPS tag them, add, comments, and post, to followers or friends instantly via mobile internet.

Since it's advent in 2004, Facebook (http://www.facebook.com/), the most popular social networking site, has grown to having more than 800 million active users globally, with, on average, more than 250 million photographs uploaded every day. More than 350 million people access it through a mobile phone [58]. Research by commercial online marketing and data collection agency comScore Media Metrix suggested that Facebook reached 73% of Americans in June 2011 [59]. With Flickr, the story is similar; Yahoo! announced in August 2011 that it had reached 51 million users and had, on average, 4.5 million photos uploaded every day. On the February 28th 2012 it had 176,605,443 geotagged photographs in total. With an integrated approach and the correct marketing and publicity, in addition to the increase of GPS-capable mobile devices it is likely that Flickr may become increasingly useful for gathering data, particularly for charismatic species.

The potential for scientific research is immense, particularly for image-based data collection where EXIF information can be mined using a custom API and identification can

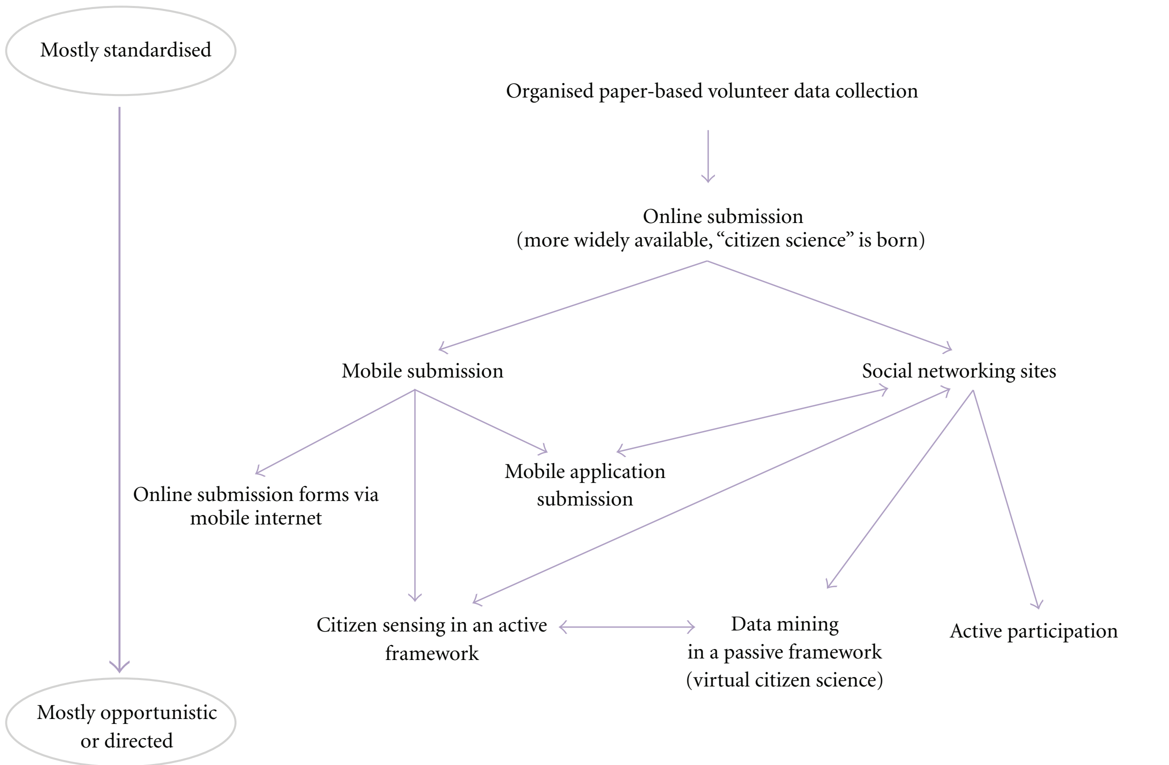
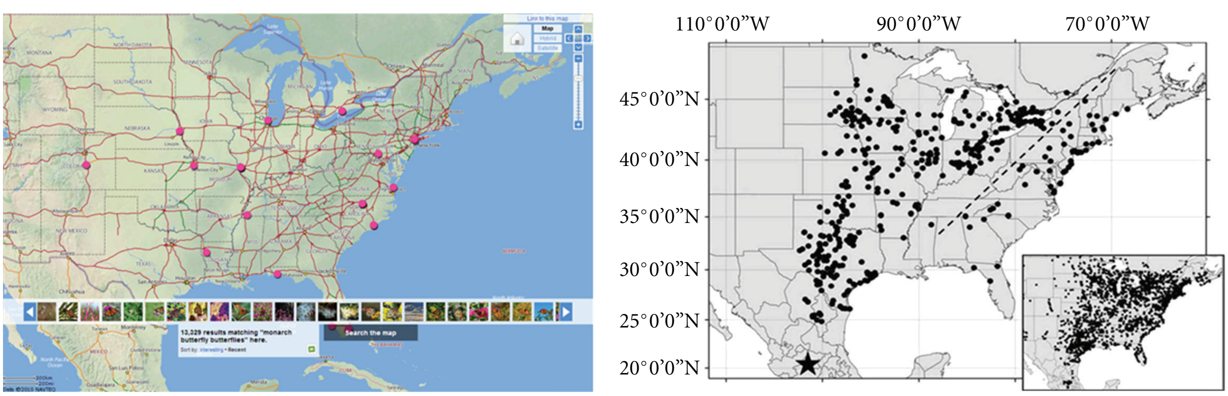


Figure 1: A brief diagram of citizen science, the diagram shows the proliferation in citizen science as new technologies have become available.



(a) (b)

Figure 2: (a) Screenshot of Flickr map displaying 250 of 13,329 geo-tagged photos tagged with "monarch butterflies butterfly" on February 28th 2012. (b) Map of Journey North roost sightings from all years combined (2005–2007). Dashed line indicates division of central and eastern flyways in analysis. Roosts in Florida were not included in the analyses. Inset map shows the locations of all Journey North participants from 1997 to 2007. Star indicates location of Mexico overwintering sites [45]. Reproduced with kind permission of Springer Science and Business Media.

be verified by trained individuals or automatic recognition software [60–63]. Figure 2 shows a small example of what Flickr can do with a simple search term "monarch butterflies butterfly" (signifying a search for either butterflies or butterfly) which pulls up 13,329 geo-tagged photos within the US (February 28th 2012), 250 of which it can plot on a map on the Flickr Website. The map seemingly holds a cursory resemblance to Howard and Davis's [45] map of monarch butterfly migration roosts (created using data from a citizen science program called Journey North which relies on a more traditional data submission process albeit via an online form—http://www.learner.org/jnorth/). Using a custom API and transposing all results onto Google Maps or other mapping software, it would be possible to limit the geo-tagged photo search by date and compare it directly with Journey North's monarch butterfly monitoring program, which has received 4078 sightings within the last year. However, without creating an API, a simple search on Flickr's advanced search facility with the search term "monarch butterfly" brought up 15,499 photographs within the same time period (using data collected on February 28th 2012). Despite being likely that a large proportion contains no useful information (i.e., not pictures of the target species) and/or is not geo-tagged, (although estimates show >40% may be geo-tagged, [64]), this suggests that if this method of data collection was further explored the number of potentially useful monarch butterfly sightings data could be greatly increased.

Currently, a general internet user's image and location uploads are predominantly limited to "events" that the user wants to share this might be "checking in" to restaurants, attractions, clubs, cinemas, or concerts, often reviewing products, or sharing visual experiences [65, 66]. Sharing these data with another user can be as simple as tagging them [66]. By exploiting social networks in this way, for ecological or biological research, many of the most common mistakes or inaccuracies that are found within volunteered data could be minimised. For example, by sharing images, and temporal and GPS data, misidentifications and location inaccuracies can be flagged and checked by trained individuals [5, 37, 64, 67, 68]. Despite this, there are very few examples of social networking sites being used actively to collect data for biological or ecological research; this may be because of confusion over copyright laws or limitations of API systems. At time of writing, there are very few examples of such usages, and those few that do exist are limited to self-contained "groups" within Flickr which search images of individual animals to export to an external catalogue for identification or use them to advertise the program and attract new submissions (Table 4).

Despite this ability to gather data quickly, they are currently underutilised for ecological or biodiversity data collection. BeeID is a program of research which used Flickr as a base for data collection [64, 67]. Researchers asked individuals to tag photographs of bees with specific searchable metatags and place location data on them if it was not already embedded. Trained individuals then confirmed species identification and marked the images as processed via the addition of a new tag. A simple custom API extracted tagged photographs from Flickr and collected the data which successfully plotted bee species distributions. Considering the project had no funding and was run by a small group of individuals with limited promotion other than on social networking sites, its success demonstrates the potential benefits of using social networking for collection of scientific data. Furthermore, the study took part before the recent integration of easily accessible location data in social networks and the continued rise of smart-phone and affordable GPS and wifi enabled camera ownership.

Another facet of Web 2.0 is the very recent addition of phone applications or "apps." These are easily integrated and simple to use; however, the release of a mobile application is not enough on its own to motivate participants and it is important to use mobile applications in an holistic approach [12]. "What's Invasive" is a very recent citizen science programme which uses a combination of a Website and custom mobile application to allow mobile devices to collect and submit information about invasive species whilst they are observing them (http://whatsinvasive.com/). Project Noah is similar in that respect but is built primarily to engage and educate individuals in addition to collecting species data through a tagging and classification system. Project Noah also incorporates "missions" to increase motivation and promote the collection of specific species sightings (http://www.projectnoah.org/).

A recently developed formatting language, Hypertext Markup Language 5 (HTML5), allows easier development across platforms and allows many of the features of mobile phone applications to be incorporated into Websites. Web pages can then be developed to contain full multimedia content that is easily accessible to popular technologies, something which some smart phones have found problematic due to limited Flash support (especially on Apple devices). In the past, this inability has limited some of the content available and increased the amount of work needed to replicate Web pages on smart-phones.

Undoubtedly, with the advent of Web 2.0 and the quickly developing technological breakthroughs, citizen science programs exploiting this technology are likely to increase exponentially in future years and should be encouraged. It is hoped that as the full potential is revealed the negative bias among the scientific community that such approaches have attracted will begin to lessen. As the population increases and we are more isolated from nature and wildlife, the use of citizen science for biodiversity studies will enable individuals to be further engaged in decision-making processes and the championing and protection of the natural environment. It is a paradigm that is evolving alongside our relationship with technology, our environment and urban ecology and cannot be ignored [69].

4. Trust and Reliability

The reluctance of the scientific community seems to predominantly stem from a mistrust of citizen science datasets due to the lack of validity assessments in academic research and published literature [70, 71]. Although many recognise that citizen science has increased the amount of data that is available, it is a concern that the quality, reliability, and overall value of these data is still preventing its adoption in many research programmes [72]. Assurance of the quality of the data is needed through rigorous scientific methods in order to allow the acceptance of citizen science data into the scientific field [20].

The literature suggests that the reliability of inherently patchy data is the most questioned aspect of citizen science. Thus, being able overcome this mistrust, a huge untapped resource of citizen scientists could be opened up, increasing the scope and insight of conducted research. Potentially, this could result in large standardised spatial and temporal datasets collected by citizen sensor networks [71]. Traditional solutions to gaining credibility are to provide reliable information or gain credentials such as qualifications; however, this works only when there are "gatekeepers" to filter information, something which is not possible with the internet on a global scale [73].

Table 4: Utilising Flickr for image-based citizen science programs.

Project title and URL	Description	Passive promotion	Active promotion	Active data searching	Project base
Whale Shark Identification (http://www.flickr.com/groups/whalesharkidentification/)	To collect images to be submitted to http://www.whaleshark.org for identification from group members and other Flickr users through the search facility. It is worth noting that this Flickr Group has been formed by a volunteer and is not officially part of the project.	Y	Y Recruits members to promote	Y	N
MantaWatch (http://www.flickr.com/groups/mantawatch/)	A place for enthusiasts to meet and a promotion tool directing people to their Website (http://mantawatch.com). Does not seem to actively recruit members or search out images of manta rays on Flickr.	Y	N	N	N
Humpback whale flukes (http://www.flickr.com/groups/humpbackflukes/)	To collect images to be submitted to http://www.coa.edu/nahwc.htm for identification from group members and other Flickr users through the search facility. The same project also has a whale catalog (http://www.flickr.com/photos/flukematcher/) located on Flickr so that individuals can manually match their sightings. A further more regional group (http://www.flickr.com/groups/northatlanticflukes/) has formed due to the volume of photos uploaded.	Y	Y Recruits members to promote	Y	N
Citizen Science: Great Blue Heron (http://www.flickr.com/groups/csgreatblueheron/)	This group aims to create a database of geo-tagged images of the Great Blue Heron, entirely run and initiated by volunteers	Y	Y Recruits members to promote	Y	Y
BeeID (http://www.flickr.com/groups/beeid/)	A completed project run by student volunteers, and overseen by a lecturer, whereby members of the public are encouraged to upload photos of UK bees (Honeybees, bumblebees, and solitary bees) to their Flickr account and “geotag” them to place them on a map, with the aim of studying distribution and phenology.	Y	Y Recruits members to promote	N	Y

The dependability of volunteer-derived data is an old problem within biology and ecology, and therefore a number of methods to help to increase the reliability of the information gathered have been developed [6, 22]. Firstly, the researchers must concisely and without jargon ask the right questions in the right way to get the quality of answer that is needed, and instructions and processes must be clear and as simple as possible [3, 9–11]. Projects are usually kept relatively simple; for example, they might include counting a few common avian species frequenting a feeding table rather than searching for rare or difficult to spot species [6, 22, 74, 75]. Projects that require higher levels of skill can be successfully developed; however, they may require additional training or longevity of participation in order to increase experience indeed, many volunteer programs document “learner” effects whereby data collectors become more accurate and correct over time [6, 10, 22, 76–80]. Some of the online citizen science programmes that Cornell University has run in the past incorporate short tests and quizzes which help in assessing a contributors’ knowledge; they have also implemented an automated meso-filter which evaluates data input and evaluates it based on already known parameters, submissions which fall out of these categories are flagged for expert review, the contributor contacted, and the entry either verified or disregarded [6, 10, 81].

Although there is not enough space to review all the literature which has been published as a result of data collected through the use of citizen science participation, literature searching has resulted in the location of over 300 instances of peer-reviewed publications. This suggests that citizen science has and will continue to produce usable forms of data (See Figure 3). As with any data, datasets should be approached with caution and “cleaned” or “scrubbed” before performing analysis to remove any obvious outliers [82]. The literature suggests, however, that if the program protocols have been properly formed and tailored to the appropriate audience data does not often differ significantly from expert data collection. Delaney et al. [50] found that

Table 5: Comparison of avian monitoring projects focused on measuring occurrence and abundance (adapted from [10]).

Results from these programs have been used in over 1000 publications						
Project	Method	Placement	Effort[a]	Extent	Interval	Participants
Audubon Christmas Bird Count	Count circle (24 km diameter)	Opportunistic	V (party hours)	International	Annual	59,918 (2008-09)
North American Breeding Bird Survey	Roadside survey (39.4 km; 50 stops)	Stratified random	S (3 min count)	International	Annual	2,749 (2009)
Project FeederWatch	Feeder counts	Opportunistic	V (2 days, hours, days)	International	Annual	9,750 (2009)
eBird	Online checklists	Opportunistic	V (hours, distance, and area)	International	Continuous	18,053
Bird Atlas	Systematic grid (100 km^2 blocks; 4 km^2 tetrads) and roving reports	Regular grid and opportunistic	S/V (roving, timed visits)	Britain and Ireland	Two visits (Winter: Nov./Dec. & Jan./Feb.; Breeding: April/May & June/July)	10,000–20,000
Common Birds Census (now replaced by Breeding Bird Survey)	Census plots (Farmland: 70 ha; Woodland: 20 ha)	Stratified random	S (territory mapping)	Britain	Annual (8–10) visits; late March–early July)	250–300

[a]Effort is considered standardised (S) or variable (V). When standardised, the protocol specifications are presented; when variable, the effort variables that were reported during sampling are presented.

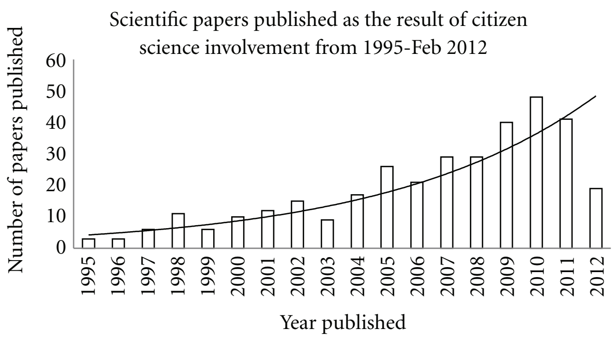


Figure 3: Numbers of published scientific papers using or resulting from citizen science data collection or involvement; it indicates an increasing trend.

far from overlooking data collection methods novices were "more careful" in their measurements and annotations, due to are being very aware of their novice status and shown in many studies to yield similar results to experts [22, 50, 83]. Delaney et al. [50] found experts and nonexperts did not come up with any significantly statistical differences, indeed students were found to be between 80 and 95% accurate with identification, with significant predictors of accuracy being their age and level of education. Dickinson et al. [10] reported that during Project FeederWatch between 2008 and 2009 they received 1,342,633 observations, out of those 378 records required "flagging" resulting in 158 records (54%) being confirmed, 45 identifications (16%) being corrected, and 88 reports (30%) being disregarded due to too little evidence.

Indeed, the very nature of gathering large sets of data results in decreased detrimental effects of "noise", greater statistical power and increased robustness, as statistical power is a function of sample sizes [22, 37]. Therefore, the common belief that volunteer collected data can only provide noisy and unreliable results that lack precision is generally incorrect [22, 37]. LePage and Francis [74] compared two citizen science programs with similar data collection protocols to test whether population patterns and distributions were temporally and spatially consistent. The study successfully showed that the two citizen science led studies, Christmas Bird Count and Project FeederWatch, had comparable trends and patterns across the same time periods, suggesting that the data was consistent and not significantly influenced by different methods and biases. The benefit of these larger datasets is that they allow researchers to draw broader conclusions across large spatial or temporal scales, enabling researchers to make inferences and robust cases for causation over a larger areas, and at a finer resolution, in contrast with small scale studies which cannot be "generalised" over greater areas [3, 6, 9, 11, 38].

It is, however, important to recognise that these datasets can be compromised by potential lack of precision, inherent biases, and uncertainties which are often present within these extensive studies [11, 22, 84]. For example, you may have more reports of species in areas that are highly populated by humans than in those that are sparsely populated, or more reports of species that are less cryptic than others. It is therefore a challenge to determine whether the data is correct or the reports are biased; this is the reason why many citizen science programs are so rigidly composed and use standardised protocols which are replicated across many stratified surveyed plots (see Table 5 and [11, 22, 84]). It is therefore important to ensure, in hypothesis driven studies, that sampling design does not introduce bias, and that counts are shaped by the data and not the ability of the observer to detect or record data [85]. This is partially why using such count data to establish index of abundance can be scientifically hazardous; however, by using capture-recapture algorithms, conversion to actual population estimates can be made and therefore data can be used to make a valid conclusion [86, 87].

Well known and successful UK citizen science-based programmes are those which are based in the public's back gardens. The British Trust for Ornithology's (BTO) and Royal Society for Protection of Birds' (RSPB's) garden-based citizen science programmes have been very successful in collecting biodiversity data, particularly on avian species. The "Garden BirdWatch" and "Big Garden Weigh-In" run by the BTO and the "Big Garden Birdwatch" and 'Make Your Nature Count survey' run by the RSPB are just a few of the citizen science programmes which encourage the recording of species which are visiting their gardens. For a full list of citizen science projects run by these organisations visit the BTO (http://www.bto.org/) and RSPB (http://www.rspb.org.uk/) websites.

These programmes have a number of key design similarities which help standardise the survey and mitigate against some of the perceived problems involved with nonexpert individuals collecting data. Indeed, they have proved to be reliable enough to result in published scientific papers. The Garden BirdWatch alone has resulted in 15 published scientific papers in addition to providing a strong set of baseline data (visit http://www.bto.org/volunteer-surveys/gbw/publications/papers for full list of publications).

To prevent confounding seasonal variation and to ensure continuity of recording effort citizen scientists are asked to record species within a given survey period, the Big Garden Weigh-In ran between the May 31st and June 5th in 2012 for example. To standardise effort the records are gathered over a particular time period, an hour is the most popular time, and many of the surveys require the species to be physically within the garden (not in a neighbouring garden or flying over). The Garden BirdWatch asks observers to repeat this recording at the same time and from the same place and of the same area for each recording session during the survey period.

Pseudoreplication is combated by removing the difference in the ability of the observer to identify different individuals; this is achieved by recording the maximum number of individual birds present at any one time within the garden. So if an observer sees one Blue Tit at the beginning of the survey but five in the middle and two towards the end of the survey, they would report it as five Blue Tits.

The species which are surveyed are also reduced to a range of easily identifiable species. The Garden Weigh-In reduces the number of birds under observation to 60 avian species which compose the core avian community. The Garden BirdWatch reduces the number further to the 42 most commonly recorded birds (nationally), with a further breakdown resulting in a list of the top ten which can have further detail added. The Big Garden BirdWatch reduces it further still providing a list of 20 more common species and ask observers to also record incidental records of other species that they might see on a separate sheet. The Garden BirdWatch goes one step further to collect additional data and provides a presence and absence record sheet for all species not mentioned.

The key difference between the RSPB and BTO's citizen science programmes is the method of collection. The RSPB has no paper-based submission format, but the BTO does, with a scanning machine which automates the data retrieval and decoding from the paper-based forms. The BTO suggests that the "relative proportions of participants submitting returns on paper and online are similar."

Neither of these programmes use social networks for more than publicity. In 2012, the BTO began the Cuckoo Tracking project, whereby tagged Cuckoo's were tracked during their migrations (http://www.bto.org/science/migration/tracking-studies/cuckoo-tracking). As part of the publicity, sightings were called for and the "hashtag" #heardacuckoo was created on the social network Twitter to publicise the project. Many individuals used the hashtag to report when they had indeed heard a cuckoo. If a tool such as CrowdMap (https://crowdmap.com/) was used to filter the tweet's with #heardacuckoo in them and verified by experts, could the conversion rate from publicity to actual record be higher?

5. The Shifting Paradigm: From "Knowledge-Driven" Analysis and Hypothesis Testing to "Data-Driven" Analysis

With the advent of the Web 2.0 world and the increase of the "citizen sensor network," there is a shifting paradigm from "knowledge-driven" analysis created by hypothesis-driven research to "data-driven" analysis, moving studies into more data-intensive science area [44, 83]. This is resulting in a new synthesis of disciplinary areas as new methods of analysis emerge to explore and identify interesting patterns that may not already be apparent; this is particularly prevalent when looking at data gathered over large spatial and temporal scales [44, 88, 89]. This approach offers valuable insights enabling further hypothesis for the discovery of underlying ecological processes. With such large datasets with such varying attributes; it is no wonder that all disciplines of science are seemingly beginning to merge into computer science as it enables scientists in varying fields to better understand complex systems [83, 90–93]. In order to better utilise citizen science collected datasets that provide a wide range of data over long periods, many researchers are moving into intelligent analysis. This may involve using novel probabilistic machine-learning statistical analysis in the form of computational modelling, or methods of analysis which include Bayesian or neural networking methods [90, 91, 93, 94]. Indeed, Link et al. [89] utilised a hierarchical model and Bayesian analyses to account for variations in effort on counts and to provide summaries over large geographic areas for a complex dataset provided by the Christmas Bird Count in America. They successfully revealed regional patterns of population change, which was then shown to be similar to data shown by the Midwinter Waterfowl Inventory in the US [89].

Currently, databases of species information are often disjunct, outdated, and incomplete, and data recording methods are often not standardised across organisational databases making reconciling datasets from different sources for studies often unreliable. This makes large scale data collection a necessity for research and the use of more complex methods of data collection an ever growing and underdemand area of study.

6. Conclusion

In our increasingly changing and evolving technological world, the presence of citizen scientists or citizen sensors who can contribute to science in more meaningful ways is allowing the rapid expansion of citizen science. Monitoring, anticipating, and mitigating large-scale threats to our biodiversity and natural world have also never been more prominent than they are now. In an increasingly urbanised world, successful monitoring of the environment is needed in the face of continuing climate and land-use change and the need to increase understanding of key ecological and environmental processes.

Citizen science and the exploitation of citizen science and sensor networks are probably one of the most important factors in being able to achieve this. The data is out there, just waiting to be understood, almost each and every person in the developed world and beyond has the potential to contribute to our understanding in a meaningful way. With the rapid progression of technology it is within our capabilities to begin this journey of understanding. It is, however, important to recognise the potential weaknesses that can result from poorly managing datasets and to preempt how the data is likely to be used and integrated beyond the original scope of the project.

It is also prudent to note something that many conservation organisations are realizing; a need to interest new generations of naturalists and enthusiasts as current recorders is an aging group with limited recruitment. By exploiting new technologies to aid recruitment of a younger generation of recorders and naturalists and educate an increasingly urbanised population, it will benefit all stake holders.

If citizen science was commonplace, how much more scientific knowledge could we discover? And in this world where people are increasingly divorced from the natural environment, how much would this influence decision making, education, and scientific thinking?

Acknowledgment

Many thanks to the University of Gloucestershire for the studentship which is supporting the author in this work. Figure 2(b) reproduced with kind permission of Springer Science and Business Media [45].

References

[1] A. Wiggins and K. Crowston, "From conservation to crowdsourcing: a typology of citizen science," in *Proceedings of the 44th Hawaii International Conference on System Sciences (HICSS '10)*, January 2011.

[2] J. P. Cohn, "Citizen science: can volunteers do real research?" *BioScience*, vol. 58, no. 3, pp. 192–197, 2008.

[3] K. S. Oberhauser and M. D. Prysby, "Citizen science: creating a research army for conservation," *American Entomologist*, vol. 54, no. 2, p. 3, 2008.

[4] S. Droege, "Just because you paid them doesn't mean their data are better," in *Proceedings of the Citizen Science Toolkit Conference*, Ithaca, NY, USA, 2007.

[5] J. Silvertown, "A new dawn for citizen science," *Trends in Ecology & Evolution*, vol. 24, no. 9, pp. 467–471, 2009.

[6] R. Bonney, C. B. Cooper, J. Dickinson et al., "Citizen science: a developing tool for expanding science knowledge and scientific literacy," *BioScience*, vol. 59, no. 11, pp. 977–984, 2009.

[7] A. Irwin, *Citizen Science: A Study of People, Expertise and Sustainable Development*, Routledge, London, UK, 1995.

[8] J. Howe, "The rise ofcrowdsourcing," *Wired Magazine*. In press.

[9] C. B. Cooper, J. Dickinson, T. Phillips, and R. Bonney, "Citizen science as a tool for conservation in residential ecosystems," *Ecology and Society*, vol. 12, no. 2, article 11, 2007.

[10] J. L. Dickinson, B. Zuckerberg, and D. N. Bonter, "Citizen science as an ecological research tool: challenges and benefits," *Annual Review of Ecology, Evolution, and Systematics*, vol. 41, pp. 149–172, 2010.

[11] V. Devictor, R. J. Whittaker, and C. Beltrame, "Beyond scarcity: citizen science programmes as useful tools for conservation biogeography," *Diversity and Distributions*, vol. 16, no. 3, pp. 354–362, 2010.

[12] N. R. Prestopnik and K. Crowston, "Citizen science system assemblages: understanding the technologies that support crowdsourced science," in *Proceedings of the 2012 iConference*, pp. 168–176.

[13] D. Brossard, B. Lewenstein, and R. Bonney, "Scientific knowledge and attitude change: the impact of a citizen science project," *International Journal of Science Education*, vol. 27, no. 9, pp. 1099–1121, 2005.

[14] D. J. Trumbull, R. Bonney, D. Bascom, and A. Cabral, "Thinking scientifically during participation in a citizen-science project," *Science Education*, vol. 84, no. 2, pp. 265–275, 2000.

[15] A. W. Crall, G. J. Newman, C. S. Jarnevich, T. J. Stohlgren, D. M. Waller, and J. Graham, "Improving and integrating data on invasive species collected by citizen scientists," *Biological Invasions*, vol. 12, no. 10, pp. 3419–3428, 2010.

[16] D. M. Aanensen, D. M. Huntley, E. J. Feil, F. Al-Own, and B. G. Spratt, "EpiCollect: linking smartphones to web applications for epidemiology, ecology and community data collection," *PLoS ONE*, vol. 4, no. 9, Article ID e6968, 2009.

[17] N. D. Lane, E. Miluzzo, H. Lu, D. Peebles, T. Choudhury, and A. T. Campbell, "A survey of mobile phone sensing," *IEEE Communications Magazine*, vol. 48, no. 9, pp. 140–150, 2010.

[18] S. Matyas, C. Matyas, C. Schlieder, P. Kiefer, H. Mitarai, and M. Kamata, "Designing location-based mobile games with a purpose—collecting geospatial data with cityexplorer," in *Proceedings of the International Conference on Advances in Computer Entertainment Technology (ACE '08)*, pp. 244–247, ACM, Yokohama, Japan, December 2008.

[19] D. Couvet, F. Jiguet, R. Julliard, H. Levrel, and A. Teyssedre, "Enhancing citizen contributions to biodiversity science and public policy," *Interdisciplinary Science Reviews*, vol. 33, no. 1, pp. 95–103, 2008.

[20] S. A. Boudreau and N. D. Yan, "Auditing the accuracy of a volunteer-based surveillance program for an aquatic invader bythotrephes," *Environmental Monitoring and Assessment*, vol. 91, no. 1–3, pp. 17–26, 2004.

[21] A. Sheth, "Citizen sensing, social signals, and enriching human experience," *IEEE Internet Computing*, vol. 13, no. 4, pp. 87–92, 2009.

[22] D. S. Schmeller, P. Y. Henry, R. Julliard et al., "Advantages of volunteer-based biodiversity monitoring in Europe," *Conservation Biology*, vol. 23, no. 2, pp. 307–316, 2009.

[23] A. M. Kitchen-Wheeler, "Visual identification of individual manta ray (Manta alfredi) in the Maldives Islands, Western Indian Ocean," *Marine Biology Research*, vol. 6, no. 4, pp. 351–363, 2010.

[24] M. E. Fernandez-Gimenez, H. L. Ballard, and V. E. Sturtevant, "Adaptive management and social learning in collaborative and community-based monitoring: a study of five community-based forestry organizations in the western USA," *Ecology and Society*, vol. 13, no. 2, article 4, 2008.

[25] D. Penrose and S. M. Call, "Volunteer monitoring of benthic macroinvertebrates: regulatory biologists' perspectives," *Journal of the North American Benthological Society*, vol. 14, no. 1, pp. 203–209, 1995.

[26] C. Evans, E. Abrams, R. Reitsma, K. Roux, L. Salmonsen, and P. P. Marra, "The Neighborhood Nestwatch program: participant outcomes of a citizen-science ecological research project," *Conservation Biology*, vol. 19, no. 3, pp. 589–594, 2005.

[27] Y. Bhattacharjee, "Citizen scientists supplement work of Cornell researchers," *Science*, vol. 308, no. 5727, pp. 1402–1403, 2005.

[28] Missouri Stream Team Project, 2011, http://www.mostreamteam.org/.

[29] ReClam the Bay, 2005, http://www.reclamthebay.org/.

[30] Shermans Creek Conservation Association, 1998.

[31] M. J. Raddick, A. S. Szalay, J. Vandenberg et al., "Galaxy zoo: exploring the motivations of citizen science volunteers," *Astronomy Education Review*, vol. 9, no. 1, pp. 010103–010118, 2010.

[32] K. Luther, S. Counts, K. B. Stecher, A. Hoff, and P. Johns, "Pathfinder: an online collaboration environment for citizen scientists," in *Proceedings of the 27th International Conference on Human factors in Computing Systems (Chi '09)*, vol. 1–4, pp. 239–248, 2009.

[33] L. S. Fore, K. Paulsen, and K. O'Laughlin, "Assessing the performance of volunteers in monitoring streams," *Freshwater Biology*, vol. 46, no. 1, pp. 109–123, 2001.

[34] D. M. Lodge, S. Williams, H. J. MacIsaac et al., "Biological invasions: recommendations for U.S. policy and management," *Ecological Applications*, vol. 16, no. 6, pp. 2035–2054, 2006.

[35] R. Lukyanenko, J. Parsons, and Y. Wiersma, "Enhancing citizen science participation in GeoWeb projects through the instance-based data model," in *Proceedings of the Spatial Knowledge and Information*, Fernie, Canada, 2011.

[36] J. E. Losey, J. E. Perlman, and E. R. Hoebeke, "Citizen scientist rediscovers rare nine-spotted lady beetle, Coccinella novemnotata, in eastern North America," *Journal of Insect Conservation*, vol. 11, no. 4, pp. 415–417, 2007.

[37] J. J. D. Greenwood, "Citizens, science and bird conservation," *Journal of Ornithology*, vol. 148, no. 1, pp. S77–S124, 2007.

[38] V. Devictor, R. Julliard, D. Couvet, and F. Jiguet, "Birds are tracking climate warming, but not fast enough," *Proceedings of the Royal Society B*, vol. 275, no. 1652, pp. 2743–2748, 2008.

[39] A. H. Hurlbert and Z. Liang, "Spatiotemporal variation in avian migration phenology: citizen science reveals effects of climate change," *Plos One*, vol. 7, no. 2, Article ID e31662, 2012.

[40] S. Altizer, W. M. Hochachka, and A. A. Dhondt, "Seasonal dynamics of mycoplasmal conjunctivitis in eastern North American house finches," *Journal of Animal Ecology*, vol. 73, no. 2, pp. 309–322, 2004.

[41] K. V. Rosenberg, J. D. Lowe, and A. A. Dhondt, "Effects of forest fragmentation on breeding tanagers: a continental perspective," *Conservation Biology*, vol. 13, no. 3, pp. 568–583, 1999.

[42] W. M. Hochachka, R. Caruana, D. Fink et al., "Data-mining discovery of pattern and process in ecological systems," *Journal of Wildlife Management*, vol. 71, no. 7, pp. 2427–2437, 2007.

[43] D. Fink, W. M. Hochachka, B. Zuckerberg et al., "Spatiotemporal exploratory models for broad-scale survey data," *Ecological Applications*, vol. 20, no. 8, pp. 2131–2147, 2010.

[44] S. Kelling, W. M. Hochachka, D. Fink et al., "Data-intensive science: a new paradigm for biodiversity studies," *BioScience*, vol. 59, no. 7, pp. 613–620, 2009.

[45] E. Howard and A. K. Davis, "The fall migration flyways of monarch butterflies in eastern North America revealed by citizen scientists," *Journal of Insect Conservation*, vol. 13, no. 3, pp. 279–286, 2009.

[46] A. K. Davis and E. Howard, "Spring recolonization rate of Monarch butterflies in eastern North America: new estimates from citizen-science data," *Journal of the Lepidopterists' Society*, vol. 59, no. 1, pp. 1–5, 2005.

[47] B. L. Sullivan, C. L. Wood, M. J. Iliff, R. E. Bonney, D. Fink, and S. Kelling, "eBird: a citizen-based bird observation network in the biological sciences," *Biological Conservation*, vol. 142, no. 10, pp. 2282–2292, 2009.

[48] T. O'Reilly, "What is Web 2.0: design patterns and business models for the next generation of software," *International Journal of Digital Economics*, no. 65, pp. 17–37, 2007.

[49] C. Wood, B. Sullivan, M. Iliff, D. Fink, and S. Kelling, "eBird: engaging birders in science and conservation," *Plos Biology*, vol. 9, no. 12, Article ID e1001220, 2011.

[50] D. G. Delaney, C. D. Sperling, C. S. Adams, and B. Leung, "Marine invasive species: validation of citizen science and implications for national monitoring networks," *Biological Invasions*, vol. 10, no. 1, pp. 117–128, 2008.

[51] J. Bughin, "The rise of enterprise 2.0," *Journal of Direct, Data and Digital Marketing Practice*, vol. 9, no. 3, pp. 251–259, 2008.

[52] E. Constantinides, C. L. Romero, and M. A. G. Boria, "Social media: a new frontier for retailers?" in *European Retail Research*, B. Swoboda, D. Morschett, T. Rudolph, P. Schnedlitz, and H. Schramm-Klein, Eds., pp. 1–28, Gabler, 2009.

[53] W. G. Mangold and D. J. Faulds, "Social media: the new hybrid element of the promotion mix," *Business Horizons*, vol. 52, no. 4, pp. 357–365, 2009.

[54] P. Klemperer, *Network Effects and Switching Costs: Two Short Essays for the New Palgrave*, Social Science Research Network, 2006.

[55] S. J. Liebowitz and S. E. Margolis, "Network externality—an uncommon tragedy," *Journal of Economic Perspectives*, vol. 8, no. 2, pp. 133–150, 1994.

[56] P. Anderson, *What is Web 2.0? Ideas, Technologies and Implications for Education*, JISC Technology & Standards Watch, 2007.

[57] S. Gaonkar, J. Li, R. R. Choudhury, L. Cox, and A. Schmidt, "Micro-blog: sharing and querying content through mobile phones and social participation," in *Proceedings of the 6th*

International Conference on Mobile Systems, Applications, and Services (Mobisys'08), pp. 174–186, 2008.

[58] Facebook, Statistics, 2011, http://www.facebook.com/press/info.php?statistics.

[59] A. Lipsman, *The Network Effect: Facebook, Linkedin, Twitter & Tumblr Reach New Height in May*, 2011, http://blogcom-.score.com/2011/06/facebook_linkedin_twitter_tumblr.html.

[60] C. W. Speed, M. G. Meekan, and C. J. A. Bradshaw, "Spot the match—wildlife photo-identification using information theory," *Frontiers in Zoology*, vol. 4, article 2, 2007.

[61] Z. Arzoumanian, J. Holmberg, and B. Norman, "An astronomical pattern-matching algorithm for computer-aided identification of whale sharks Rhincodon typus," *Journal of Applied Ecology*, vol. 42, no. 6, pp. 999–1011, 2005.

[62] M. J. Kelly, "Computer-aided photograph matching in studies using individual identification: an example from Serengeti cheetahs," *Journal of Mammalogy*, vol. 82, no. 2, pp. 440–449, 2001.

[63] C. J. R. Anderson, N. Da Vitoria Lobo, J. D. Roth, and J. M. Waterman, "Computer-aided photo-identification system with an application to polar bears based on whisker spot patterns," *Journal of Mammalogy*, vol. 91, no. 6, pp. 1350–1359, 2010.

[64] R. Stafford, A. G. Hart, L. Collins et al., "Eu-social science: the role of internet social networks in the collection of bee biodiversity data," *PloS one*, vol. 5, no. 12, p. e14381, 2010.

[65] M. Ebner and M. Schiefner, "Microbloggging—more than fun?" in *Proceedings of the IADIS Mobile Learning Conference*, Algarve, Portugal, 2008.

[66] M. Ames and M. Naaman, "Why we tag: motivations for annotation in mobile and online media," in *Proceedings of the 25th SIGCHI Conference on Human Factors in Computing Systems (CHI '07)*, vol. 1-2, pp. 971–980, May 2007.

[67] C. L. Kirkhope, R. L. Williams, C. L. Catlin-Groves et al., "Social networking for biodiversity: the BeeID project," in *Proceedings of the International Conference on Information Society (i-Society '10)*, 2010.

[68] Y. F. Wiersma, "Birding 2.0: citizen science and effective monitoring in the web 2.0 world," *Avian Conservation and Ecology*, vol. 5, no. 2, 2010.

[69] E. Paulos, "Designing for doubt: citizen science and the challenge of change," in *Engaging Data: First International Forum on the Application and Management of Personal Electronic Information*, MIT, 2009.

[70] M. F. Goodchild, "Citizens as sensors: the world of volunteered geography," *GeoJournal*, vol. 69, no. 4, pp. 211–221, 2007.

[71] M. F. Goodchild, "Citizens as voluntary sensors: spatial data infrastructures in the world of web 2.0," *Journal of Spatial Data Infrastructures Research*, vol. 2, p. 8, 2007.

[72] A. Flanagin and M. Metzger, "The credibility of volunteered geographic information," *GeoJournal*, vol. 72, no. 3-4, pp. 137–148, 2008.

[73] T. A. Callister Jr., "Media literacy: on-ramp to the literacy of the 21st century or cul-de-sac on the information superhighway," *Advances in Reading/Language Research*, vol. 7, p. 17, 2000.

[74] D. Lepage and C. M. Francis, "Do feeder counts reliably indicate bird population changes? 21 Years of winter bird counts in Ontario, Canada," *The Condor*, vol. 104, no. 2, pp. 255–270, 2002.

[75] J. J. Lennon, P. Koleff, J. J. D. Greenwood, and K. J. Gaston, "Contribution of rarity and commonness to patterns of species richness," *Ecology Letters*, vol. 7, no. 2, pp. 81–87, 2004.

[76] Y. Bas, V. Devictor, J. P. Moussus, and F. Jiguet, "Accounting for weather and time-of-day parameters when analysing count data from monitoring programs," *Biodiversity and Conservation*, vol. 17, no. 14, pp. 3403–3416, 2008.

[77] W. L. Kendall, B. G. Peterjohn, and J. R. Sauer, "First-time observer effects in the North American Breeding Bird Survey," *Auk*, vol. 113, no. 4, pp. 823–829, 1996.

[78] F. Jiguet, "Method learning caused a first-time observer effect in a newly started breeding bird survey," *Bird Study*, vol. 56, no. 2, pp. 253–258, 2009.

[79] J. R. Sauer, B. G. Peterjohn, and W. A. Link, "Observer differences in the North-American Breeding Bird Survey," *Auk*, vol. 111, no. 1, pp. 50–62, 1994.

[80] R. E. McCaffrey, "Using citizen science in urban bird studies," *Urban Habitats*, vol. 3, no. 1, p. 16, 2005.

[81] R. Caruana, M. Elhawary, A. Munson et al., "Mining citizen science data to predict prevalence of wild bird species," in *Proceedings if the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '06)*, pp. 909–915, ACM, Philadelphia, Pa, USA, August 2006.

[82] E. Rahm, "Data cleaning: problems and current approaches," *IEEE Data Engineering Bulletin*, vol. 23, p. 11, 2000.

[83] H. B. Newman, M. H. Ellisman, and J. A. Orcutt, "Data-intensive e-science frontier research," *Communications of the ACM*, vol. 46, no. 11, pp. 68–77, 2003.

[84] M. J. G. Hopkins, "Modelling the known and unknown plant biodiversity of the Amazon Basin," *Journal of Biogeography*, vol. 34, no. 8, pp. 1400–1411, 2007.

[85] C. A. Lepczyk, "Integrating published data and citizen science to describe bird diversity across a landscape," *Journal of Applied Ecology*, vol. 42, no. 4, pp. 672–677, 2005.

[86] C. Eraud, J. M. Boutin, D. Roux, and B. Faivre, "Spatial dynamics of an invasive bird species assessed using robust design occupancy analysis: the case of the Eurasian collared dove (*Streptopelia decaocto*) in France," *Journal of Biogeography*, vol. 34, no. 6, pp. 1077–1086, 2007.

[87] B. K. Williams, J. D. Nichols, and M. J. Conroy, *Analysis and Management of Animal Populations*, Academic Press, London, UK, 2002.

[88] P. H. Crowley, "Resampling methods for computation-intensive data analysis in ecology and evolution," *Annual Review of Ecology and Systematics*, vol. 23, no. 1, pp. 405–447, 1992.

[89] W. A. Link, J. R. Sauer, and D. K. Niven, "A hierarchical model for regional analysis of population change using Christmas Bird Count data, with application to the American Black Duck," *The Condor*, vol. 108, no. 1, pp. 13–24, 2006.

[90] S. Reddy, D. Estrin, and M. Srivastava, "Recruitment framework for participatory sensing data collections," in *Proceedings of the 8th International Conference on Pervasive Computing*, P. Floréen, A. Krüger, and M. Spasojevic, Eds., pp. 138–155, Springer, Berlin, Germany, 2010.

[91] Y. Jun, W. Weng-Keen, and R. A. Hutchinson, "Modeling experts and novices in citizen science data for species distribution modeling," in *Proceedings of the 10th IEEE International Conference on Data Mining (ICDM '10)*, pp. 1157–1162, December 2010.

[92] P. Cohen and N. Adams, "Intelligent Data Analysis in the 21st Century," in *Proceedings of the 8th International Symposium on Intelligent Data Analysis: Advances in Intelligent Data Analysis VIII (IDA '09)*, N. Adams, C. Robardet, A. Siebes, and J. -F. Boulicaut, Eds., pp. 1–9, Springer, Berlin, Germany, 2009.

[93] D. Z. Wang, E. Michelakis, M. N. Garofalakis, and J. M. Hellerstein, "BayesStore: managing large, uncertain data repositories

with probabilistic graphical models," in *Proceedings of the VLDB Endowment*, vol. 1, no. 1, pp. 340–351, 2008.

[94] W. A. Link and J. R. Sauer, "Seasonal components of avian population change: joint analysis of two large-scale monitoring programs," *Ecology*, vol. 88, no. 1, pp. 49–55, 2007.

Hindawi Publishing Corporation
International Journal of Zoology
Volume 2012, Article ID 724835, 13 pages
doi:10.1155/2012/724835

Research Article

The Accuracy of Behavioural Data Collected by Visitors in a Zoo Environment: Can Visitors Collect Meaningful Data?

Rachel L. Williams,[1] Sue K. Porter,[2] Adam G. Hart,[1] and Anne E. Goodenough[1]

[1] *Department of Natural and Social Sciences, University of Gloucestershire, Cheltenham GL50 4AZ, UK*
[2] *Wildfowl and Wetlands Trust, Slimbridge, Gloucestershire GL2 7BT, UK*

Correspondence should be addressed to Rachel L. Williams, rachel.williams_31@hotmail.com

Received 24 February 2012; Accepted 22 June 2012

Academic Editor: Simon Morgan

Volunteer data collection can be valuable for research. However, accuracy of such data is often a cause for concern. If clear, simple methods are used, volunteers can monitor species presence and abundance in a similar manner to professionals, but it is unknown whether volunteers could collect accurate data on animal behaviour. In this study, visitors at a Wetlands Centre were asked to record behavioural data for a group of captive otters by means of a short questionnaire. They were also asked to provide information about themselves to determine whether various factors would influence their ability to collect data. Using a novel analysis technique based on PCA, visitor data were compared to baseline activity budget data collected by a trained biologist to determine whether visitor data were accurate. Although the response rate was high, visitors were unable to collect accurate data. The principal reason was that visitors exceeded the observation time stated in the instructions, rather than being unable to record behaviours accurately. We propose that automated recording stations, such as touchscreen displays, might prevent this as well as other potential problems such as temporal autocorrelation of data and may result in accurate data collection by visiting members of the public.

1. Introduction

Animal behaviour data are important across the field of biological sciences, from evolution and population biology to ethology in captive or domesticated animals. However, collecting these data is time consuming. Given that the duration of data collection for behavioural studies can range from several weeks [1, 2] to several years [3], funding professional researchers can be prohibitively expensive for many studies, especially those conducted by zoological parks and wildlife organisations [4, 5]. However, animal behaviour is of considerable interest to the general public (or at least a subset of the public with environmental and zoological interests), and many people spend considerable time observing animals as a hobby (e.g., watching pets, wild birds, or animals in zoos). Professionals could use this interest to recruit volunteers to record animal behaviour.

There are many advantages of using volunteers to collect data. Volunteers can collect data at little or no financial cost to the organisation running the project [4–6]; indeed large numbers of untrained members of the public have been collecting biodiversity data for wildlife organisations for several decades. For example, in 2011, over 600,000 members of the public took part in the Royal Society for the Protection of Birds' "Big Garden Birdwatch" [7]. Several studies have shown that volunteer-collected data on, for example, species identification and quantifying abundance, can be as accurate as basic biodiversity data recorded by scientists [4, 6, 8, 9], especially when projects offer basic training and are closely supervised by scientists. Moreover, several methods have been developed to enhance the accuracy of volunteer-run surveys, either in terms of the methods used to collect the data or in subsequent analysis [4, 10–14]. Collection of behavioural data, however, is subject to a certain degree of interpretation and may be more complex to record than counting or identifying species. It is not known whether the quality of volunteer-collected behavioural data would be sufficient to calculate accurate activity budgets or to test behavioural ecology hypotheses.

Monitoring animal behaviour is particularly important in zoos because of the importance of animal welfare [15, 16]. Zoos may encourage their zookeepers to participate in research [17] but data collection often cannot be a priority amongst the zookeepers' daily husbandry activities [18]. Research activities can be supplemented with undergraduate and postgraduate students under the supervision of lecturers and scientists, with no financial cost for the zoos involved [19, 20], but while this provides useful and reliable data, it relies on the availability of students and on University course content.

An alternative approach could be to use zoo visitors to collect data on a voluntary basis. The benefits of asking zoo visitors to collect data while they visit could be numerous. Zoos are popular attractions worldwide, attracting more than 700 million people each year [21], so there is no shortage of potential volunteers. Many visitors have a keen interest in animals and wildlife conservation [22, 23], and this could be a strong incentive to participate in research that may benefit the animals they are observing. Furthermore, behavioural data could be collected almost continuously throughout the day as and when visitors pass the animal enclosures. This should create a database from which daily activity budgets can be calculated. Finally, interactive activities create more positive experiences for visitors when compared to passive exhibit viewing [24], so an activity such as this could make the zoo more attractive to its visitors.

While some research suggests that zookeepers' casual observations throughout the day provide a good indication of the overall activity budgets of the animals [18, 25, 26], and keepers are generally well acquainted with individual animals and their behaviours, they may not be acquainted with recording behaviour in a scientific and rigorous manner. It also seems reasonable to assume that the vast majority of visitor-based "volunteers" would have no prior experience of collecting behavioural data and it would be logistically difficult, or impossible, to train and/or supervise them while they collect data. However, if visitors are able to collect accurate data on captive animals, there is a potential for volunteer projects to collect behavioural data on wild animals, especially where there are large concentrations of people and animals, such as in nature reserves or game parks. The aim of this study is to determine whether visitors can collect accurate data on the behaviour of a small group of animals in a captive environment. Visitor data were compared to data collected by a trained biologist.

2. Methods

2.1. Study Site. The study was conducted at the Wildfowl and Wetlands Trust (WWT) centre at Slimbridge, Gloucestershire, UK (OS grid reference SO722047). A group of three female captive North American river otters (*Lontra canadensis*) were selected for the study because of their popularity with visitors and the fact that this species demonstrated a rich suite of behaviours during the daily opening hours of the centre (R. L. Williams pers. obs.). It was important that visitors could see the otters in order to record their behaviour, and the layout of the otter enclosure facilitated this. Large panels of clear glass around the enclosure allowed visitors to view the otters easily from the walkway that spanned the front of the enclosure (Figure 1). There was also a small indoor sleeping chamber in which visitors could see the otters through small glass windows in a walkthrough tunnel. Otters could access all parts of the enclosure at any time of the day, and no parts of the enclosure were closed during routine cleaning of the exhibit.

2.2. Ethogram Data

2.2.1. Ethogram Construction and Scientific Data Collection. To determine whether visitors could record data that would accurately represent the otters' behaviour, reliable baseline data were required for comparison. A biologist with experience in collecting behavioural data (RLW) created an ethogram as per Martin and Bateson [27] to record the otters' behaviour based on prior observations in a pilot study. Behaviour categories were adapted from a behavioural study done by Anderson et al. [24] on a similar species (Asian small-clawed otters—*Aonyx cinerea*). Behaviours were grouped into simple, easily definable, categories to ensure that members of the public should be able to recognise them in the latter part of the study (Table 1). The study took place over 7 days during the opening hours of the park (10 am until 5 pm). Each hour was divided into six 10 minute periods and the otters' behaviour was recorded during two randomly selected 10-minute periods each hour [28]. An instantaneous scan sampling method [27–29] was used to record the behaviour of each of the 3 otters systematically every 10 s during the recording periods. This was the shortest interval in which data could be recorded by watching each otter consecutively. By using this sampling technique for each of the otters, the problem of missing out individual behaviours was minimised and an overall activity budget for all three otters could also be calculated. Subtle differences in size and coat colouration were used to distinguish each otter to calculate individual activity budgets. If an individual otter was out of view at any time during the recording period, it was noted as such. In total, 16.5 h of data were collected for each otter, with a data point collected from each otter simultaneously, giving 1,980 ethogram observations per otter (6 recordings per minute, that is, one every 10 seconds, $\times 20$ minutes of observation per hour $\times 16.5$ hours in total $= 1,980$). This sample size is comparable to those used in studies of a similar nature [18, 30].

2.2.2. Interobserver Variability. To examine the potential for interobserver variability in the collection of behavioural data, a second biologist (herein referred to as CK; not an author of this study and independent from its planning and prior implementation but with the same level of experience as RLW) collected ethogram data over one day, during exactly the same recording periods (14×10 min). The paired data were then compared.

Figure 1: Otter enclosure at Slimbridge, a photograph taken from the front of the enclosure and showing the visitors' viewpoint.

2.3. Questionnaires

2.3.1. Otter Behaviour Questionnaire. The ethogram was simplified to a multiple-choice questionnaire to determine whether visitors could collect accurate data on otter behaviour. The instructions on the questionnaire were as clear, concise, and self-explanatory as possible, as recommended by previous studies [6, 8, 10, 12, 31]. Visitors had to fill in basic information (e.g., write the time down, answer "yes" or "no" if they could see otters inside and/or outside), and tick the behaviours they saw when the otters were outside (i.e., not in the sleeping chamber) during a 30 s period. This method was adapted from the one-zero sampling method in that all behaviours which were observed within the interval were ticked once (1) and those that were not observed were not ticked (0). It is recognised that the two datasets differed not only in who had collected the data (biologist or visitors) but how the data had been collected (ethogram instantaneous scan sampling or questionnaire extended one-zero sampling, resp.). The differences in data collection methods were undertaken for good reason-one-zero sampling was the easiest type of sampling for visitors (and thus the most likely to be reliable) whereas instantaneous scan sampling is a more robust method for generating data for activity budgets. Therefore, although it could be argued that different methods will give different results, the study aimed to determine whether visitor-collected data (at its simplest) could be compared to maximally robust and reliable data, validating the approach taken.

The layout of the questionnaire was an important consideration [32]. Colour photographs were used to illustrate each of the behaviours with the exception of "other", which was represented by a question mark with space underneath for visitors to write down what they had seen. Visitors were not asked to distinguish between individual otters, because identifying them reliably would have been very difficult given the short recording period and subtlety of the differences between otters. Consequently, they were requested to record all of the behaviours they observed, regardless of which individual was performing the behaviour. The "out of view" category from the ethogram was not included in the questionnaire because visitors did not know how many otters were in the enclosure. If they could not see any of the otters, they should have answered "no" to the questions asking whether they could see any otters inside or outside.

Visitors were asked how long they spent at the otter enclosure overall to determine whether this was related to the number of behaviours recorded, and because this could be a potential indication that visitors might be spending longer than the requested 30 s recording data. Visitors were asked some anonymous personal information questions (e.g., their age group, whether they had volunteered before, whether they were a member of a wildlife organisation) to determine whether any of these factors influenced their ability to record accurate data. Finally, visitors were required to indicate how many people had helped them fill in the questionnaire.

The study took place over 8 consecutive days, for 7 hours each day. Visitor data were collected for a day more than the ethogram data because of logistical issues when undertaking both activities was not possible. However, analysis of daily otter activity budgets after the data were collected showed that this did not affect the results. The study was advertised using A3-sized posters at the entrance of the centre and near the otter enclosure, and was promoted by the mammal keeper during the twice daily otter feeding demonstrations (11.30 am and 3.30 pm). Visitors approaching the otter enclosure were asked whether they would be willing to fill in a questionnaire as part of a research project on otter behaviour. No other details were given unless visitors asked questions, as the aim of the study was to determine whether visitors could collect data without supervision. In order to compare ethogram- and questionnaire-derived data, both were collected on the same days (in order to ensure consistent activity levels of the otters—Anderson et al. [24]). The study was carried out on four days before the school holidays and on four days during the school holidays. This allowed a comparison between uptake of the questionnaire during quiet and busy periods at the centre, as well as increasing the range of different visitors filling in the questionnaire (e.g., more families during school holidays).

2.3.2. Visitor Segmentation Questionnaire. The WWT developed a questionnaire as part of a survey to learn more about their visitors, and this was used as a complementary tool in this study [33]. This questionnaire (named the visitor segmentation questionnaire) was stapled behind the otter behaviour questionnaire, but was optional so that length of the two combined questionnaires did not deter visitors from participating. It consisted of a list of questions with the instruction "tick the statement that best describes you". The questions concerned topics such as motivations for visiting the centre, personal interests and affinity for nature, and preferences for various animals at the centre. Analysis of the results determined which "segment" a visitor belonged to (Table 2) and, subsequently, allowed examination to test whether different segments of visitors could record otter behaviour more effectively than others.

Table 1: Ethogram used by a trained biologist to record simple otter behaviours.

Behaviour	Comments and additional information
Inside	"Inside" is not a behaviour, but it was necessary to record this so that the period of time that the otters spent inside was included in the activity budget (it was speculated that visitors may underrecord otters when they were inside—Section 4).
Swimming	In water, not interacting with other otters and/or showing signs of play.*
Eating	This occurred mainly during twice-daily public demonstrations.
Playing	Any playful interaction with another otter (such as chasing, play fighting) or playing alone (diving/rolling in the water, playing with an object).*
Walking or running	As stated.
Grooming	Self-grooming or mutual grooming (if mutual grooming occurred, all otters involved were recorded as grooming).
Rolling	Rolling on land.
Sitting or lying down	Inactive animal (included pausing for a few seconds but also sleeping outside).
Fighting	This was never recorded with the ethogram, though the otters did display aggressive behaviour over food on one occasion (outside a recording period), so it is possible that visitors could have recorded this.
Other	Any behaviour not mentioned above, for example, sprainting, climbing a tree, and drinking.
Out of view	If an otter was not observable at any point during a sampling interval such that its behaviour could not be recorded (i.e., under the pedestrian walkway or hidden in vegetation).

* See Section 4 for comments about the differentiation of swimming and playing.

Table 2: Segmentation pen portraits—Modified and adapted from WWT visitor segmentation report [33].

Visitor segment	Description and comments
Learn together families	They believe in life-long learning for their family. Accessing the outside plays an important role in their leisure time, and they are generally open to all forms of nature, rather than visiting specifically to see birds.
Fun time families	Doing something that entertains and satisfies their children is the main priority in their day out. If their children learn something along the way, then this is an added bonus.
Social naturalists	Their interest in nature is broad; it is not about acquiring detailed knowledge on specific species but more about simply enjoying any kind of wildlife.
Interested naturalists	Interested naturalists are not active birdwatchers but visit to improve their knowledge and learn new things, driven by a broad interest in the natural world.
Interested birders	For interested birders, trips in the outside are a significant part of their life, and the majority are active birdwatchers. Whilst they are mainly looking to develop their interests, their interest in birds is often tied into other hobbies such as walking, photography, and painting.
Social birders	Social birders are seeking to spend quality time with other people in natural surroundings where they are guaranteed to see interesting birds.
Expert birders	Expert birders are applied birdwatchers who tend to take their hobby relatively seriously. This segment has the most knowledge about the WWT's wider conservation activities.
Sensualists	Experiencing the outside is essential to sensualists' lives; to them, it is food for the soul and is a space in which they can relax and experience nature's beauty.
Social day-outers	Wildlife and the outside are not of prime interest to them; their main focus is to spend quality time with others in a nice environment.

2.4. Data Processing and Analysis

2.4.1. Uncorrected and Corrected Data. When data were entered into a spreadsheet, two copies were made: an uncorrected version with data exactly as they were recorded by visitors and a corrected version, whereby any mistakes visitors had made that were noticed by RLW were rectified when possible or omitted from the dataset if the whole questionnaire was unusable (c. 10% of the questionnaires were affected). Mistakes that resulted in exclusion from the corrected dataset included writing the wrong time (pers. obs.), not answering all of the questions, and ticking all of the boxes haphazardly (such questionnaires were usually filled in by young children—pers. obs.). Questionnaires that could be rectified were those in which visitors had interpreted a behaviour as "other" when it could be reclassified as one of the categories listed, for example, "kissing" or "licking" = grooming; "going through tunnel" = playing, and so forth. These datasets are henceforth referred to as uncorrected visitor data and corrected visitor data.

2.4.2. Calculating Activity Budgets. Ethogram data and questionnaire data were converted into activity budgets to indicate the percentage occurrence of specific behaviours as per Stafford et al. [30]. An activity budget was calculated for each individual otter and for the whole group (using ethogram data), as well as for the group of otters using visitor data (using corrected and uncorrected data). In addition to the full questionnaire datasets, various subsets were extracted for separate analysis, for example, for each visitor segment and from adapted or standardised datasets (see below).

2.4.3. Adaptation of the Visitor Datasets and Extraction of Subsets . In addition to the full activity budgets mentioned above, activity budgets were also calculated with the behaviours playing and swimming combined into one category because these behaviours often overlapped. This was similar to the adaptations of Margulis and Westhus [18] where "swim" and "stereotypic swim" were combined to allow the comparison of keeper-collected data and scientist data on brown bear (*Ursus arctos*) behaviour.

There was a disparity in the number of visitors at different times of day, which could have led to an underrepresentation of inside in the mornings when there were fewer questionnaires completed (because there were fewer visitors in the centre) and an overrepresentation of eating when many questionnaires were filled in during the otter demonstrations. To reduce the effect of pseudoreplication and temporal autocorrelation (visitors recording the same behaviours at the same time) that may result from this, an average activity budget was calculated over each half hour period taking into account the number of questionnaires answered in each period. Given the varying length of time that visitors had the questionnaire (including filling in the segmentation questionnaires) it was not logistically possible to calculate an average from the questionnaires over a shorter time interval than 30 min, and in some cases, autocorrelation between questionnaires was likely. The effects of this possible autocorrelation are discussed below.

Separate activity budgets were also calculated from subsets of questionnaires extracted from the complete dataset. These were based on the personal information questions at the end of the behaviour questionnaire. Activity budgets were calculated based on the removal of all questionnaires that had been filled in by a child aged 10 or under from the initial dataset (because children may have difficulty giving accurate answers [34]), as well as separate subsets for the visitors who had prior experience volunteering and for those who had none, and for visitors who were members of a wildlife organisation and for those who were not.

2.4.4. PCA and Analytical Framework. To compare the ethogram activity budgets with the activity budgets calculated for the visitor datasets and subsets, bootstrapped principal components analysis (PCA) was conducted in the *R* statistical package [35], following methods in Stafford et al. [30]. Rather than plotting each activity budget on a two-dimensional scatterplot (as in conventional PCA), this approach involved plotting the mean value of calculated principal components in three dimensions with the radius of the resulting sphere, or "bubble", indicating the confidence radius. Plots were constructed using the RGL library and rgl.sphere function for *R* [36]. Each bubble represented the overall activity budget, with the centre representing the mean of the first three principal components and the radius representing the 95% confidence interval. Statistical inferences were made on the basis that overlapping bubbles signify no significant difference between the activity budgets represented by the bubbles while no overlap indicates significant differences in the activity budgets ($\alpha = 0.05$). In order for the plot to be reliable, the cumulative proportion of the variance explained by the first three principle components (i.e., those used to create the plots) needs to be greater than 0.95 [30]; in this study, all values exceeded 0.95.

A chi-square test for association was performed to test whether the number of behaviours recorded related to the length of time spent at the otter enclosure. The corrected visitor data were used to calculate the number of behaviours recorded, and any questionnaires where the question regarding time spent at the enclosure was left blank were excluded. Number of behaviours recorded were combined into 5 categories for the chi-square test (0, 1-2, 3-4, 5-6, and 7-8) and time periods were classed as less than 2 mins, 2–5 mins, 6–10 mins, and over 10 mins. It is worth noting that, although visitors could have recorded up to 10 behaviours, this did not occur (one visitor did record 9 behaviours, but this was excluded from the analysis because the visitor was a young child and data accuracy was questionable).

2.5. Simulations to Test Accuracy of Visitor-Collected Data. The selection of the time period in which the visitors were asked to collect data was based on the concept that a 30 s period would capture more data than a single instantaneous scan, yet would not be likely to result in all behaviours being observed; hence an estimate of frequency of behaviours could be obtained using this method. Given that preliminary observations indicated that visitors vastly exceeded this time period (see below), a computer simulation was developed to determine if the 30 s sampling period would produce comparable data to ethogram recordings given assumptions that incorrect identification of behaviour and temporal autocorrelation of the data did not exist (i.e., data were collected perfectly, except for the time of recording). The simulation was constructed using *R* [35]. The simulation was parameterised according to the relative probability of the behaviours, as collected from ethogram recordings, making the assumption that the ethogram data collected in this study were an accurate representation of the otters' activity budget (see results, Figure 2).

The simulation produced a random number (score) between 1 and 100, which corresponded to a particular behaviour based on the proportion of its occurrence (see results for details, but otters were seen swimming 11% of the time, so a score between 1 and 11 would correspond to the behaviour "swimming"). After this initial score had been set, the simulation ran with a timestep of the

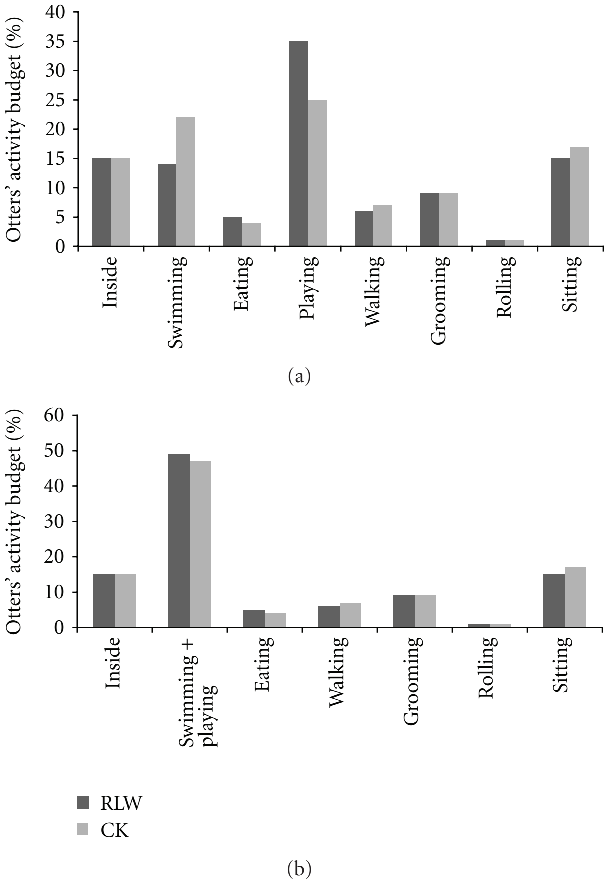


Figure 2: (a) Comparison of otters' activity budgets calculated from ethogram data collected by two biologists (RLW and CK) over one day. Note: categories "fighting" and "other" are not displayed on the graph because neither occurred on that day. (b) As above, swimming and are playing combined as one category.

simulation of 5 s. At each timestep, the score was modified by adding or subtracting a second, randomly generated number (between 3 and −3 from a uniform distribution), from the current score. This new score then indicated the behaviour of the otter at the next timestep. In practise, this meant that successive time steps normally resulted in the same behaviours being recorded, which corresponded to observations on behaviour (i.e., behavioural inertia is more likely than behavioural change).

To parameterise this alteration (named the "change by" variable), results from the ethogram recordings were used. Results indicated that the otters performed on average 3.6 behaviours in a 10 min period. Therefore, we systematically changed the "change by" variable, and for each value, we simulated 100,000 individuals 10 min periods (with sampling every two 5 s timesteps—equating to the 10 s recording periods that were used in this study) to produce a number of behaviours as close as possible to 3.6. The "change by" variable of 6 (i.e., between −3 and 3) produced the most accurate representation, producing an average of 3.5 behaviours over 10 min. (when the "change by" variable was 7 (±3.5), the model produced an average number of behaviours of 3.8, and when 5 (±2.5) produced an average of 3.2 behaviours).

We next simulated data that represented 30 s of sampling by visitors. Although these simulated data were free from confounds such as temporal autocorrelation and misidentification of behaviours, they would give an accurate indication of whether the 30 s recording period would have allowed visitors to collect accurate data on the otters' activity budget. As such, we simulated 574 visitor responses (the same number collected in the study). We compared simulated data and real visitor-collected data in terms of the number of behaviours recorded in a questionnaire to examine the average length of time that visitors may have recorded data for. We also compared the 30 s simulated visitor data to ethogram data and real visitor data using modified PCA or "bubble" analysis, to determine whether recording behaviour for 30 s would result in significant differences to either of these recording methods.

3. Results

3.1. Interobserver Variability. The activity budgets collected by the two biologists were very similar except for the categories of playing (35% for RLW and 25% for CK) and swimming (14% for RLW and 22% for CK). Because playing and swimming were sometimes difficult to differentiate (playing often occurred in water), the differences between the two activity budgets were less apparent when these categories were combined as a single category (Figures 2(a) and 2(b)). There was no significant difference between activity budgets collected by the two biologists. However, when playing and swimming were combined, the bubbles overlapped more, indicating greater similarity (Figures 3(a) and 3(b)).

3.2. Uptake of Questionnaires and Potential Errors. In total, 574 questionnaires were collected during the study. A very low number of visitors declined to fill in the questionnaire when they were asked (estimated at <5%), and the main reason given for this was that they did not have time. Of the questionnaires collected, 39.2% were collected outside of school holidays and 60.8% during the school holidays, reflecting the increase in visitor numbers in the centre. Some visitors left various questions unanswered in the otter behaviour questionnaire (Table 3). The segmentation questionnaire was completed by 62.4% of visitors who had filled in the otter behaviour questionnaire, but of these, 5.6% could not be used because visitors had not followed the instructions and had ticked more than one answer, meaning that they could not be classified into a visitor segment.

While the questionnaires were being filled in, personal observations indicated that visitors were watching the otters for longer than 30 s. This was reflected in the responses to the question concerning the length of time visitors had spent at the enclosure. A chi-square test showed that the length of time a visitor spent at the enclosure affected the number of behaviours recorded ($\chi^2 = 41.7$, $df = 12$, $P < 0.001$). This was because visitors who stayed at the otter enclosure for shorter lengths of time recorded significantly

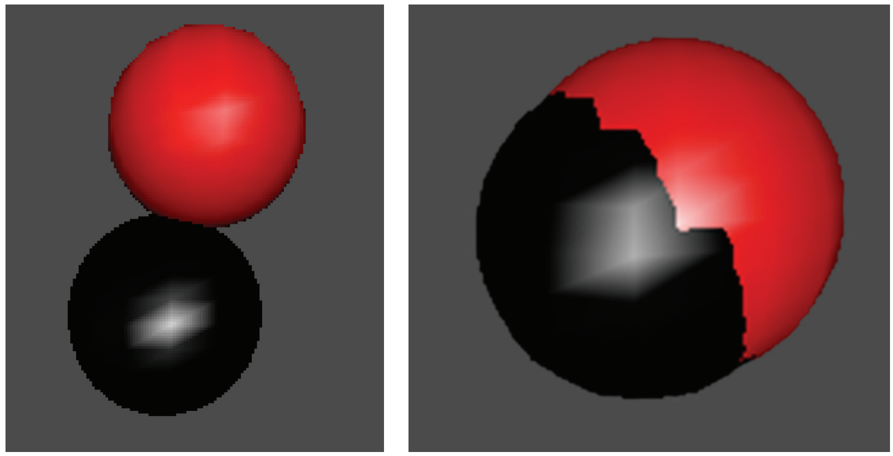


FIGURE 3: Results of bootstrapped PCA examining differences between ethogram data collectedby two biologists for the group of otters over one day. Black = RLW, red = CK. Cumulative proportionof variance explained by first 3 principal components > 0.999. (b) as above but with playing and swimming combined.

TABLE 3: Percentage of questions not answered in the otter behaviour questionnaire.

Question	Questionnaires where this was left unanswered
What time is it?	0.2%
Approximately how long have you spent at the otter enclosure in total today?	5.7%
Are you, or someone who helped fill in this questionnaire a member of any wildlife charities?	8.3%
Have you or anyone who helped fill in this questionnaire volunteered or done something to help any wildlife charities? (e.g., habitat improvement, wildlife surveys, helped at events, raised money, etc.)	11.6%
What age are you/the people who helped fill in this questionnaire? Write down the number of people in each age group.	9.9%

fewer behaviours than those who stayed at the enclosure for longer (mean number of behaviours recorded: <2 mins = 2.14; 2–5 mins = 2.34; 6–10 mins = 2.93, >10 mins = 3.33).

3.3. Comparing Ethogram Activity Budgets with Activity Budgets Calculated from Visitor Data. The otters' activity budget calculated using ethogram data consisted mainly of time spent inside (28%), followed by playing (21%) (Figure 4). "Other" behaviours (e.g., sprainting, drinking, climbing. . .), and rolling amounted to the smallest proportion of the activity budget (2%). Fighting is not represented in the ethogram activity budget, but visitors did record fighting (1%), and it was observed during the study (outside of the randomly allocated observation periods). Compared to the ethogram data, visitors underrecorded sitting, time spent inside and playing and overrecorded all of the other behaviours, with the exception of "other" in the corrected visitor data, which was identical to the ethogram data. The most noticeable differences between ethogram and visitor data lie between time spent inside (28% for ethogram data and 11% for visitor data) and swimming (10% for ethogram data and 25% for visitor data).

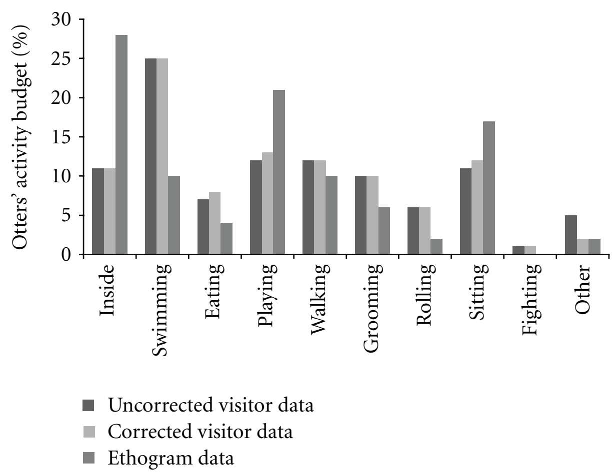


FIGURE 4: Differences in otters' activity budgets calculated using corrected and uncorrected visitor data and ethogram data.

There were significant differences between ethogram data and visitor data, but there were no significant differences between uncorrected visitor data and corrected visitor data (Figure 5). Additionally, there were no significant differences between each individual otter and the average taken for the group, so to simplify subsequent analyses, only corrected visitor data and ethogram data for the group of otters were used. Significant differences also occurred between ethogram data and data collected by different visitor segments, but there were no significant differences between the behavioural data recorded by different types of visitor (as quantified using the visitor segments used in the analysis: learn together families, fun time families, sensualists, social naturalists and expert birders, note: other segments could not be used because of small sample sizes) (Figure 6).

There was a significant difference between ethogram data and visitor data, but no significant difference between corrected visitor data before and after questionnaires filled in by children were excluded from the dataset. There was no significant difference between visitors who had prior experience volunteering, or were a member of a wildlife organisation and those who were not. All visitor datasets were still significantly different to the ethogram dataset (Figures 7(a) and 7(b)). There were still significant differences between ethogram and visitor data when playing and swimming were combined in the activity budgets and when visitor data was reclassified taking into account time periods in which the data had been collected (Figures 7(c) and 7(d)).

3.4. Simulation of Test Accuracy of Visitor Data Collection Methods. The average number of behaviours recorded by visitors in the study was 2.9, whereas the average number of behaviours recorded in the simulation running for 30 s was 1.4. Changing the length of time that visitors took to record behaviours in the simulation indicated that visitors may have watched the otters for up to 8 min, instead of following the instructions and recording behaviour for 30 s. Comparing the overall behaviour of all three otters combined using bootstrapped PCA demonstrated that there was no significant difference in overall behaviour when observations took place for 30 s (from simulated data) and the real ethogram data, but when compared with the longer 8 min observation period or the visitor collected data, significant differences to the ethogram data occurred (Figure 8).

4. Discussion

4.1. Visitors Cannot Accurately Collect Behavioural Data. The ethogram method used to determine otter activity budgets was repeatable between trained biologists, and this suggests that it is a reliable way of determining activity budgets. However, visitors were unable to collect accurate data on the otters' behaviour regardless of which visitor segment they were in, their age, prior experience volunteering or whether they were a member of a wildlife organisation. This did not differ when behaviours that overlapped (playing and swimming) were combined in the analysis, nor when much of the potential pseudoreplication caused by varying numbers of visitors throughout the day was removed. It may seem intuitive that an "expert birder" with experience of collecting scientific data on birds may be more likely to collect accurate data than a "fun time family" that is on a recreational trip, but this was not the case in this study.

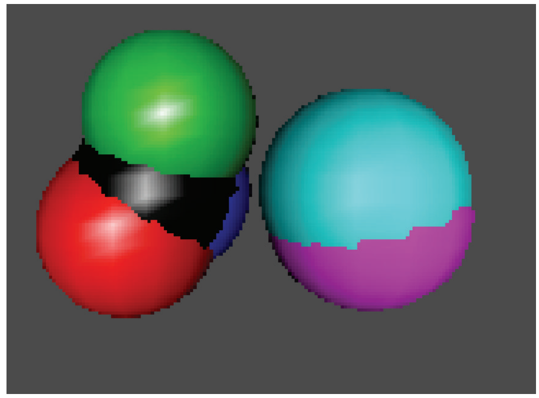


Figure 5: Results of bootstrapped PCA examining differences between ethogram and visitor data. Black = ethogram data for group of otters, red = ethogram data for otter 1, green = ethogram data for otter 2, dark blue = ethogram data for otter 3, light blue = corrected visitor data, and pink = uncorrected visitor data. Cumulative proportion of variance explained by first 3 principal components = 0.995.

4.2. Where Did They Go Wrong?

4.2.1. Ignoring the Instructions. One of the most important instructions on the questionnaire was the length of time required to observe the otters for. This length of time was chosen because it was thought to be short enough not to deter visitors from participating and would allow the recording data as and when visitors walked past the enclosure. Ease of data collection and reliability were both a key aspect of this study because visitors were assumed to be untrained. Therefore, 30 s was considered to be a reasonable length of time for visitors to scan the otter enclosure and be able to identify behaviours while imposing a time limit so that all visitors should spend approximately the same length of time recording data. Results of the simulation model of visitors undertaking 30 s sampling periods when filling in questionnaires showed that this length of time should have resulted in the accurate representation of the otters' activity budgets.

Despite the instruction to watch for 30 s being underlined and in bold font, most visitors did not follow this and recorded data for much longer than 30 s (pers. obs.). When visitors stayed longer at the otter enclosure, they ticked significantly more behaviours. This is probably one of the main reasons why their activity budgets were incorrect. In some cases, visitors admitted watching for longer. One visitor ticked rolling and wrote "when arrived," indicating that they felt this was an interesting behaviour and that they should record it, even though it was not in their 30 s recording period. Another visitor wrote "the otters came out at 10.36," which also indicates that they watched for longer than 30 s but may have thought that adding extra detail would benefit

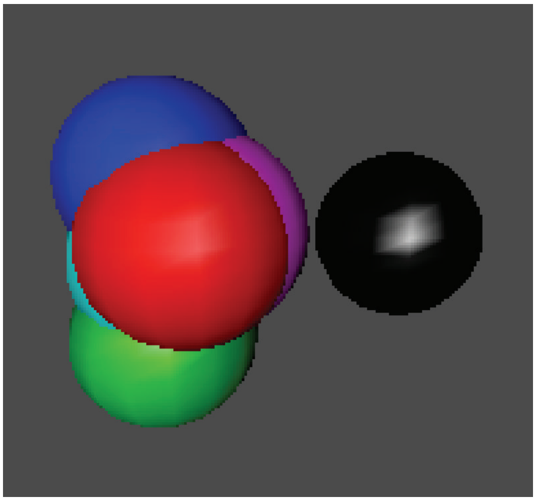


Figure 6: Results of bootstrapped PCA examining differences between ethogram data and different visitor segments. Black = ethogram data for group of otters, red = fun time families, green = sensualists, dark blue = social naturalists, light blue = expert birders, and pink = learn together families. No other visitor segments were included, since in total they contained <20 responses. Pairwise comparisons between social naturalists and sensualists also indicated no significant differences occurred between these categories. Cumulative proportion of variance explained by first 3 principal components = 0.997.

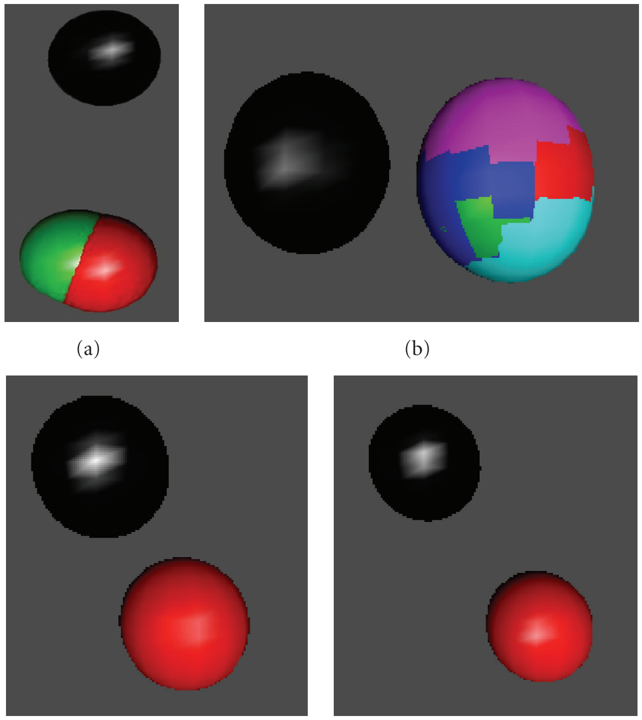
the study. At the end of one questionnaire that had been filled in by a parent and child (where all but one of the boxes had been ticked), the parent wrote, "hence saw all of the above because watched for a long time." Another visitor wrote that they "saw the otters outdoors earlier" so had filled their questionnaire in for a previous time (based on their memory of what they saw the otters do) as well as the present (when the otters were indoors), thus confounding their results. Some visitors demonstrated attention to detail by adding detailed notes on their questionnaires. However, these details are often impossible to analyse unless they can be reclassified, and this process can be time consuming (pers. obs). It seems that attention to detail and enthusiasm, while generally considered key attributes for volunteering, can hinder the quality of behavioural data collected.

4.2.2. Making Mistakes and Adding Extra Details. Occasionally, visitors admitted that they were wrong on their questionnaires, despite understanding the instructions. One visitor ticked rolling but wrote "in water" next to the box despite the fact that the behaviour was entitled "rolling—e.g. on soil or rocks", another ticked sitting but specified that the otters were indoors. However, only the obvious mistakes could be removed from the corrected dataset, and it is highly likely that some mistakes remained undetected (i.e., if visitors wrongly interpreted behaviours or deliberately ticked boxes even though they had not seen a particular behaviour). It was impossible to measure this. Furthermore, the question "What age are you/the people who helped fill in this questionnaire? Write down the number of people in each age group" could not be analysed because visitors misunderstood the question. Most visitors wrote down the number of people in their party, regardless of whether or not they had helped fill in the questionnaire.

The fact that visitors underrecorded sitting and time spent inside may be because these could be ignored if they appeared less interesting for visitors than more active behaviours. Sitting generally occurred for short periods of time (with otters pausing for a few seconds), in which case visitors could have missed this. The underrecording of time spent inside may have been caused by visitors missing otters inside if some of the otters were outside. If this was the case, visitors often observed the otters that were outside and did not check the sleeping chamber (pers. obs.). Another contributing factor could be that otters spent more time inside during quiet times when there were no visitors around to record this (early morning and late afternoon). The underrecording of playing is probably correlated with the overrecording of swimming; it is likely that some visitors confused the two behaviours and ticked swimming instead of playing when otters were playing in the water (Figures 2(a) and 2(b)). Playing may have been difficult for some visitors to interpret. Indeed, most "other" behaviours that were reclassified in the corrected dataset were reclassified as playing. However, removing mistakes and omissions and grouping behaviours did not change the overall results. This suggests that misidentification of behaviours by visitors was not the prime reason for the differences between ethogram and visitor activity budgets.

4.2.3. Item Nonresponse. Item nonresponse, in which a questionnaire is returned with one or more questions unanswered, can have an impact on results of a survey but these impacts are difficult to measure [37–39]. There could be various reasons why some visitors left questions blank (Table 3). For example, the visitor who missed out the question asking for the time may not have been able to find out what the time was as they did fill in all of the other questions. Boredom or rushing to finish the questionnaire may have been reasons why 1.6% of visitors filled in the time and ticked behaviours but did not answer any other questions that appeared later in the questionnaire [40]. It is also possible that some of the visitors who did not answer questions on the second page did not realise they were there, despite the staple and instruction "please turn over" in bold and underlined at the bottom of the first page: some visitors only realised this when another visitor pointed it out to them (pers. obs.). Another possibility is that visitors may not have wanted to fill in the questionnaire but felt obliged to do so out of politeness and as a result, may have rushed through the questions, missing some out.

This lack of attention to detail could be caused by the fact that the questionnaire was *impromptu*: visitors were on a day out not expecting to have to concentrate on a task. They may also have been distracted by the surrounding environment (e.g., by their children or by other visitors). Slightly more visitors avoided answering the question about volunteering than the question about being a member of a wildlife



(a) (b) (c) (d)

Figure 7: (a) Results of bootstrapped PCA examining differences between ethogram data, corrected visitor data, and uncorrected visitor data when all questionnaires filled in by children were removed from the dataset. Black = ethogram data for group of otters, red = children's questionnaires removed from corrected visitor data, and green = corrected visitor data. Cumulative proportion of variance explained by first 3 principal components >0.999. (b) As above but examining visitor segments. Black = ethogram data for group of otters, red = corrected visitor data, green = visitors who had previous experience volunteering, dark blue = visitors who did not have prior experience volunteering, light blue = visitors who were members of a wildlife organisation, and pink = visitors who were not members of a wildlife organisation. Cumulative proportion of variance explained by first 3 principal components = 0.995. (c) As above but examining ethogram data for group of otters and corrected visitor data when playing and swimming were combined. Black = ethogram data for group of otters and red = corrected visitor data. Cumulative proportion of variance explained by first 3 principal components >0.999. (d) As above but examining ethogram data and visitor data with standardised time periods. Black = ethogram data for group of otters and red = corrected visitor data. Cumulative proportion of variance explained by first 3 principal components = 0.987.

organisation or charity (Table 3). This may be because the membership question can be more easily interpreted, as membership to the WWT is well advertised throughout the centre and 57% of all visitors to the centre during the study were members of WWT. The volunteering question may confuse those who are unfamiliar with the idea of volunteering; one visitor said that she considered visiting the centre as volunteering (pers. comm.).

4.2.4. Temporal Autocorrelation of the Data. Questionnaires were handed to visitors as and when they arrived at the otter enclosure. As such, it is highly likely that some of the otters' behaviours were simultaneously recorded by many visitors, especially at busy times such as during the feeding demonstrations. While it would have been possible to hand out only one questionnaire at a time, such an approach would reduce the uptake of the questionnaire, and also would have a negative influence on visitor experience, with visitors either waiting a long time to participate or feeling left out if they could not participate. In a zoo environment, it would be very difficult to fully control the spread of questionnaires over time because of the irregular flow of visitors, not only at different times of day (e.g., when the centre first opens or when visitors are hurrying to leave before the closing time), but also in adverse weather conditions when visitors would be less likely to want to fill in a questionnaire. Additionally, there were often more visitors at the enclosure when the otters were active, with large crowds often attracting passers by because the formation of a crowd could indicate that the otters were doing something interesting or unusual (pers. obs.). In this study, the averaging of data over 30 min periods helped reduce autocorrelation effects due to the effects mentioned previously, but would not completely eliminate them if there was a difference in recorder effort within a 30 min period.

However, the effects of temporal autocorrelation on the results of this study appear minimal. Firstly, "standardised" data (where an average activity budget was calculated over each 30 min period taking into account the number of

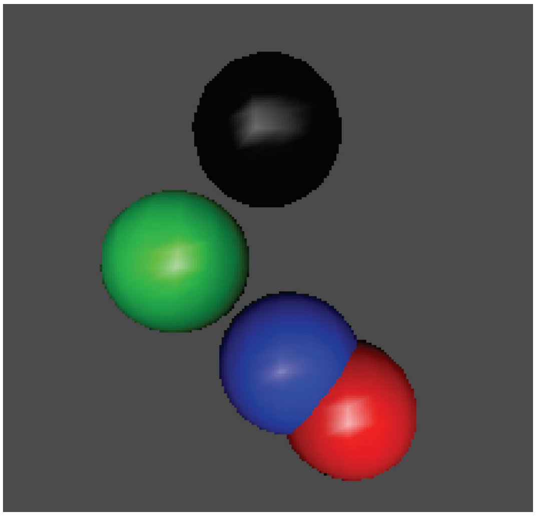


Figure 8: Results of bootstrapped PCA examining differences between real data and simulated data. Black = real visitor data, red = real ethogram data, green = simulated visitor data where data were collected for 8 min, and blue = simulated visitor data where data were collected for 30 s. Cumulative proportion of variance explained by first 3 principal components = 99.8.

questionnaires answered) and "unstandardised" data both differed significantly from ethogram data. Secondly, when data were simulated (and autocorrelation effects were eliminated) results corresponding to visitors collecting data for a long period of time (8 min) were highly significantly different from ethogram recordings. Hence, it appears that it was the length of time in which visitors recorded behaviour that was the largest source of error, rather than potential errors inherent to the sampling design used. Nevertheless, methods to eliminate temporal autocorrelation and enhance the visitor experience are given in the Recommendations Section.

4.3. A Success: The High Questionnaire Uptake Rate. The questionnaire uptake rate may not have been so high if the questionnaires had not been handed out in person [41]. Indeed, very few visitors were observed picking up a questionnaire themselves when the questionnaires were laid out on a wall next to the otter enclosure, despite posters advertising the study. In this situation, children were more curious than adults, often picking up questionnaires and filling them in of their own accord. Curiosity is a strong motivational force in children [42–44] and it is often believed that curiosity decreases with age [44], which may explain why fewer adults picked questionnaires up. Distributing questionnaires in the manner described in this study could cause logistical problems for zoos (for financial and temporal reasons discussed in Section 1). However, it may be possible that handing questionnaires upon entry to the park along with a quick explanation or instruction leaflet could be a suitable method to increase participation, similar to the method described in Dillman [41].

Uptake rate may be less high when animals are out of view or in an indoor area. As discussed previously, otters were less popular with visitors when they were inside, visitors walked past and/or did not see the point of filling in the questionnaire until it was explained that it was important to find out how much time the otters were spending inside. This has been discussed in previous studies. Indeed, Altman [45] and Anderson et al. [24] found that zoo visitors paid more attention to an animal's behaviour when the animals were most active compared to when they were less active or inactive. Jackson [46] and Johnston [47] found that visitors spent less time in front of enclosures where animals were inactive. Additionally, mammals are the most popular class in zoos [48], and larger animals may be preferred by visitors over smaller animals [49]. It is possible that a behavioural study would not prove as popular with visitors if it involved less appealing classes or species. Indeed, Hoff and Maple [50] found that some visitors deliberately avoided going to reptile exhibits.

4.4. Recommendations for Further Study. A visitor who had completed the questionnaire made the following comment: "you could tell us more about the otters than we could tell you". This statement underlies the concept of volunteer data collection: a scientist's work can be more reliable than that of a volunteer, as was the case in this study. However, it is the large number of volunteers that can make them a powerful tool for research. Although the method in this study did not allow visitors to collect accurate activity budgets, it did have some success. The high uptake rate suggests that getting visitors to collect data on active and entertaining animals can be successful. Public engagement and distributing the questionnaires by hand also undoubtedly had a major influence on the uptake rate.

Several improvements could be made in future research. When asking volunteers to collect behavioural data, it is important that behaviours are simple enough that volunteers can distinguish them without confusion. Clear instructions are needed when designing questionnaires, but in situations where a time limit is necessary, it is important to try to facilitate this to ensure that methods are followed as closely as possible, perhaps by providing a large clock in front of the enclosure. A time limit could also be imposed with the use of technology, for example, through multimedia or interactive video screens, which have previously been used in zoos and aquaria to convey information to visitors [51–53]. This type of technology has also been used by the National Marine Aquarium in Plymouth, UK to allow visitors to collect data on fish in an exhibit (pers. obs). Visitors could also collect data with the use of smart phone technology as this has already been used for other types of volunteer data collection [54]. Technology such as this may also reduce the number of questions that are unanswered by imposing a response, or could be used to eliminate any temporal autocorrelation of responses by either only having a single display, or by accurately recording the time of the response, so replication in time can be removed.

Overall, many of the aims of volunteering were completed in this study as visitors were keen to participate, enjoyed observing the otters, gave positive feedback, and asked questions about the study. Visitors were generally able to recognise different behaviours and recorded a rare behaviour that the scan sampling method did not detect [27]. They were also often eager to provide detailed notes on their observations. The "*ad libitum*" behaviour sampling method may be more suited to volunteers as it would remove the need for a restrictive time limit and would allow volunteers to record behaviours as they wished. This technique is commonly used in preliminary studies or to record rare but important events [27]. However, data collected in this manner would be difficult to analyse and could not be used to calculate activity budgets. New data collection techniques need to be tested if volunteers are to be used to collect behavioural data effectively.

Acknowledgments

Many thanks are due to Ann Nicol for providing the authors with the WWT visitor segmentation questionnaire, John Crooks for his help setting up and advertising the study, and the anonymous reviewer and academic editor for their helpful comments to improve the paper. Thanks also are due to our three colleagues who helped hand out questionnaires: Dr. Rick Stafford, Christina Catlin-Groves, and Claire Kirkhope.

References

[1] B. M. Md-Zain, N. M. Yen, and I. A. Ghani, "Daily activity budgets and enrichment activity effect on chimpanzees (Pan troglodytes) in captivity," *Sains Malaysiana*, vol. 37, no. 1, pp. 15–19, 2008.

[2] P. A. Rees, "Activity budgets and the relationship between feeding and stereotypic behaviors in asian elephants (Elephas maximus) in a zoo," *Zoo Biology*, vol. 28, no. 2, pp. 79–97, 2009.

[3] A. Berger, K. M. Scheibe, K. Eichhorn, A. Scheibe, and J. Streich, "Diurnal and ultradian rhythms of behaviour in a mare group of Przewalski horse (Equus ferus przewalskii), measured through one year under semi-reserve conditions," *Applied Animal Behaviour Science*, vol. 64, no. 1, pp. 1–17, 1999.

[4] D. S. Schmeller, P. Y. Henry, R. Julliard et al., "Advantages of volunteer-based biodiversity monitoring in Europe," *Conservation Biology*, vol. 23, no. 2, pp. 307–316, 2009.

[5] R. Stafford, A. G. Hart, L. Collins et al., "Eu-social science: the role of internet social networks in the collection of bee biodiversity data," *PLoS ONE*, vol. 5, no. 12, Article ID e14381, 2010.

[6] P. G. Finn, N. S. Udy, S. J. Baltais, K. Price, and L. Coles, "Assessing the quality of seagrass data collected by community volunteers in Moreton Bay Marine Park, Australia," *Environmental Conservation*, vol. 37, no. 1, pp. 83–89, 2010.

[7] RSPB, The Big Garden Birdwatch, 2011, http://www.rspb.org.uk/birdwatch/results.aspx.

[8] J. Foster-Smith and S. M. Evans, "The value of marine ecological data collected by volunteers," *Biological Conservation*, vol. 113, no. 2, pp. 199–213, 2003.

[9] C. Kremen, K. S. Ullman, and R. W. Thorp, "Evaluating the quality of citizen-scientist data on pollinator communities," *Conservation Biology*, vol. 25, no. 3, pp. 607–617, 2011.

[10] W. R. T. Darwall and N. K. Dulvy, "An evaluation of the suitability of non-specialist volunteer researchers for coral reef fish surveys. Mafia Island, Tanzania—a case study," *Biological Conservation*, vol. 78, no. 3, pp. 223–231, 1996.

[11] S. R. Engel and J. R. Voshell, "Volunteer biological monitoring: can it accurately assess the ecological condition of streams?" *American Entomologist*, vol. 48, pp. 164–177, 2002.

[12] C. Newman, C. D. Buesching, and D. W. Macdonald, "Validating mammal monitoring methods and assessing the performance of volunteers in wildlife conservation—'Sed quis custodiet ipsos custodies?'," *Biological Conservation*, vol. 113, no. 2, pp. 189–197, 2003.

[13] P. Y. Henry, S. Lengyel, P. Nowicki et al., "Integrating ongoing biodiversity monitoring: potential benefits and methods," *Biodiversity and Conservation*, vol. 17, no. 14, pp. 3357–3382, 2008.

[14] J. K. Szabo, P. A. Vesk, P. W. J. Baxter, and H. P. Possingham, "Regional avian species declines estimated from volunteer-collected long-term data using List Length Analysis," *Ecological Applications*, vol. 20, no. 8, pp. 2157–2169, 2010.

[15] S. P. Hill and D. M. Broom, "Measuring zoo animal welfare: theory and practice," *Zoo Biology*, vol. 28, no. 6, pp. 531–544, 2009.

[16] A. M. Claxton, "The potential of the human-animal relationship as an environmental enrichment for the welfare of zoo-housed animals," *Applied Animal Behaviour Science*, vol. 133, no. 1-2, pp. 1–10, 2011.

[17] C. Reed, Keeping Research Alive—The role of zookeepers in research at Drusillas Park, BIAZA 8th Annual Symposium on Zoo Research, Colchester Zoo, 2006, http://tinyurl.com/cjxh33r.

[18] S. W. Margulis and E. J. Westhus, "Evaluation of different observational sampling regimes for use in zoological parks," *Applied Animal Behaviour Science*, vol. 110, no. 3-4, pp. 363–376, 2008.

[19] Bristol Conservation and Science Foundation, Zoo research: previous research projects, 2008, http://www.bcsf.org.uk/previous-research-projects.

[20] Colchester Zoo, Research projects, 2012, http://www.colchester-zoo.co.uk/index.cfm?fa=research.list.

[21] M. Gusset and G. Dick, "The global reach of zoos and aquariums in visitor numbers and conservation expenditures," *Zoo Biology*, vol. 30, no. 5, pp. 566–569, 2011.

[22] G. Broad, "Visitor profile and evaluation of informal education at Jersey Zoo," *Dodo*, vol. 32, pp. 166–192, 1996.

[23] R. Ballantyne, J. Packer, K. Hughes, and L. Dierking, "Conservation learning in wildlife tourism settings: lessons from research in zoos and aquariums," *Environmental Education Research*, vol. 13, no. 3, pp. 367–383, 2007.

[24] U. S. Anderson, A. S. Kelling, R. Pressley-Keough, M. A. Bloomsmith, and T. L. Maple, "Enhancing the zoo visitor's experience by public animal training and oral interpretation at an otter exhibit," *Environment and Behavior*, vol. 35, no. 6, pp. 826–841, 2003.

[25] W. Canino and D. Powell, "Formal behavioral evaluation of enrichment programs on a zookeeper's schedule: a case study with a polar bear (Ursus maritimus) at the Bronx Zoo," *Zoo Biology*, vol. 29, no. 4, pp. 503–508, 2010.

[26] L. Collins, K. McDonald, and R. Stafford, "The role of casual observations by keepers in detecting repetitive behaviours in Amur leopards (Panthera pardus orientalis)," in *Proceedings of the 13th Annual BIAZA Research Symposium*, Bristol, UK, 2011.
[27] P. Martin and P. Bateson, *Measuring Behaviour: An Introductory Guide*, Cambridge University Press, 2007.
[28] M. Stamp Dawkins, *Observing Animal Behaviour: Design and Analysis of Quantitative Data*, Oxford University Press, 2007.
[29] P. Lehner, *Handbook of Ecological Methods*, Cambridge University Press, 2nd edition, 1996.
[30] R. Stafford, A. E. Goodenough, K. Slater et al., "Inferential and visual analysis of ethogram data using multivariate techniques," *Animal Behaviour*, vol. 83, no. 2, pp. 563–569, 2012.
[31] D. Sewell, T. J. C. Beebee, and R. A. Griffiths, "Optimising biodiversity assessments by volunteers: the application of occupancy modelling to large-scale amphibian surveys," *Biological Conservation*, vol. 143, no. 9, pp. 2102–2110, 2010.
[32] A. N. Oppenheim, *Questionnaire Design, Interviewing and Attitude Measurement*, Continuum, London, UK, 1992.
[33] M. Hargreaves McIntyre, Wildfowl and Wetlands Trust Visitor Segmentation report, commissioned by WWT, (Internal report), 2011.
[34] N. Borgers, J. Hox, and D. Sikkel, "Response quality in survey research with children and adolescents: the effect of labeled response options and vague quantifiers," *International Journal of Public Opinion Research*, vol. 15, no. 1, pp. 83–94, 2003.
[35] R Core Development Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2011.
[36] D. Adler and D. Murdoch, rgl: 3D Visualization Device System (OpenGL), R Package Version 0.77. 325, 2008, http://rgl.neoscientists.org.
[37] R. Ferber, "Item nonresponse in a consumer survey," *Public Opinion Quarterly*, vol. 30, no. 3, pp. 399–415, 1996.
[38] R. M. Durand, H. J. Guffey, and J. M. Planchon, "An examination of the random vs. non random nature of item omissions," *Journal of Marketing Research*, vol. 20, no. 3, pp. 305–313, 1983.
[39] M. Denscombe, "Item non-response rates: a comparison of online and paper questionnaires," *Journal of Social Research Methodology*, vol. 12, no. 4, pp. 281–291, 2009.
[40] A. I. Kraut, A. D. Wolfson, and A. Rothenberg, "Some effects of position on opinion survey items," *Journal of Applied Psychology*, vol. 60, no. 6, pp. 774–776, 1975.
[41] D. A. Dillman, "The design and administration of mail surveys," *Annual Review of Sociology*, vol. 17, pp. 225–249, 1975.
[42] R. W. Peterson and L. F. Lowery, "A study of curiosity factors in first grade children," *Science Education*, vol. 52, no. 4, pp. 347–352, 1968.
[43] J. A. Jenkins, "An experimental investigation of the effects of structured science experiences on curiosity among fourth grade children," *Journal of Research in Science Training*, vol. 6, no. 2, pp. 128–135, 1969.
[44] A. Chak, "Teachers' and parents' conceptions of children's curiosity and exploration," *International Journal of Early Years Education*, vol. 15, no. 2, pp. 141–159, 2007.
[45] J. D. Altman, "Animal activity and visitor learning at the zoo," *Anthrozoos*, vol. 11, no. 1, pp. 12–21, 1998.
[46] D. M. Jackson, "Animal activity and presence of docent interaction: visitor behavior at Zoo Atlanta," *Visitor Behavior*, vol. 9, no. 1, article 16, 1994.
[47] R. J. Johnston, "Exogenous factors and visitor behavior a regression analysis of exhibit viewing time," *Environment and Behavior*, vol. 30, no. 3, pp. 322–347, 1998.
[48] A. Moss and M. Esson, "Visitor interest in zoo animals and the implications for collection planning and zoo education programmes," *Zoo Biology*, vol. 29, no. 6, pp. 715–731, 2010.
[49] P. I. Ward, N. Mosberger, C. Kistler, and O. Fischer, "The relationship between popularity and body size in zoo animals," *Conservation Biology*, vol. 12, no. 6, pp. 1408–1411, 1998.
[50] M. P. Hoff and T. L. Maple, "Sex and age differences in the avoidance of reptile exhibits by zoo visitors," *Zoo Biology*, vol. 1, no. 3, pp. 263–269, 2005.
[51] D. Semani, T. Bouwmans, C. Frelicot, and P. Courtellemont, "Automatic fish recognition in interactive live videos," in *Proceedings of the 6th World Conference on Systemics, Cybernetics and Informatics, Image, acoustic, Speech and Signal Processing III (SCI '02)*, vol. 14, pp. 94–99, 2002.
[52] H. Hlavacs, F. Gelies, D. Blossey, and B. Klein, "A ubiquitous and interactive zoo guide system," *Lecture Notes in Computer Science*, vol. 3814, pp. 235–239, 2005.
[53] P. Lindemann-Matthies and T. Kamer, "The influence of an interactive educational approach on visitors' learning in a Swiss zoo," *Science Education*, vol. 90, no. 2, pp. 296–315, 2006.
[54] D. M. Aanensen, D. M. Huntley, E. J. Feil, F. Al-Own, and B. G. Spratt, "EpiCollect: linking smartphones to web applications for epidemiology, ecology and community data collection," *PLoS ONE*, vol. 4, no. 9, Article ID e6968, 2009.

Hindawi Publishing Corporation
International Journal of Zoology
Volume 2012, Article ID 980520, 4 pages
doi:10.1155/2012/980520

Research Article

Information Flows in Community-Based Monitoring Exercises in the Ecuadorian Amazon

Johan A. Oldekop,[1,2] Nathan K. Truelove,[1] Santiago Villamarín,[3] and Richard F. Preziosi[1]

[1] *Faculty of Life Sciences, The University of Manchester, Manchester M13 9PT, UK*
[2] *School of the Environment, Washington State University, Vancouver, WA 98686, USA*
[3] *Museo Ecuatoriano de Ciencias Naturales, Quito 17078976, Ecuador*

Correspondence should be addressed to Johan A. Oldekop, j.oldekop@gmx.net

Received 3 February 2012; Accepted 24 May 2012

Academic Editor: Simon Morgan

Community-based monitoring schemes provide alternatives to costly scientific monitoring projects. While evidence shows that local community inhabitants can consistently measure environmental changes, few studies have examined how learned monitoring skills get passed on within communities. Here, we trained members of indigenous Kichwa communities in the Ecuadorian Amazon to measure fern and dung beetle species richness and examined how well they could pass on the information they had learned to other members of their community. We subsequently compared locally gathered species richness data to estimates gathered by trained biologists. Our results provide further evidence that devolved monitoring protocols can provide similar data to that gathered by scientists. In addition, our results show that local inhabitants can effectively pass on learned information to other community members, which is particularly important for the longevity of community-based monitoring initiatives.

1. Introduction

Community-based monitoring schemes (CBMS) combine local traditional knowledge with existing organizational systems to measure ecological changes [1, 2]. Because CBMS can increase local understanding of environmental issues [3], they are considered capacity building exercises that provide evidence for local management decisions [4].

Evidence shows that, with appropriate training, CBMS can provide precise data on environmental processes. Danielsen et al. [5] show that trained community members are able to accurately monitor biomass and logging activities in India, Tanzania, and Madagascar. Similarly, Oldekop et al. [6] show that community inhabitants in Ecuador can use simple and cost-effective methodologies to provide fern species richness estimates that accurately reflect biodiversity patterns observed by scientists.

What has not yet been addressed, however, is how information gained by those attending training schemes is passed on to other community members. In other words, we do not know whether trained community members can train other people within their communities. This information is important for the creation of long-term and decentralized CBMS, where the majority of the collection and interpretation of data is directly managed by local communities and stakeholders [1, 2].

Here, we use a CBMS exercise with indigenous Kichwa communities in the Ecuadorian Amazon to assess the ability of locally trained community members to train other residents within their communities. Specifically, we compare species richness estimates of two biodiversity indicators, ferns [7] and dung beetles [8] (henceforth beetles), gathered by scientists, community members trained by scientists, and community members trained by the community members originally trained by scientists.

2. Methods

Exercises took place in the communities of San José de Payamino (henceforth Payamino) and Chontococha in

August and November 2008. The communities are located within the Sumaco Biosphere Reserve and are classified as areas of tropical forest [9].

Four men from each community took part in the exercises and received a typical regional day wage each ($10 per day) for the duration of each exercise (four days). Participants in each community were divided into two groups: an expert-trained group and a community-trained group. The expert-trained group received fern and dung beetle identification training from the scientist group, which was composed of two Ph.D. students (JAO, NKT) from The University of Manchester with several months experience conducting biodiversity monitoring of ferns and beetles in the region. The community-trained group received fern and dung beetle identification training from the expert-trained group. Information was therefore passed from the scientist group to the expert-trained group to the community-trained group. Expert-trained group participants were chosen on a volunteer basis, whereas the expert-trained group recruited the community-trained group participants. Preliminary results were presented during community meetings on subsequent visits in 2009.

3. Training

Evidence shows that training schemes increase the accuracy of CBMS [6]. Our study, therefore, only focused on comparisons between participants who had received training. Expert-trained groups received fern identification training for several weeks while working as field assistants with JAO and NKT, who conducted a larger regional biodiversity assessment. Expert-trained groups were taught how to differentiate beetle species during a single 30 min session. In the case of both ferns and beetles, the expert-trained groups were given information on the key physical characteristics of each taxonomic group but were not given specific information to differentiate between specific species. Once trained, the expert-trained groups were asked to recruit and train the community-trained groups using their choice of methods. While community-trained groups received training on fern identification several days after the expert-trained group had finished working with JAO and NKT, dung beetle identification training for both expert- and community-trained groups occurred during the same day. Despite having no time limit, training for each indicator lasted approximately 15 min in both communities and consisted of field visits to review ferns and sessions examining beetle specimens.

4. Sampling

In each community, the different groups (scientist, expert-trained and community-trained) sampled ferns and beetles along three 500 m transects situated in primary forest. Ferns were sampled along each transect in 10 equally spaced 5 × 5 m quadrats; groups were specifically asked not to remove ferns or break off samples for comparisons. Groups sampled transects in random order and were not allowed to sample transects before previous groups had finished.

Table 1: Matched-pairs analysis results showing differences in species richness estimates between scientist, expert-trained, and community-trained groups.

Indicator	Community	Richness estimates*
Fern richness	Payamino	Scientist: 5.8^a
		Expert trained: 7.6^b
		Community trained: 5.9^a
	Chontacocha	Scientist: 5.6^a
		Expert trained: 5.4^a
		Community trained: 4.6^b
Beetle richness	Payamino	Scientist: 11.2^b
		Expert trained: 8.2^a
		Community trained: 8.6^a
	Chontacocha	Scientist: 10.7^a
		Expert trained: 8.8^b
		Community trained: 13.2^c

*Values that are not connected with the same letter differ significantly from each other.

Beetles were sampled using dung-baited pit-fall traps placed in each quadrat. Traps were collected after 24 hrs, and beetles were stored in 95% ethanol. Expert-trained and community-trained groups were then asked to determine the species richness of each trap. An expert taxonomist (SV) confirmed beetle species richness after the exercise.

5. Analysis

Correlations were analyzed separately for each community and each indicator. The accuracy, the amount by which groups over or underestimated species richness, was analyzed using paired t-tests. All analyses were performed in JMP8 (SAS Institute Inc.).

6. Results

With the exception of beetle richness in Chontacocha (Figure 1(d)) all richness estimates between scientist and expert-trained groups correlated significantly and positively (Figure 1(a)–(c)). All correlations between expert-trained and community-trained groups were positive and significant (Figure 1(e)–(h)). Only beetle species richness estimates in Payamino (Figure 1(j)) and fern species richness estimates in Chontacocha (Figure 1(k)) were significantly and positively correlated between the scientist and community-trained groups. Both expert- and community-trained groups over and under estimated species richness but there is no discernable pattern (Table 1).

7. Discussion

Results show substantial positive and significant correlations between the different groups, suggesting that information was passed successfully between the groups. These results confirm previous findings that CBMS can show similar

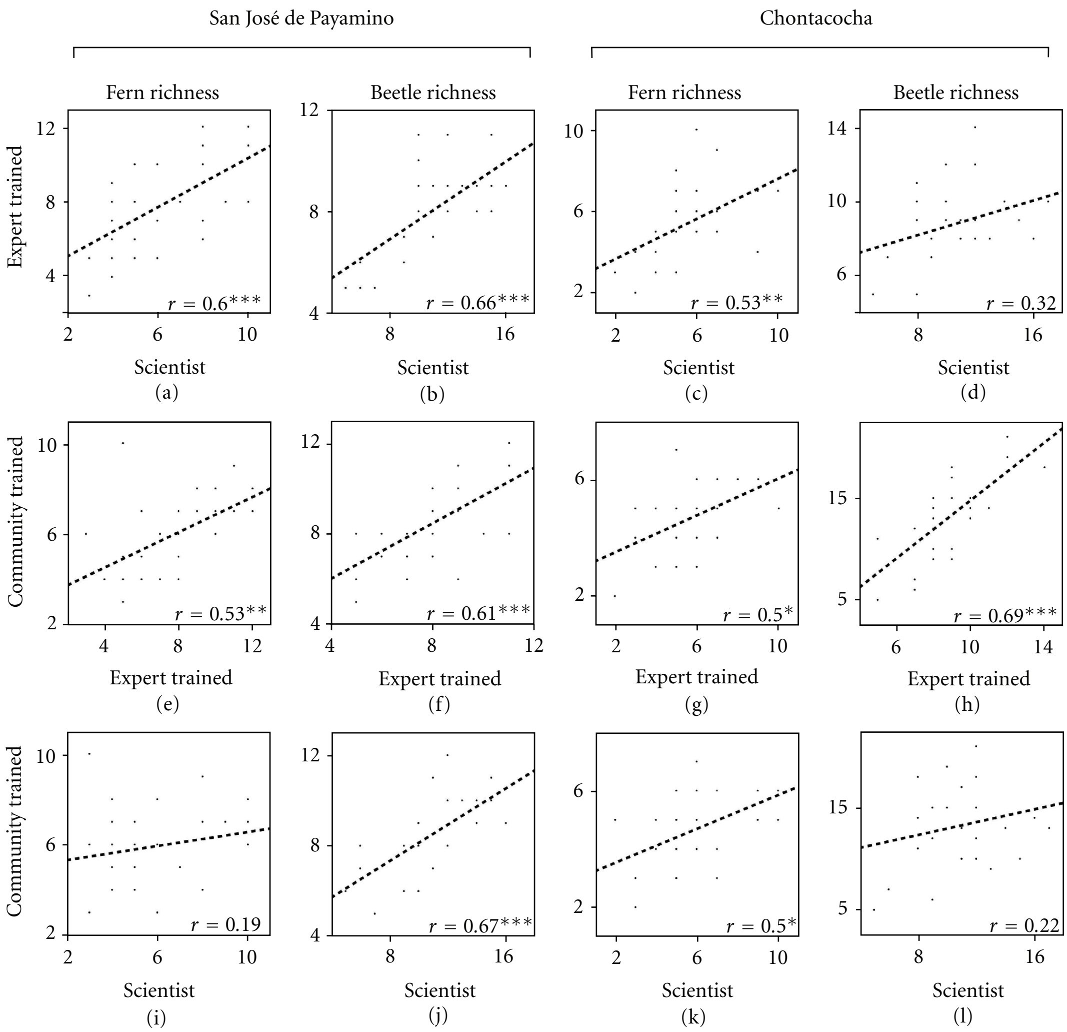


FIGURE 1: Correlations between scientist, expert-trained, and community-trained groups ($^{*}P < 0.05$, $^{**}P < 0.005$, and $^{***}P < 0.0005$).

trends to those found by scientists [4–6]. Beetle richness estimates between scientist and expert-trained groups showed no significant correlation in Chontacocha. Expert-trained groups, however, only received dung beetle identification training for 30 min. While Payamino's significant correlations between the scientist and expert-trained group suggest that community members have the ability to learn complex taxonomic information in a short time, a single 30 min session might not always be sufficient. Conversely, fern species identification training of the expert-trained group took place over several weeks, and the correlations between the scientist and the expert-trained group were positive and significant in both communities. While long training schemes are not always feasible, there is evidence that CBMS participants can learn taxonomic identification skills in substantially shorter periods of time [6].

Correlations between expert- and community-trained groups suggest that the information flow is particularly strong between community members. Poor beetle richness estimate correlations between the scientist and the community-trained groups in Chontacocha might be explained by a poor information flow between the scientist and the expert-trained group. The nonsignificant correlation of fern richness estimates between the scientist and the community-trained group in Payamino, however, suggest that there is a significant loss of information passed from the expert-trained to the community-trained groups. Despite the potential for information loss at these two points of communication, our data suggest that there is a large potential for community members to train other members within their communities, but that the way that information is transmitted to those individuals trained by professional scientists is critical.

Significant over- and underestimations of fern and dung beetle richness suggest significant errors in the accuracy of species richness estimates. Non-experts taking part in monitoring exercises commonly fail to recognize certain species as being either the same or different [10]. Differences between groups in our data are likely due to similar species "lumping" and "splitting" events.

The success of monitoring schemes relies on adapting methodologies to specific needs. While CBMS might not necessarily provide as detailed information as monitoring exercise performed by trained scientists [11, 12], they can lead to quicker decision making [5, 13]. An important factor influencing the longevity of CBMS is their ability to become less dependent on external expertise and resources. Although only based on a few comparisons, our data suggest that participants are remarkably good at passing on learnt information. Of key importance, however, is how initial information gets passed on from scientific experts to local practitioners and community-based monitoring initiatives. If we, the scientific community, can devise simple and accurate training methodologies that can be easily taught, learned, and implemented, then CBMS can provide a powerful and locally relevant tool to measure changes in biodiversity, natural resources, and ecosystem services.

Acknowledgments

The authors thank the participants and the communities of Payamino and Chontacocha for allowing them to work on their lands. Ecuadorian Ministry of the Environment research permit no. 005-08 IC-FAU-DNBAPVS/MA.

References

[1] F. Danielsen, M. M. Mendoza, P. Alviola et al., "Biodiversity monitoring in developing countries: what are we trying to achieve?" *Oryx*, vol. 37, no. 4, pp. 407–409, 2003.

[2] F. Danielsen, N. D. Burgess, A. Balmford et al., "Local participation in natural resource monitoring: a characterization of approaches," *Conservation Biology*, vol. 23, no. 1, pp. 31–42, 2009.

[3] H. T. Andrianandrasana, J. Randriamahefasoa, J. Durbin, R. E. Lewis, and J. H. Ratsimbazafy, "Participatory ecological monitoring of the Alaotra wetlands in Madagascar," *Biodiversity and Conservation*, vol. 14, no. 11, pp. 2757–2774, 2005.

[4] J. G. Mueller, I. H. B. Assanou, I. D. Guimbo, and A. M. Almedom, "Evaluating rapid participatory rural appraisal as an assessment of ethnoecological knowledge and local biodiversity patterns," *Conservation Biology*, vol. 24, no. 1, pp. 140–150, 2010.

[5] F. Danielsen, N. D. Burgess, P. M. Jensen, and K. Pirhofer-Walzl, "Environmental monitoring: the scale and speed of implementation varies according to the degree of peoples involvement," *Journal of Applied Ecology*, vol. 47, no. 6, pp. 1166–1168, 2010.

[6] J. A. Oldekop, A. J. Bebbington, F. Berdel, N. K. Truelove, T. Wiersberg, and R. F. Preziosi, "Testing the accuracy of non-experts in biodiversity monitoring exercises using fern species richness in the Ecuadorian Amazon," *Biodiversity and Conservation*, vol. 20, no. 12, pp. 2615–2626, 2011.

[7] R. Pardini, D. Faria, G. M. Accacio et al., "The challenge of maintaining Atlantic forest biodiversity: a multi-taxa conservation assessment of specialist and generalist species in an agro-forestry mosaic in Southern Bahia," *Biological Conservation*, vol. 142, no. 6, pp. 1178–1190, 2009.

[8] T. A. Gardner, J. Barlow, I. S. Araujo et al., "The cost-effectiveness of biodiversity surveys in tropical forests," *Ecology Letters*, vol. 11, no. 2, pp. 139–150, 2008.

[9] H. Navarrete, *Helechos comunes de la Amazonía baja Ecuatoriana*, Simbioe, Quito, Ecuador, 2001.

[10] I. Oliver and A. J. Beattie, "A possible method for the rapid assessment of biodiversity," *Conservation Biology*, vol. 7, no. 3, pp. 562–568, 1993.

[11] P. Z. Goldstein, "How many things are there? A reply to Oliver and Beattie, Beattie and Oliver, Oliver and Beattie, and Oliver and Beattie," *Conservation Biology*, vol. 11, no. 2, pp. 571–574, 1997.

[12] F. T. Krell, "Parataxonomy vs. taxonomy in biodiversity studies—pitfalls and applicability of "morphospecies" sorting," *Biodiversity and Conservation*, vol. 13, no. 4, pp. 795–812, 2004.

[13] J. van Rijsoort and Z. Jinfeng, "Participatory resource monitoring as a means for promoting social change in Yunnan, China," *Biodiversity and Conservation*, vol. 14, no. 11, pp. 2543–2573, 2005.

Hindawi Publishing Corporation
International Journal of Zoology
Volume 2012, Article ID 348653, 11 pages
doi:10.1155/2012/348653

Research Article

The Maryland Amphibian and Reptile Atlas: A Volunteer-Based Distributional Survey

Heather R. Cunningham,[1] Charles A. Davis,[1] Christopher W. Swarth,[2] and Glenn D. Therres[3]

[1] *The Natural History Society of Maryland, P.O. Box 18750, Baltimore, MD 21206, USA*
[2] *Jug Bay Wetlands Sanctuary, 1361 Wrighton Road, Lothian, MD 20711, USA*
[3] *Wildlife and Heritage Service, Maryland Department of Natural Resources, 580 Taylor Avenue, Annapolis, MD 21401, USA*

Correspondence should be addressed to Glenn D. Therres, gtherres@dnr.state.md.us

Received 5 March 2012; Accepted 19 April 2012

Academic Editor: Richard Stafford

Declines of amphibian and reptile populations are well documented. Yet a lack of understanding of their distribution may hinder conservation planning for these species. The Maryland Amphibian and Reptile Atlas project (MARA) was launched in 2010. This five-year, citizen science project will document the distribution of the 93 amphibian and reptile species in Maryland. During the 2010 and 2011 field seasons, 488 registered MARA volunteers collected 13,919 occurrence records that document 85 of Maryland's amphibian and reptile species, including 19 frog, 20 salamander, five lizard, 25 snake, and 16 turtle species. Thirteen of these species are of conservation concern in Maryland. The MARA will establish a baseline by which future changes in the distribution of populations of native herpetofauna can be assessed as well as provide information for immediate management actions for rare and threatened species. As a citizen science project it has the added benefit of educating citizens about native amphibian and reptile diversity and its ecological benefits—an important step in creating an informed society that actively participates in the long-term conservation of Maryland's nature heritage.

1. Introduction

Amphibian and reptile species are among the most threatened groups of vertebrate animals [1, 2]. Factors that lead to population declines are habitat alteration and loss, invasive species, disease, environmental pollution, commercial collection, and climate change [1, 3]. The lack of thorough understanding of regional distribution patterns of amphibian and reptile populations can limit our ability to predict how species will respond to these factors [4]. An additional challenge to the protection and conservation of amphibian and reptile species (also called herps) is the overall negative perception by the public towards these organisms [5]. There is a pervasive attitude that these organisms are unimportant [5]. However citizen science projects, defined as projects where citizens participate in scientific research [5], have the potential to advance the protection of amphibian and reptile species. Specifically, citizen science-based atlas projects can efficiently assemble distribution information across large spatial scales while increasing environmental awareness in the general public about the ecological importance of herpetofauna. Through participation in atlas projects citizens play an important role in the long-term protection and conservation of amphibians and reptiles.

Currently 93 native species of amphibians and reptiles occur in Maryland (20 anurans, 21 salamanders, 27 snakes, 19 turtles, and six lizards). The diversity of native herpetofauna is, in part, an outcome of the three physiographic provinces in Maryland: Appalachian, Piedmont, and Coastal Plain. Some species are restricted to particular provinces. For example, the Red-bellied Watersnake (*Nerodia erythrogaster*) and Eastern Tiger Salamander (*Ambystoma tigrinum*) are restricted to the Coastal Plain, and the Mountain Earthsnake (*Virginia valeriae pulchra*) and Green Salamander (*Aneides aeneus*) are found exclusively in the Appalachian Province. In 1975 Harris Jr., [6] compiled the most recent set of comprehensive maps for Maryland herpetofauna.

That publication maps the historic distribution of reptile and amphibian species in Maryland from the early 1900s through the mid 1970s using sightings and locality records of specimens collected and held in private collections, universities, museums, and with the Natural History Society of Maryland [7].

Since the Harris publication [6], Maryland has become much more urbanized, and natural lands have diminished. Additionally, land management practices, animal disease distribution, water pollution abatement practices, or climate have altered habitat suitability and population fitness of Maryland's herpetofauna. For example, researchers in Maryland found that particular timber management practices, specifically cutting and burning of small patches of forest can result in decreased local diversity of herpetofauna [8]. The amphibian chytrid fungus, *Batrachochytrium dendrobatidis,* which is responsible for the disease chytridiomycosis, now occurs in the Chesapeake and Ohio Canal National Historic Park, where researchers documented this pathogen in two species of stream-associated salamanders [9]. Changes to climate will inherently alter the hydrologic cycle [10]—a direct concern to habitat quality of regional herpetofauna. Further, Maryland researchers observed that in urban environments amphibians are attracted to stormwater retention ponds where they become exposed to trace metal contamination from accumulated runoff [11]. In addition to changing environmental conditions, relative to other states in the United States, Maryland ranks in the top 25th percentile for number of native amphibian and reptile species at risk of extinction [12]. Maryland has 11 species of reptiles and eight amphibians that have state conservation status of endangered, threatened, or in-need-of-conservation [13]. All of these factors reinforce the critical need to understand the current distribution patterns of amphibian and reptile species within the state.

Ecological atlases are one of the many types of citizen science projects that have been successfully conducted throughout the world [14]. Atlases show the distributions of organisms such as birds, butterflies, or plants for a given geographic area. Specifically, they are presence-only data sets of spatially explicit species occurrence data [15]. Most atlases use a predefined grid for sampling, employ a sampling protocol, have minimum requirements for the submission of data [16], and rely on citizen scientists to collect the data. The experience level of atlas volunteers can range from amateurs, with little scientific training, to scientists who specialize in the species of interest.

Atlas projects ultimately produce maps of the distribution of focal species. Repeated atlas efforts and comparisons can detect species' distributional shifts. Atlas data can inform scientists about these shifts and provide important information for focused studies on the causes. The first ecological atlas using a systematic approach to collecting field data was conducted in the United Kingdom for plants [17, 18]. In North America, many breeding bird atlases have been conducted in accordance with guidance provided by the North American Breeding Bird Atlas Committee [19] and subsequently published. Additionally, atlases for reptiles and amphibians have been completed or are in progress in several states in North America (see review [4]). In recent years, the methodology for citizen involvement and data collection for atlas projects has evolved with advances in new technology (e.g., the internet, social media, and digital photography). Current projects use the internet extensively as a means to gather and distribute information [4].

The Maryland Amphibian and Reptile Atlas project (MARA) is a citizen science project with the goal to document the current distribution of all amphibian and reptile species in Maryland using a systematic and repeatable approach during the five-year period, 2010–2014. Another important goal is to provide current information on the location and status of rare or threatened species. MARA data will inform management strategies for the immediate and long-term conservation and protection of Maryland's herpetofauna. Surveying large areas, such as an entire state, requires significant volunteer assistance and provides an opportunity to recruit and train novice participants, and thereby raises general ecological literacy by increasing awareness, skills, understanding, and knowledge of the natural world.

The Maryland Department of Natural Resources (MDNR) and the Natural History Society of Maryland (NHSM) cosponsor the project. The MDNR is the chief government agency responsible for conservation of the herpetofauna of the state. The NHSM, established in 1929, supports the community of amateur and professional naturalists within the state and during the intervening years has organized and published research on Maryland herpetology.

2. Approach

Conducting the atlas project entails four principal challenges: defining the survey methods, preparing data handling strategies, recruiting volunteers, and managing the volunteer network. The methodological foundation for MARA is based on herpetofauna atlases conducted in other states and two Maryland breeding bird atlases: 1983–1987 and 2002–2006 [4, 20–22]. Volunteer recruitment and management today can take advantage of online social networking to establish and maintain the research community for the project. We are implementing the atlas project by using these strategies.

2.1. The Atlas Survey Grid. Building an atlas on a grid base helps to meet the objective of using a systematic and repeatable method. The MARA uses a grid based on US Geological Survey 7.5-minute topographic quadrangle maps (called quads) divided into six equal blocks (Figure 1). The blocks are designated as northeast, northwest, central east, central west, southeast, and southwest. Each atlas block covers approximately 25 km^2. The state of Maryland includes 1,300 of these blocks within all or portions of 260 quads. This is the same grid system that was used successfully in the two previous breeding bird atlases conducted in Maryland [20, 21].

2.2. Search Effort Targets. In 2009, prior to initiating the MARA project, a pilot study was conducted in one Maryland

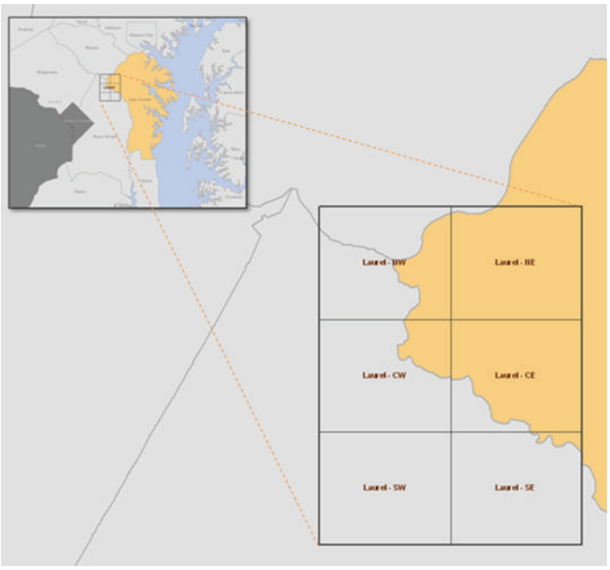


Figure 1: Example of a US Geological Survey 7.5-minute topographic quadrangle map divided into six atlas blocks.

county to develop the procedures and goals [23]. The pilot study helped to determine possible discovery rates of organisms per unit of survey effort. Providing an estimate of search effort is one way of standardizing results, assuming that the relationship between sampling effort and sample size is consistent [16]. In an earlier Maryland study, researchers documented that time-area searches were as or more effective than intensive trapping methods for documenting herpetofaunal presence [24]. To help assure dispersed geographic coverage across the state, the MARA Steering Committee established two goals for adequate coverage based on number of species discovered (at least ten species per atlas block and 25 species per quad) and the amount of time spent actively searching (at least 25 hours of active searching within each quad). Once these thresholds are reached in a block or quad, then surveyors should move to another less thoroughly searched area. In some blocks, which are highly urbanized, or dominated by one habitat type (i.e., Chesapeake Bay), reaching the ten species threshold may not be possible. In those instances, once the quad thresholds are met, surveyors are encouraged to move to a new quad. Cumulative time spent searching in a particular block is captured within the database as a record of effort for future comparisons. These guidelines balance the probable species numbers to be detected with the expected volunteer capacity to search the state within the project interval.

2.3. Time Period of Survey. Deciding the interval of the survey is strictly an estimate of the time to complete the effort based on the expected recruitment and effort of volunteers, as well as the anticipated detection rate of the herps. Ideally, this is the narrowest time interval possible so that the results will be least affected by any changes in distribution that may happen during the survey period. We chose a five-year period based on assessments of other state herpetological atlases (e.g., for Maine [25]), experiences from the pilot study, and experience with five-year survey periods used during the two Maryland bird atlases [20, 21].

2.4. Acquiring the Data. Data collected during active searching or incidental observations are recorded on a standard data sheet. Among the information recorded on the data sheet is locality and observer information; additionally the data sheet contains a checklist of all native species in the state by common names based on Crother et al. [26]. Surveyors complete one data sheet per year for each atlas block surveyed. Herpetological occurrence records are obtained by fundamentally two approaches: "active searching" or "incidental observations." Active searching is the main source of atlas data and involves intentional looking for reptiles and amphibians. There are no standardized methods used during active searching. Data are collected through listening surveys for calling frogs and toads, searching various habitats, turning over logs and cover boards, scanning ponds for turtles, turning over rocks along streams, and so forth.

We measure active searching by recording the amount of time a surveyor conducts searches in a given atlas block on each date of survey work. Incidental observations are sightings that observers make when they are not engaged in a formal, active searching survey. Incidental records are added to the database just as active searches, except no survey time is noted.

2.5. Additional Identity Verification Information. For most common species, no additional information is required although photographs for each occurrence record are encouraged similar to procedures of the Carolina Herp Atlas [4]. However, verification is required for certain species and surveyors complete an additional data form for those observations. Verification is required for rare species to obtain exact location information or for difficult-to-identify species so that we can be certain that identifications are accurate and correct. For example, photographic or written verification is required for Common Five-lined Skink (*Plestiodon fasciatus*) and Broad-headed Skink (*P. laticeps*) which are closely-related species and require close inspection of head scales to differentiate them. Similarly, the two species of gray treefrogs (*Hyla versicolor* and *H. chrysoscelis*) require vocal recordings because these identical species can only be distinguished by their calls. Verification is required for all records of eggs or larvae of any species. Photodocumentation is the main practice to verify species identity, but written documentation can be considered. In addition to those species for which verification is required, we encourage photodocumentation of each species found within a quad. Employing verification reduces misidentification of species, a serious error that can occur in citizen scientist projects [16].

2.6. Verification Committee. All photos, audio recordings of anuran calls, and verification forms are reviewed by a verification committee who are experienced with Maryland's herpetofauna (Figure 2). Photos and audio recordings are

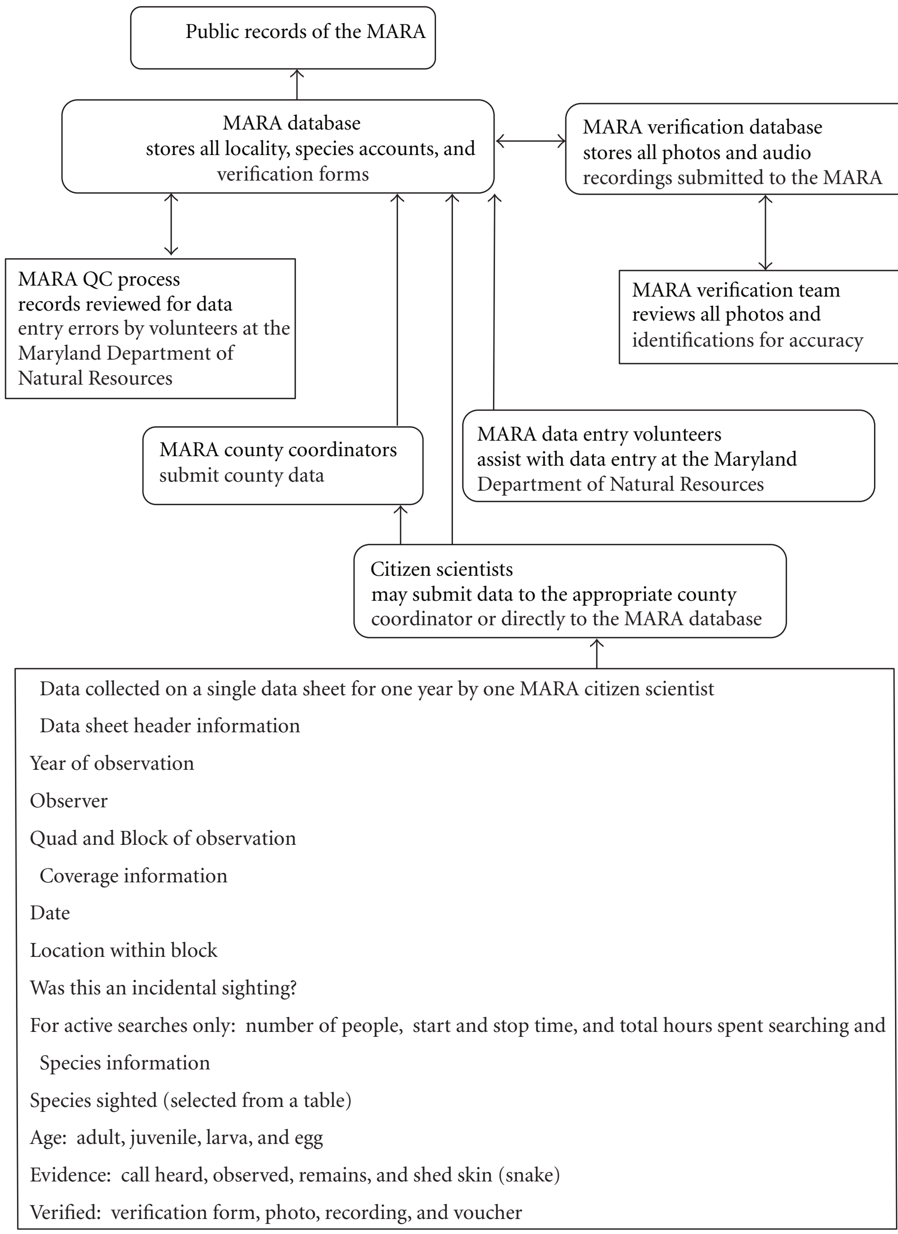


Figure 2: Data recorded and flow through the Maryland Amphibian and Reptile Atlas from collection to public release.

reviewed independently by each member of the committee via an online system. Only the photo or sound recording is provided to the reviewer with no additional information (no species name or observer name), except the county of occurrence. If the first three committee members to vote agree on the species identity, the record is confirmed; otherwise, the record is placed in a process to resolve the disputed record. Records are labeled "confirmed," if sufficient evidence was provided to the verification committee; "accepted" if no verification was required; or "unconfirmed" if the species required verification but none was provided or the verification committee could not positively determine the species' identity.

2.7. Database. Data are managed through a central MARA database, which was developed and maintained by the MDNR (https://webapps02.dnr.state.md.us/mara/default.aspx). Surveyors can enter data and submit verification photographs through online access to the database (Figure 2). County coordinators can access summary statistics here, view additional data summaries, and grant access for surveyors. Similarly, the verification committee accomplishes their duties through an associated portal to the submitted photographs. The central database is updated regularly with the decisions of the verification committee. Summary tables of the data are available on line to the volunteers, county coordinators, steering

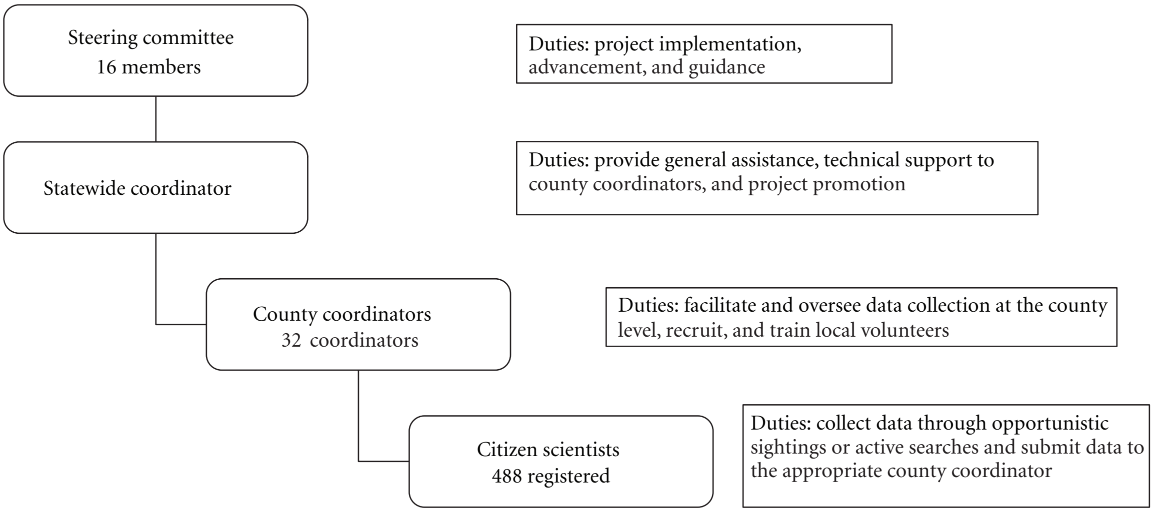


Figure 3: Organizational structure of the Maryland Amphibian and Reptile Atlas project.

committee members, and the public. This real-time data helps volunteers and coordinators plan for the field season and check on previous records. Allowing volunteers to examine the collected data is an important educational aspect of citizen science [27]. The MARA database will be used to produce distribution maps for all the reptile and amphibian species documented during the atlas project.

2.8. State Project Steering Committee. Recruiting, organizing, and mobilizing volunteers to achieve the project goals and objectives required oversight at statewide and local levels (Figure 3). Oversight of the MARA project is guided by a steering committee of 16 members. Successful citizen science projects require a development team comprising multiple partners and disciplines [27]. Three members cochair the steering committee, including representatives of the two primary project sponsors. The initial function of the steering committee was to develop the protocols of the distributional survey and recruit volunteer coordinators. As the atlas project got underway, the role of the steering committee shifted to project implementation and resolving technical issues. The steering committee meets monthly.

2.9. State Project Coordinator. To assist the county coordinators with recruiting, training, and motivating volunteers, and to develop strategies for collecting field data, a statewide coordinator was hired in August 2010. The statewide coordinator also conducts outreach efforts to promote the project, and to recruit and train additional volunteers. Additionally, the statewide coordinator produces a monthly project newsletter, educational materials relevant to the MARA project, and maintains the MARA website and social networking site. Federal wildlife grant money provided to the MDNR has been used to fund the statewide coordinator through a contract with the NHSM.

2.10. County Coordinators. Maryland contains 23 counties, and the MARA project chose to organize field data collection at this level. The county coordinators are integral to the success of the statewide effort. Coordinators work at the local level to recruit volunteers and to oversee the collection of field data within their counties. One to two coordinators were recruited for each county.

The county coordinators employ various strategies to achieve adequate coverage within their counties. They all recruit and coordinate volunteers to collect field data. A few counties are coordinated by employees of county park and recreation agencies that have access to networks of volunteers interested in reptiles and amphibians. These coordinators rely heavily on that network to obtain data. Other coordinators rely on a few dedicated volunteers to collect most of the field data within their county.

For the county coordinators, an annual meeting has been held each February beginning in 2010 to prepare for the upcoming field season. Topics of discussion at these meetings include how to recruit volunteers, strategies for achieving adequate coverage of blocks and quads, field techniques for finding reptiles and amphibians, and success stories presented by some of the county coordinators themselves. This meeting is an important annual check on progress of the project and a rally for the upcoming field season.

2.11. Recruiting Volunteers. Volunteers were recruited to the MARA project using various methods. Articles describing the MARA project and the need for volunteers were published in newspapers and nature club newsletters. Volunteers were also recruited at wildlife and nature festivals hosted at nature centers. MARA information was displayed before organizations such as the Maryland Association of Environmental and Outdoor Education in 2011 and 2012. Additionally, the two previous Maryland breeding bird atlases provided an existing network of citizen scientists experienced in atlas methods from which to recruit volunteers for the MARA through appeals at the Maryland Ornithological Society annual conferences in 2011 and 2012 and articles in their newsletter. Volunteers are also recruited through Facebook and Volunteer Match (http://www.volunteermatch.org/).

2.12. Training Volunteers. To aid the volunteers, several resources were developed to explain data collection and to ensure that data are assigned to the proper atlas block and quad. A handbook [28] was developed and provided to volunteers. The handbook explains the purpose of the project, grid system, data forms, techniques for finding reptiles and amphibians, and recommended references for species identification. The handbook also discusses health and safety precautions that are relevant to amphibian and reptile surveys. Measures for protecting habitat, animals, and surveyors are discussed in the handbook, covered at MARA training sessions and in the monthly project newsletter.

Paper copies of atlas blocks and a digital overlay of the grid system for use with Google Earth were made available to the volunteers. These and other resources are available on the project website (http://marylandnature.org/mara/) to guide data collection. In addition, county coordinators frequently held training sessions at the beginning of each field season to offer hands-on or technical training to volunteers. In some counties, coordinators held public hikes during which volunteers gained practical training while actively surveying for amphibians and reptiles.

2.13. Retaining Volunteers. As a means of retaining volunteers, two outreach products were developed. A monthly newsletter was initiated in November 2010. The newsletter encourages submissions from MARA participants. Photographs of interesting amphibian and reptile species encountered are highlighted in the newsletter in addition to accounts of experiences of MARA volunteers in the field. When county coordinators receive the newsletter by email, they send it to their local volunteers, and the newsletters are available to the general public through the project website. We established a Facebook page (https://www.facebook.com/MDHerpAtlas) in October 2011 to provide a forum for online exchange of project information and to encourage volunteers to communicate with each other. Volunteers are invited to post photographs and amphibian and reptile sighting information on the Facebook page.

3. Preliminary Results

3.1. Volunteer Recruitment. From January 1, 2010 to February 22, 2012, 488 citizen scientists have registered with the MARA database and contributed observations. This is a conservative estimate of the total number of actual contributors to the MARA, since this includes neither landowners that have granted access to private property nor those who assisted registered volunteers with active searching.

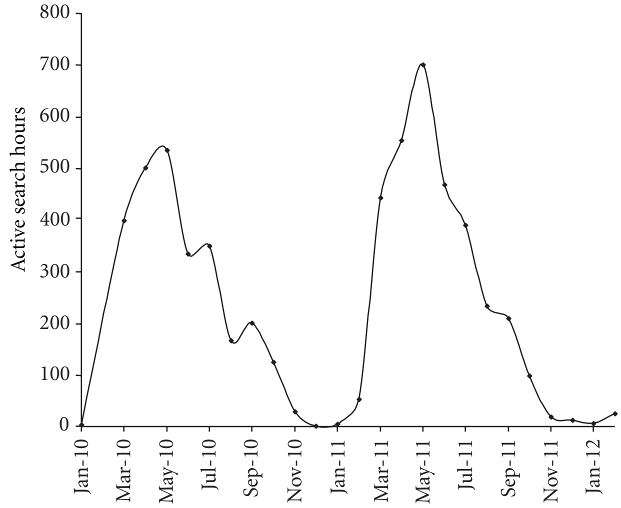


Figure 4: Survey hours for the Maryland Amphibian and Reptile Atlas.

3.2. Observation Counts. Surveyors have reported 13,919 occurrence records. This total includes 457 records from one county that were generated during the 2009 pilot study [23]. The mean number of occurrence records submitted per volunteer was 28.52 ± 106.13 (mean ± standard deviation, SD) with about 21% (5.89 ± 20.72) of the records being incidental observations (Table 1). Some observers contribute many more sightings than others. For instance, three observers submitted over 1,000 records each, five submitted between 500 and 1,000 each, and 39 submitted between 50 and 500 records. Sixty-nine percent of the volunteers submitted fewer than ten records each, and of those volunteers 186 submitted just one observation. Eighty percent of the records were contributed by 10% of the surveyors.

3.3. Time Effort Progress. Fifty-two percent (128/246) of quads have reached the minimum coverage goals of 25 species or 25 survey hours. Registered and unregistered volunteers have reported a total of 12,671 survey hours, but the number of survey hours reported per quad and block varied (Table 2). For instance, the total number of hours spent actively searching by all surveyors in particular quads ranged from 0 to 649 hours. The mean number of survey hours, or active searching hours, per registered participant was 12.48 ± 42.80 hours (Table 1). Survey effort occurred primarily throughout the spring and summer months, although records have been reported from every month of the year (Figure 4).

3.4. Areal Coverage Progress. Surveyors have collected data from 232 quads and 1,302 blocks (Figure 5). Thirty percent (74/246) of the quads have reached the minimum coverage goal of 25 species. Overall, surveyors observed 20.43 ± 10.44 and 8.27± 7.82 species per quad and block, respectively (Table 2). Forty-one percent (542/1,302) of the blocks have reached the minimum coverage goal of 10 species. On

Table 1: Number of occurrence records and search effort of registered MARA volunteers.

Volunteer contribution	Mean (per volunteer)	Mode	Frequency of mode	Minimum	Maximum	SD
Records submitted	28.52	1	186	0.00	1069.00	106.13
Opportunistic records	5.89	1	224	0.00	280.00	20.72
Active search hours	12.49	0	302	0.00	389.75	42.80
Quads surveyed	2.98	1	334	1.00	82.00	6.92
Blocks surveyed	5.53	1	316	1.00	284.00	18.76

Table 2: Species observed and person-hours of effort for blocks and quads.

Assessment	Count (quads/blocks)	Mean	Minimum	Maximum	SD
Species reported per quad	232	20.34	1	48.00	10.44
Species reported per block	1301	8.27	0	43.00	7.82
Person-hours per quad	246	51.51	0	649.00	79.55
Person-hours per block	1301	9.74	0	522.50	24.41

Table 3: Mean number of presence records per class/order per quad. Number of quads = 232.

Class/order	Mean	SD
Amphibia/Anura	27.43	26.03
Amphibia/Caudata	6.59	8.35
Reptilia/Squamata-Lacertilia	1.73	3.65
Reptilia/Squamata-Serpentes	12.62	13.70
Reptilia/Testudines	11.42	11.37

average, each registered participant collected data in 2.98 ± 6.92 quads and 5.53 ± 18.76 blocks (Table 1).

3.5. Species Detection Progress. Eighty-five of Maryland's 93 native amphibian and reptile species have at least one occurrence record in the MARA database. Distributional data have been recorded for 19 frog, 20 salamander, five lizard, 25 snake, and 16 turtle species. Volunteers also located 12 nonnative amphibian and reptile species. Anuran records comprise 45% (6,371/13,919) of the records in the database (Figure 6). The mean number of anuran records in surveyed quads was 27.43 ± 26.03 (Table 3). The second most commonly sighted group per quad was snakes, with salamanders and lizards sighted less frequently per quad. The most commonly reported species, per taxonomic order, were Spring Peeper (*Pseudacris crucifer*; 946 records), Red-backed Salamander (*Plethodon cinereus*; 431 records), Common Five-lined Skink (*Plestiodon fasciatus*; 212 records), Eastern Ratsnake (*Pantherophis alleghaniensis*; 617 records), and Eastern Box Turtle (*Terrapene carolina*; 665 records; Figure 3). Several records have also been collected on species listed as threatened or endangered in Maryland (Table 4). For example, volunteers submitted occurrence records for the state endangered Mountain Earthsnake (*Virginia valeriae pulchra*) from two previously undocumented locations. In addition, volunteers submitted two records for the state endangered Rainbow Snake (*Farancia erytrogramma*).

Voucher photographs or audio recordings accompanied 37% (5,203) of the submitted records. Through February 22, 2012, the verification committee has reviewed 4,406 records and determined that only 4% of the submitted records were misidentified.

4. Discussion

Public participation in the MARA resulted in a total of 13,919 occurrence records, in just 25 months. The MARA compares well with other successful herpetofauna atlases including the Georgia Herp Atlas [29] which collected 7,452 records during five years and the Carolina Herp Atlas [4] (currently underway) which collected 11,663 records during its first 36 months.

Though the number of registered participants was nearly 500 through the first two years of the MARA project, the majority of species records were submitted by less than 10% of the volunteers. This is consistent with the Carolina Herp Atlas [4]. However, the value of the few records submitted by the majority of the volunteers is also important. Their discoveries included difficult to find species, including salamanders and small snakes, and most of the nonnative species. Typically, volunteers participate in citizen scientist projects because they enjoy being able to put their skills to use in searching for and identifying species of the target group [16]. However, finding enough of these volunteers throughout the project area is the challenge in any widespread atlas effort.

The MARA project currently has a solid volunteer corps who freely contribute to the project by serving as county coordinators, steering committee members, and field workers. However, retaining volunteers for the duration of the project requires regular communication between the project management team and data collectors. This is important to ensure success of the project. Regular updates on progress are communicated via the project website, monthly electronic newsletter, and a social networking website. This communication is vital to ensure the participants remain interested and feel that the data being collected are being used [16]. The monthly newsletter and social networking site

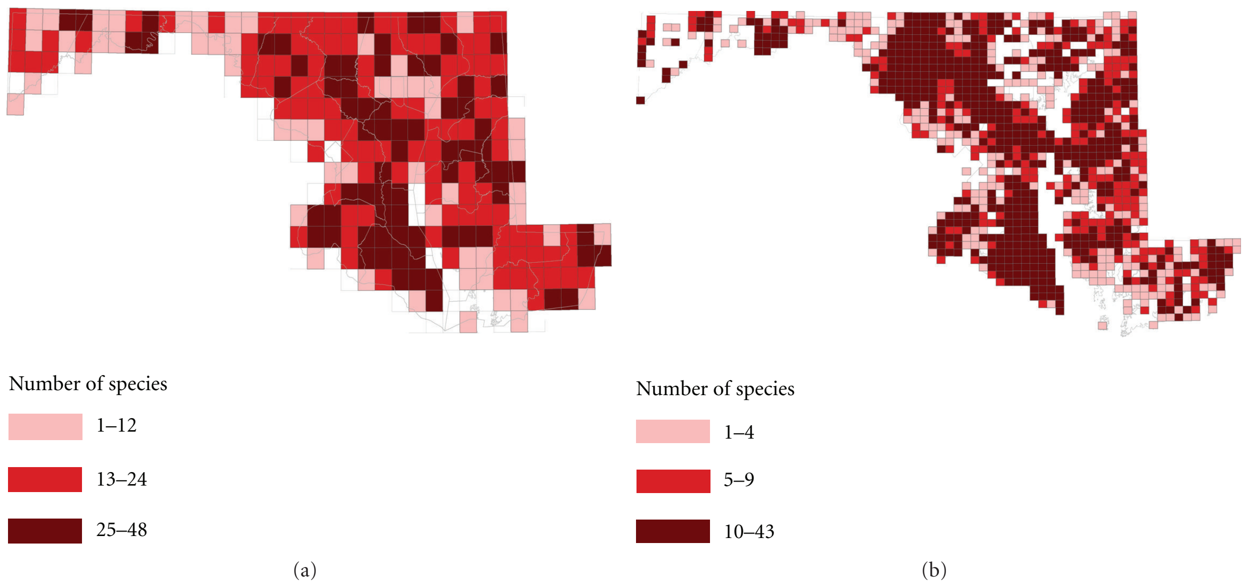


Figure 5: Results of the Maryland Amphibian and Reptile Atlas as of February 20, 2012. The number of species recorded in each quad (a) and block (b) is shown in a categorized series.

Table 4: Number of occurrence records of amphibian and reptile species submitted to the MARA that the state of Maryland lists as endangered (E) or threatened (T).

Class/order	Species	State status	Number of records
Amphibia/Anura	Barking Treefrog (*Hyla gratiosa*)	E	2
	Eastern Narrow-mouthed Toad (*Gastrophryne carolinensis*)	E	1
Amphibia/Caudata	Eastern Hellbender (*Cryptobranchus alleganiensis*)	E	1
	Eastern Tiger Salamander (*Ambystoma maculatum*)	E	5
	Green Salamander (*Aneides aeneus*)	E	5
Reptilia/Squamata-Serpentes	Mountain Earthsnake (*Virginia valeriae pulchra*)	E	3
	Rainbow Snake (*Farancia erytrogramma*)	E	2
Reptilia/Testudines	Bog Turtle (*Glyptemys muhlenbergii*)	T	27
	Green Sea Turtle (*Chelonia mydas*)	T	1
	Kemp Ridley Sea Turtle (*Lepidochelys kempii*)	E	5
	Leatherback Sea Turtle (*Dermochelys coriacea*)	E	1
	Loggerhead Sea Turtle (*Caretta caretta*)	T	14
	Northern Map Turtle (*Graptemys geographica*)	E	3

have been valuable tools to retain volunteers and increase communication among them.

To achieve consistent coverage, many atlas programs have standard benchmarks based on the number of species sampled within a given block or the hours of effort [14]. The MARA has set minimum coverage goals of 25 species or 25 survey hours per quad and 10 species per block. Currently, the MARA is on course to achieve these goals statewide by the end of 2014. We have achieved the minimum coverage goal of 25 active search hours within 52% of quads and the coverage goal of at least 10 species per block in 41% of blocks. With repeated atlases, often spanning twenty years or more, effort and the change in effort for individual blocks are crucial sources of variation that should be accounted for. When programs have no requirement for minimum effort, studies may be biased, resulting in overreporting of rare species, underreporting of common species, and failure to report repeated sightings because they are not deemed as interesting by the observer [14]. Further, people may simply stop sampling when there are no interesting organisms to be seen.

Atlasing efforts play an important role in biodiversity conservation by providing essential data on the occurrence of species [16]. The majority of records are anuran, most likely because anurans can be detected by sight and sound. To date, salamander and lizard records are not well represented in the MARA. The cryptic nature of salamanders and strict seasonal activity patterns of particular species (e.g., ambystomid species) likely contribute to the low number of salamander records. Of the six lizard species occurring within

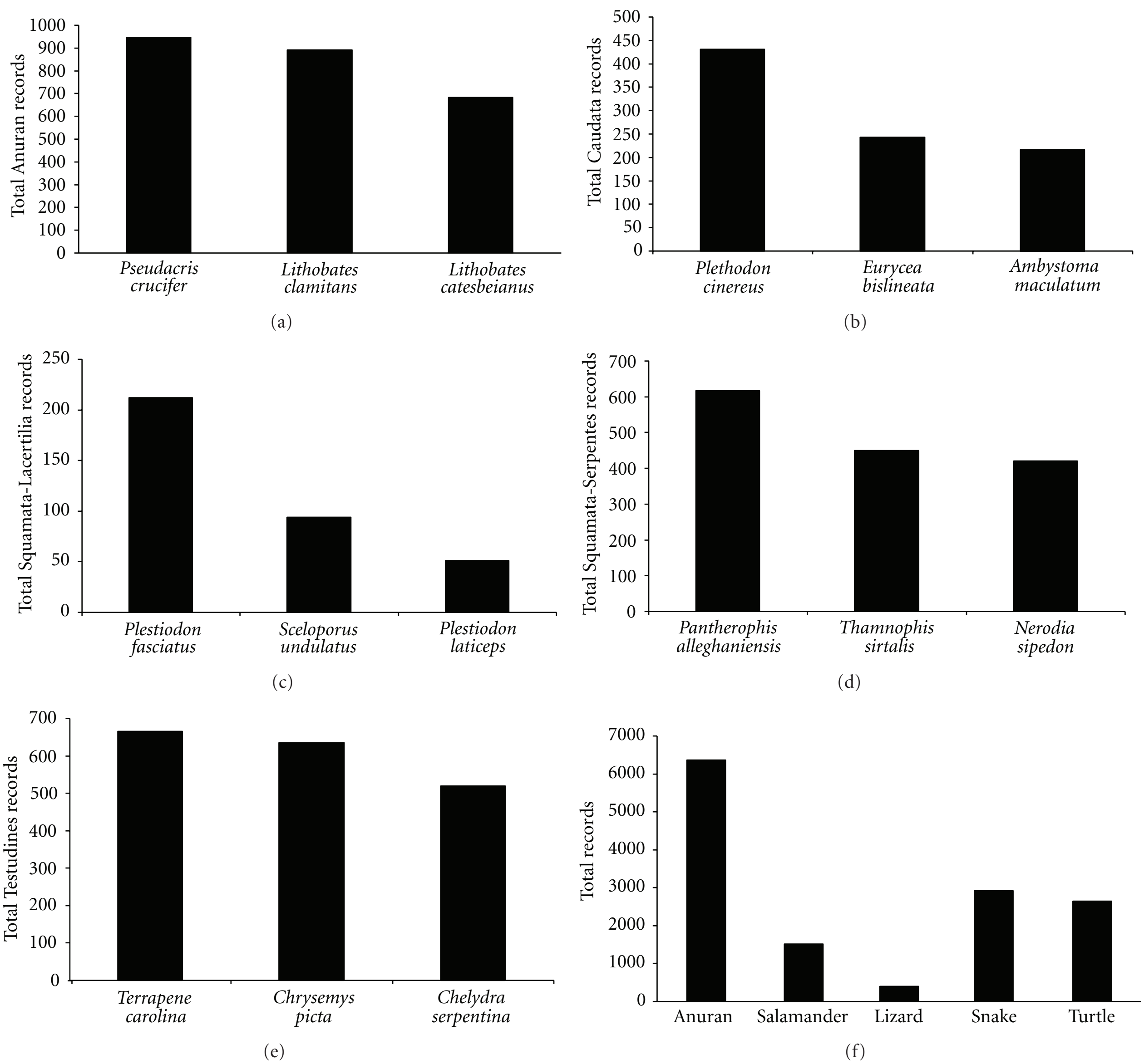


Figure 6: The most commonly reported species of frog (a), salamander (b), lizard (c), snake (d), turtle (e), and total number of records per group (f) through February 22, 2012.

Maryland, several have very restricted ranges which may explain the paucity of records. Several strategies have been helpful to increase records for these species. The availability of real-time data through the MARA database has allowed participants to be informed of the disparity in the record counts for the different groups. Information on how to survey for the underrepresented groups was shared with participants, resulting in an increase in record counts for those groups from 2010 to 2011.

Ultimately, the usefulness of atlas projects depends on the quality and quantity of data collected as well as the standardization of sampling methods and the appropriateness of the scale of sampling for the research question. An important function of our verification process is its capacity to quantify and correct error rates. The verification committee detected two primary error types during the early phase of the project: (1) animals were misidentified prior to data submission and (2) errors when submitting data to the online database, such as wrong photos submitted with a record or wrong species identity selected during data entry. The verification procedure enabled these errors to be identified and corrected prior to the records being finalized and the communication systems in place allowed feedback to the citizen scientists to reduce further errors. For example, in response to these findings, the MARA newsletter included, over the course of several months, articles containing information and techniques to identify the species that were found to be frequently misidentified and others discussing database usage and data entry.

Citizen science is perhaps the only practical way to achieve the geographic reach required to document ecological patterns and address ecological questions at scales relevant to species range shifts, broad-scale population trends, and impacts of environmental processes like landscape and climate change [14].

The MARA will establish a baseline by which future changes in the distribution of populations of native herpetofauna can be assessed. This project will be used to inform long-term conservation and protection strategies for Maryland's amphibian and reptile species. The MARA provides an opportunity for citizens to actively learn about native species while collecting valuable distributional data that the Natural History Society of Maryland and the Maryland Department of Natural Resources will use for the conservation and protection of Maryland's amphibians and reptiles. Educating citizens about native amphibian and reptile diversity and its ecological benefits is an important step in creating an informed society that actively participates in the long-term conservation and protection of Maryland's natural heritage.

Acknowledgments

The MARA is funded in part by State Wildlife Grant funds provided to the state wildlife agencies by US Congress and administered through the Maryland Department of Natural Resources' Wildlife and Heritage Service. Thanks go to all the MARA county coordinators, steering committee members, verification committee members, and volunteers who generously gave their time to the project. The authors extend gratitude to all the private landowners and public land managers who granted land access for surveys. Thanks are extended to Lynn Davidson and Bob Swan for developing the MARA database.

References

[1] J. W. Gibbons, D. E. Scott, T. J. Ryan et al., "The global decline of reptiles, deja vu amphibians," *BioScience*, vol. 50, no. 8, pp. 653–666, 2000.

[2] S. N. Stuart, J. S. Chanson, N. A. Cox et al., "Status and trends of amphibian declines and extinctions worldwide," *Science*, vol. 306, no. 5702, pp. 1783–1786, 2004.

[3] J. P. Collins and A. Storfer, "Global amphibian declines: sorting the hypotheses," *Diversity and Distributions*, vol. 9, no. 2, pp. 89–98, 2003.

[4] S. J. Price and M. E. Dorcas, "The Carolina Herp Atlas: an online, citizen-science approach to document amphibian and reptile occurrences," *Herpetological Conservation and Biology*, vol. 6, no. 2, pp. 287–296, 2011.

[5] J. W. Gibbons, "The management of amphibians, reptiles, and small mammals in North America: the need for an environmental attitude adjustment," in *Management of Amphibians, Reptiles, and Mammals in North America*, R. C. Szaro, Ed., USDA Forest Service General Technical Report RM-166, pp. 4–10, Rocky Mountain Forest Range Experiment Station, Fort Collins, Colo, USA, 1988.

[6] H. S. Harris Jr., "Distributional survey (Amphibia/Reptilia): Maryland and the District of Columbia," *Bulletin of the Maryland Herpetological Society*, vol. 11, no. 3, pp. 73–167, 1975.

[7] H. S. Harris Jr., "The past history of documenting the distribution of amphibians and reptiles of Maryland and the District of Columbia," *Bulletin of the Maryland Herpetological Society*, vol. 45, no. 1, pp. 14–16, 2009.

[8] R. F. McLeod and J. E. Gates, "Response of herpetofaunal communities to forest cutting and burning at Chesapeake Farms, Maryland," *American Midland Naturalist*, vol. 139, no. 1, pp. 164–177, 1998.

[9] E. H. Campbell Grant, L. L. Bailey, J. L. Ware, and K. L. Duncan, "Prevalence of the amphibian pathogen *Batrachochytrium dendrobatidis* in stream and wetland amphibians in Maryland, USA," *Applied Herpetology*, vol. 5, no. 3, pp. 233–241, 2008.

[10] R. T. Brooks, "Potential impacts of global climate change on the hydrology and ecology of ephemeral freshwater systems of the forests of the northeastern United States," *Climatic Change*, vol. 95, no. 3-4, pp. 469–483, 2009.

[11] R. E. Casey, A. N. Shaw, L. R. Massal, and J. W. Snodgrass, "Multimedia evaluation of trace metal distribution within stormwater retention ponds in suburban Maryland, USA," *Bulletin of Environmental Contamination and Toxicology*, vol. 74, no. 2, pp. 273–280, 2005.

[12] B. A. Stein, *States of the Union: Ranking America's Biodiversity*, NatureServe, Arlington, Va, USA, 2002.

[13] Maryland Natural Heritage Program, *Rare, Threatened and Endangered Animals of Maryland*, Maryland Department of Natural Resources, Wildlife and Heritage Service, Annapolis, Md, USA, 2010.

[14] J. L. Dickinson, B. Zuckerberg, and D. N. Bonter, "Citizen science as an ecological research tool: challenges and benefits," *Annual Review of Ecology, Evolution, and Systematics*, vol. 41, pp. 149–172, 2010.

[15] A. M. Dunn and M. A. Weston, "A review of terrestrial bird atlases of the world and their application," *Emu*, vol. 108, no. 1, pp. 42–67, 2008.

[16] M. P. Robertson, G. S. Cumming, and B. F. N. Erasmus, "Getting the most out of atlas data," *Diversity and Distributions*, vol. 16, no. 3, pp. 363–375, 2010.

[17] E. H. Perring and S. M. Walters, *Atlas of the British Flora*, Botanical Society of the British Isles, T. Nelson, London, UK, 1962.

[18] P. T. Harding, "National species distribution surveys," *Monitoring for Conservation and Ecology*, pp. 133–154, 1991.

[19] North American Ornithological Atlas Committee, *Handbook for Atlasing American Breeding Birds*, Vermont Institute of Natural Science, Woodstock, Vt, USA, 1990.

[20] C. S. Robbins and E. A. T. Blom, Eds., *Atlas of the Breeding Birds of Maryland and the District of Columbia*, University of Pittsburgh Press, Pittsburgh, Pa, USA, 1996.

[21] W. G. Ellison, Ed., *Atlas of the Breeding Birds of Maryland and the District of Columbia*, Johns Hopkins University Press, Baltimore, Md, USA, 2nd edition, 2010.

[22] New Jersey Endangered and Nongame Species Program, *New Jersey's Herp Altas Project: Herp Atlas Volunteer Training Manual*, New Jersey Department of Environmental Protection, NJ, USA, 2002.

[23] G. D. Therres, C. A. Davis, and C. W. Swarth, "Grid-based Herp Atlas using active searching: a pilot project," Final Report, Maryland Department of Natural Resources, Annapolis, Md, USA, 2011.

[24] D. H. Foley III and S. A. Smith, *Comparison of two herpetofaunal inventory methods and an evaluation of their use in a volunteer-based statewide reptile and amphibian atlas project*,

Final Report, Maryland Department of Natural Resources, Wye Mills, Md, USA, 1999.

[25] M. L. Hunter Jr., J. Albright, and J. Arbuckle, Eds., *The Amphibians and Reptiles of Maine*, Maine Agricultural Experiment Station Bulletin 838, 1992.

[26] B. I. Crother, J. Boundy, F. T. Burbrink et al., *Scientific and Standard English Names of Amphibians and Reptiles of North America North of Mexico, with Comments Regarding Confidence in Our Understanding*, SSAR Herpetological Circular 37, 6th edition, 2008.

[27] R. Bonney, C. B. Cooper, J. Dickinson et al., "Citizen science: a developing tool for expanding science knowledge and scientific literacy," *BioScience*, vol. 59, no. 11, pp. 977–984, 2009.

[28] R. Gauza and D. Smith, *Maryland Amphibian and Reptile Atlas (MARA) Training Handbook*, Maryland Department of Natural Resources, Natural History Society of Maryland, 2010.

[29] J. B. Jensen, C. D. Camp, W. Gibbons, and M. J. Elliott, Eds., *Amphibians and Reptiles of Georgia*, University of Georgia Press, Athens, Ga, USA, 2008.

Hindawi Publishing Corporation
International Journal of Zoology
Volume 2012, Article ID 710710, 6 pages
doi:10.1155/2012/710710

Research Article

Reviving a Legacy Citizen Science Project to Illuminate Shifts in Bird Phenology

Jessica Zelt,[1] Jason Courter,[2] Ali Arab,[3] Ron Johnson,[2] and Sam Droege[1]

[1] *USGS, Patuxent Wildlife Research Center, BARC-East, Building 308, 10300 Baltimore Avenue, Beltsville, MD 20705, USA*
[2] *College of Agricultural, Forest, and Environmental Sciences, Clemson University, Lehotsky Hall Clemson, SC 29634, USA*
[3] *Department of Mathematics and Statistics, Georgetown University, St. Mary's Hall, 37th and O Streets, NW Washington, DC 20057, USA*

Correspondence should be addressed to Jessica Zelt, jzelt@usgs.gov

Received 2 March 2012; Accepted 26 April 2012

Academic Editor: Richard Stafford

Climate change has been of high interest to both the scientific community and the public at large since the phenomenon was first suggested. Subsequently, and with growing evidence of its impending ramifications, numerous studies have attempted to illuminate climate change impacts on bird migration. Migration is a key event in the annual cycle in the reproductive success of birds, and changes in migration in response to climate may indicate that species populations are at risk. Previous studies report earlier arrival dates in response to climate change in many bird species, although specific mechanisms are often difficult to explain at broad spatial and temporal scales. Using a newly revived dataset of historical migration cards for over 870 species and spanning 90 years throughout North America, we are developing an historical baseline of bird arrival dates to compare with contemporary records. Here we chronicle the history and reemergence of the North American Bird Phenology Program. We present two case studies illustrating how data from this program has been used to model historical arrival dates of Ruby-Throated Hummingbird (*Archilochus colubris*) and Purple Martin (*Progne subis*) throughout eastern North America. Our results show the importance of considering spatial and temporal variability in understanding patterns of bird spring arrivals.

1. Introduction

Climate change is a leading threat to the survival of species and integrity of ecosystems [1], and in response to climate change, bird migration times are changing throughout much of the world [2]. Shifting spring arrivals have been documented in North America, Eurasia, and Australia for a broad range of species and locations [3]. Not only are birds showing advancement in their arrivals to breeding grounds, but also are departing later in the autumn [4–7]. Migratory changes have been correlated to changing temperatures in wintering grounds, breeding grounds, and along migration routes and to a variety of other variables such as rainfall, humidity, and wind speed [1]. Migration must align with favorable environmental conditions to reduce the stress incurred through mistiming arrivals with maximal food abundance and vegetation growth [8, 9]. If not, migratory changes can negatively impact reproduction and survival [10]. Phenology, the study of the timing of seasonal life cycle stages, is one of the simplest ways to track biological changes of species in response to climate change [3]. Because phenological sequences in birds often rely on environmental stimuli to spur arrivals and departures, studies of bird phenology are often used by ecologists to assess the possible impact of climate change on wildlife species [11].

Precise relationships between climate variables and migration times remain difficult to assess because many phenological records are from small geographic areas, recent time periods, and/or contain only small subsets of species. While citizen science efforts have long provided important information for studying long-term range shifts in birds over expansive geographic ranges (e.g., the Christmas Bird Count, the Breeding Bird Survey), a similar large-scale effort that provides data to assess changes in phenology has long been lacking.

In this paper we introduce or, to a few, reintroduce, the North American Bird Phenology Program (BPP), a citizen science program, unsurpassed in geographic, taxonomic, and chronological extent; the BPP houses the most comprehensive legacy dataset in the world for bird migration. This data set contains several million historical arrival and departure records for migratory birds, collected between 1881 and 1970 by leading naturalists and the general public. Although first envisioned to study the distribution and migration of birds, it is now being adapted to investigate shifts in bird arrivals over time, a question with enormous ecological implications. Here we chronicle the history and reemergence of the BPP and present two case studies to demonstrate the potential usefulness of this dataset. Of particular interest to this special edition is the use of citizen volunteers to both collect the original data and now input historical records into digital databases for future analysis.

2. History of the BPP

In 1881, Wells Woodbridge Cooke, a Mississippi teacher interested in the seasonal movements of birds, started a regional cooperative collection of records with acquaintances and colleagues throughout the Mississippi Valley, to document first arrival dates, dates of highest abundance, and last seen dates of all migratory birds [12]. With the support of the American Ornithologists' Union (AOU), Cooke expanded the cooperative which grew to 3,000 participants at the program's height [12]. Participants consisted of both prominent naturalists as well as citizens interested in collecting data for a scientific study. This growth also expanded the geographic range of the volunteer network to include the entire United States, Canada, and a portion of the West Indies. The program was then picked up, in the late 1880s, by the newly formed United States Department of Agriculture (USDA) Biological Survey, where participation peaked, before being passed onto the predecessors of the United States Fish and Wildlife Service (USFWS) and later the United States Geological Survey (USGS) in the Department of the Interior. Cooke relocated to the Washington D.C. area with the survey and became a naturalist for the USDA using the records collected to understand migratory bird patterns, geographic ranges, and basic avian behavior. After Cooke's death in 1916, the program was subsequently guided by other coordinators whom carefully tended and added to the records in the files. Throughout the period of data collection, protocol was instructed and guided by the program coordinator and remained largely consistent though the formatting changed with technological advances. In sum, records on bird migration, breeding, wintering, and distribution for over 870 species were collected over a 90 year span and served as the basis for the AOU's Check-list of North American Birds and sources of information for the first ornithological field guides [12–14]. Although the federal government actively maintained the program, participation gradually declined, and in 1970 the program came to an end as private sector bird watching groups became interested and active in maintaining bird distribution and migration records. This once formidable program was largely forgotten after years of little use recognition and sat idle for over 40 years. In March of 2008, the program was revived in response to nationwide concerns about changes in bird phenology as a result of climate change and the unique opportunity these cards offered to providing documented evidence of possible changes. Funding was limited; however, because the majority of records were handwritten in various formats, they could not simply be converted into a usable digital form using optical character recognition. Records were therefore scanned as image files, later to be displayed online through a data entry interface for transcription by the general public. After months of scanning records in-house, the public website was launched in February of 2009. The program currently relies on a growing worldwide network of over 2,500 volunteers to complete a double-blind transcription of each record, which upon matching is sent into a custom built database. To date, over one million records have been scanned and 500,000 cards transcribed online. Once validated, the records will be accessible online by biologists, managers, and members of the general public.

3. Volunteer Recruitment, Training, and Data Management

Volunteers for the BPP program have been recruited through media outreach, presentations at local and national ornithological meetings, and by word-of-mouth. Volunteers of all abilities and ages are invited to become BPP transcribers, as long as they have Internet access and a web browser [15]. After registering with the program, each volunteer must watch a 15-minute training video that explains how to transcribe different versions of the migration cards. Then, using a web-based application, observation card images are selected from the pool of cards available for transcription by the volunteer. An observation card can be selected at random or filtered for a desired state or species. Volunteers transcribe the contents of the cards by filling in fields of an online form. Each card is entered by at least two independent volunteers for quality assurance. When the two entries for the card match, the data are sent into the BPP database. If the two entries do not match, or if a transcriber marks the card as a "problem card," the card is flagged and reviewed in the BPP office.

Volunteers can track their individual progress as well as the progress of the volunteer community through an expandable window on the transcription screen and through charts viewable on the main website. Volunteers are ranked based on the number of cards they submit and receive recognition for their contribution through email, certificates, prizes, and mention in the BPP monthly newsletter.

4. Case Study 1: Modeling Ruby-Throated Hummingbird Arrival Dates across Eastern North America

Ruby-throated Hummingbirds (*Archilochus colubris*) are charismatic, neotropical migrants that have fascinated naturalists for centuries [15]. They are easily identified and the

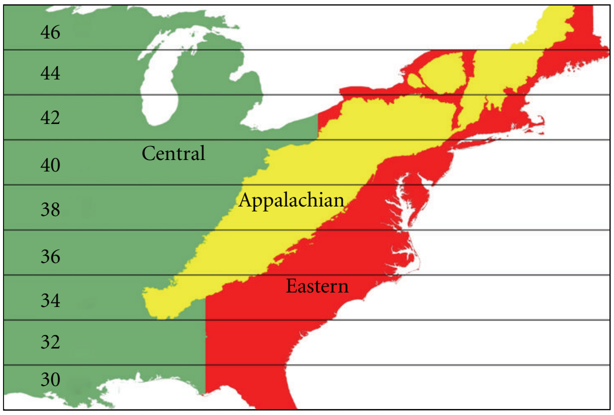


Figure 1: Our study area (29–47°N, 67–95°W) divided into three regions based on classifications used by the Breeding Bird Survey (http://www.pwrc.usgs.gov/bbs/) and the Environmental Protection Agency (Level III Ecoregions; http://www.epa.gov/wed/pages/ecoregions/level_iii_iv.htm).

only regularly reported hummingbird in eastern North America, making them suitable targets for long-term monitoring programs. Hummingbirds winter in Central America between northern Panama and southern Mexico and most migrate across the Gulf of Mexico, arriving at their breeding grounds in North America between February–April [15] where they help pollinate at least 31 plant species [16]. Recent studies indicate that hummingbirds are arriving earlier to their breeding grounds than in historical time periods in some places (e.g., Maine [17]; Massachusetts [18, 19]; South Dakota [20]; and New York [18]), but not others (Minnesota [20]). Proposed mechanisms for such changes include climate change, increasing hummingbird populations, and an increase in popularity of backyard bird-feeding. It remains difficult to explain possible mechanisms for changes, however, when examining migration at site- or region-specific scales.

Recent hummingbird migration patterns are becoming well-documented throughout North America, thanks to emerging networks of citizen science observers reporting first arrivals online through popular websites such as Journey North (http://www.learner.org/jnorth/), hummingbirds.net, and eBird (http://ebird.org/content/ebird/). At the same time, changes in climate [21, 22] and land-use [23, 24] have also been well documented as mapping technology (e.g., United States Geological Survey Land Cover Institute http://landcover.usgs.gov/usgslandcover.php) and historical climate data have become widely accessible (e.g., the Nature Conservancy's Climate Wizard http://www.climatewizard.org/). Both migration and environmental data are needed to understand how meteorological conditions and land-use influence bird migration at broad spatial scales, but to understand bird migration across both space and time, we need to better understand how bird migration occurred historically. Until recently, a continent-wide baseline to compare with recent arrivals has been largely unavailable.

The objective of this study is to demonstrate how data from the recently revitalized North American Bird Phenology Program (BPP) can be used to generate an historical understanding of migration in Ruby-throated Hummingbirds. This understanding could help prompt future studies that assess changes in bird migration at broad spatial and temporal scales in response to climate and land-use changes.

First arrival dates of Ruby-throated Hummingbirds in eastern North America (29–47°N, 67–95°W) were transcribed from arrival cards reported through the North American Bird Phenology Program. Each arrival location was then assigned a location (i.e., latitude, longitude, and altitude) based on the centroid of the reported arrival city using the GPS visualizer geocoding service (http://www.gpsvisualizer.com/). Longitudinally, arrival records east of 95°W (approximate range limit for Ruby-throated Hummingbirds) were divided into Central, Appalachian, and Eastern Regions (Figure 1). Regions were delineated based on categories used by the Breeding Bird Survey and the Environmental Protection Agency (Level III Ecoregions; http://www.epa.gov/wed/pages/ecoregions/level_iii_iv.htm).

In sum, we analyzed 5,065 first arrival records from the BPP card files between 1880 and 1969. We used multiple regression to assess the effects of latitude, longitude, and altitude on hummingbird arrival dates in the eastern United States and ANOVA to assess the differences in mean hummingbird arrival dates by region (Central, Appalachian, Eastern) while including latitude as a covariate. It should be noted that these findings are preliminary and based on statistical methods with strict assumptions about the data (e.g., independence) that are difficult to meet due to potential dependence structures over space and time.

Our model (adjusted $r^2 = 0.66$, $F_{3,5061} = 3267.1$, $P < 0.0001$) indicated that latitude, longitude, and altitude were related to hummingbird arrival dates. In summary, from 1880–1969, hummingbirds arrived 3.4 days later for every 1° increase in latitude, 1.2 days later for every 10° longitude increment, moving from west to east, and 7.5 days later for every 1000 m increase in elevation. Mean arrival dates also differed by region (Figure 1), with birds arriving 1.3 ± 0.33 S.E. days earlier in the central United States than in the Appalachian Region ($P < 0.0001$) and 1.1 ± 0.27 S.E. days earlier in the central United States than in the Eastern Region ($P < 0.0001$). No significant differences in arrival dates were noted between the Appalachian and Eastern Regions ($P = 0.52$; Figure 2).

Earlier arrivals in the central United States (Figure 2) could be explained by a difference in travel distance with central migrants travelling directly north from Gulf States where many hummingbirds make landfall (i.e., Alabama, Mississippi, Louisiana) [15] and eastern migrants travelling northeast, a less direct route. Another possibility is that central migrants may have migrated over land (i.e., Mexico and Texas) [15], perhaps requiring less time to refuel than their eastern counterparts making dangerous and exhausting trips across the Gulf of Mexico. Later arrivals in mountainous regions are not surprising [25] given that melting snowpack and cooler spring temperatures can delay spring phenology at high latitudes [26].

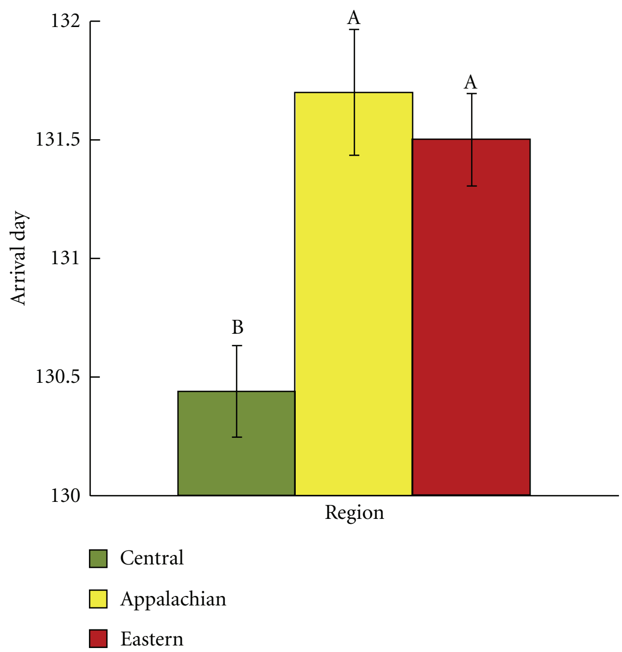


Figure 2: Difference in mean first arrival dates (±S.E.) of Ruby-throated Hummingbirds in the Central ($N = 2002$), Appalachian ($N = 1075$), and Eastern Regions ($N = 1988$; see Figure 1 for region designations). Arrival days are expressed in "day of year" and corrected for leap years; for example, "130" corresponds to May 10. Inset letters represent differences that are significant at the $P < 0.05$ level.

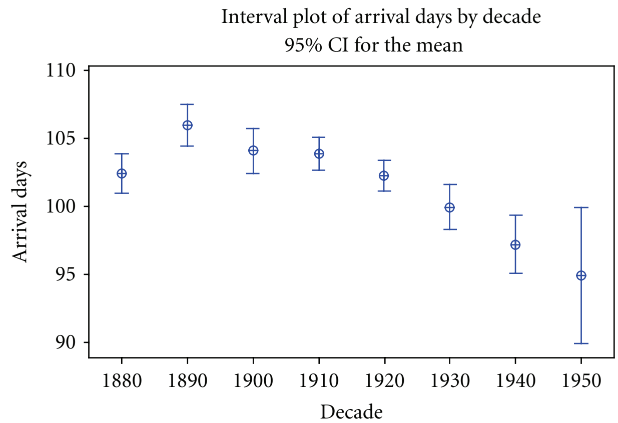


Figure 3: Interval plot for the mean arrival days for decades 1880–1950. Intervals show the 95% confidence interval for the mean decadal arrival dates. Arrival days are expressed in "day of year" and corrected for leap years; for example, "90" corresponds to March 29. The sample sizes for decades are 501 (1880s), 565 (1890s), 700 (1900s), 1126 (1910s), 1152 (1920s), 738 (1930s), 444 (1940s), 119 (1950s).

Our findings highlight the importance of considering spatial variables such as altitude and region in studies of phenology and provide a basis for a number of future research questions. For example, are hummingbirds tracking recent climate change events? Does climate change slow or speed up migration in hummingbirds? Do hummingbirds increase stopover periods in areas such as mountains that are disproportionately affected by warming climates [27]. Where are changes in migration most pronounced and are these changes synchronized across food webs (e.g., birds, insects, plants)? Have these changes impeded or enhanced pollination services hummingbirds provide?

Efforts such as the North American Bird Phenology Program help provide the missing puzzle piece of data for understanding historical migration patterns. As recent migration data emerge along with more sophisticated tools to assess climate and land use changes of the past century, the availability of historical data allows us to better understand changes in bird phenology and related global changes at broad temporal and spatial scales.

5. Case Study 2: Changes in Purple Martin Arrival Dates

Purple Martins (*Progne subis*) are the largest member of the swallow family in North America. They spend their nonbreeding season in Brazil and migrate to North America to nest. Adults commonly return to the same nesting sites where they were successful in previous years. These adult martins arrive to their nesting grounds a few weeks before subadults arrive to establish new nesting sites. Once a pair is established, they cooperate equally in building the nest out of mud, grass, and twigs. This species is of special interest to birders, in large part, because of the close proximity of their nesting sites to human settlements [28].

Based on an initial dataset of 5,345 Purple Martin arrival records provided by the BPP, arrival dates were aggregated over each decade, from the 1880s through the 1950s. We found that the arrival dates for 1920s, 1930s and 1940s were statistically earlier than the arrival dates for the 1880s through 1910s (Figure 3) (a randomized block analysis of variance (ANOVA) of the arrival days using "decade" as the main effect and "latitudinal band" as the blocking effect showed the main effect (decade) to be statistically significant at 0.05 level ($P < 0.0001$)). Reforestation in the northeast during the beginning of the 20th century [29] and increasing use of artificial martin houses [30] may have increased martin populations during this time and may partially explain migratory advancements noted [31]. Increased competition for nest cavities with introduced species, such as European Starlings (*Sturnus vulgaris*) and House Sparrows (*Passer domesticus*) [32], may also be associated with advancing martin migration dates. Finally, it should also be noted that the relatively large standard error for the 1950s data in Figure 3 is due to relatively lower sample size (119 observations) than the previous decades.

We conducted a change point analysis of the mean arrival dates for latitudinal degree groups (total of 15 groups) in the 1880s–1950s. The groups were assigned so each group had data from each decade. The change point analysis was conducted based on the method proposed by Allen and Nice [31] and the functions available through the package MCMCpack in the freely distributed software R (http://www.r-project.org/). Using this method, for all time series and latitudinal groups, the most probable change points selected were the 1890s and 1900s. In most cases, there was a decline in the mean arrival dates after the 1900s.

6. Discussion of Future Analysis Using BPP Data

Previous approaches to predict species' responses to climate change have generally addressed migratory changes at broad spatial or temporal scales, but rarely both, because of the insufficient availability of long-term or spatially extensive data sets [1]. Also, there is increasing interest to understand potential association between spatial, temporal, or spatiotemporal "change-points" or "tipping-points" of environmental processes, such as the bird arrival process, and changes in the earth's climate. Needs for these datasets are urgent and the BPP dataset provides a unique baseline that more recent data on bird phenology can be compared. The original BPP dataset covers 90 years, and by integrating it with contemporary records of first arrival dates, we can evaluate 130 years of spatial, temporal, and interspecific variation in bird phenology across the United States and Canada. This could provide unprecedented insight into the migratory behavior of birds and its relationship to climate and environmental changes.

Future research studies are being planned to investigate changes in arrival dates of species at site-based and national scales while using elevation and climate variables to explain temporal and spatial variation in migration times and strengthen the conclusions presented here. This dataset could facilitate understanding of ecosystem responses to climate change and be used to develop decision support tools for avian and habitat management plans in the face of climate variability and directional change. These data will also contribute to our understanding of patterns and processes related to bird-related diseases, such as avian influenza, that can be affected by changing migration phenology and may also impact human health. In addition, these data could assist modelers in developing species vulnerability assessments—phenology can indicate the likelihood that a population will decline in size in response to climate change [33, 34]. Because phenology is related to population distributions, these data could be used to create tools for ecological forecasting and risk assessment for imperiled species as well as helping researchers identify which species are more adaptable to projected changes and likely to maintain sustainable populations.

Historical data, such as those preserved by the BPP, are of special importance because they are irreplaceable. This particular dataset provides a wealth of information about bird migration, and converting it to a digitally usable format is the first step to making it widely useful and accessible. Collecting and then transcribing original records over 130 years later has only been possible by relying on participants from around the world to assist in this process. Through standardized protocols and the collective effort of the public, the scope of this monitoring program has been magnified beyond any capability the government would have had to collect or transcribe data. As such, we have created a public/private partnership in which a dedicated team of governmental employees and academics have created an appropriate home and structure for the data, and members of the general public have collected and transcribed data; each making important contributions to a project that will ultimately help us better understand large-scale ecological processes that affect us all.

Acknowledgments

The authors are much indebted to the original observers who collected the extensive data used here as well as the thousands of volunteers who have contributed to transcribing the data into our database. The North American Bird Phenology Program relies heavily on the work of volunteers around the globe and we could not accomplish what we have without them. There are many individuals besides the authors of this paper who have provided guidance and support to the BPP, including Chandler Robbins, Jane Fallon, John R. Sauer, Kevin Laurent, Kinard Boone, and Lynda Garrett. The development of the North American Bird Phenology Program has been supported with grants from the National Oceanic and Atmospheric Administration, United States Fish and Wildlife Service, the United States Geological Survey, and the Maryland Ornithological Society. They also thank Clemson University, with additional support from a Carolina Bird Club grant, for funding the work presented in the hummingbird case study.

References

[1] P. E. Hulme, "Adapting to climate change: is there scope for ecological management in the face of a global threat?" *Journal of Applied Ecology*, vol. 42, no. 5, pp. 784–794, 2005.

[2] E. Lehikoinen, T. H. Sparks, and M. Zalakevicius, "Arrival and Departure Dates," *Advances in Ecological Research*, vol. 35, pp. 1–31, 2004.

[3] Intergovernmental Panel on Climate Change (IPCC), "Impacts, adaptation and vulnerability," in *Contribution of Working Group II to the Fourth Assessment Report of the IPCC*, M. L. Parry, O. F. Canziani, J. P. Palutikof, P. J. van der Linden, and C. E. Hanso, Eds., pp. 98–101, Cambridge University Press, Cambridge, UK, 2007.

[4] L. Jenni and M. Kéry, "Timing of autumn bird migration under climate change: advances in long-distance migrants, delays in short-distance migrants," *Proceedings of the Royal Society B*, vol. 270, no. 1523, pp. 1467–1471, 2003.

[5] C. Parmesan and G. Yohe, "A globally coherent fingerprint of climate change impacts across natural systems," *Nature*, vol. 421, no. 6918, pp. 37–42, 2003.

[6] T. L. Root, J. T. Price, K. R. Hall, S. H. Schneider, C. Rosenzweig, and J. A. Pounds, "Fingerprints of global warming on wild animals and plants," *Nature*, vol. 421, no. 6918, pp. 57–60, 2003.

[7] A. Menzel, T. H. Sparks, N. Estrella et al., "European phenological response to climate change matches the warming pattern," *Global Change Biology*, vol. 12, no. 10, pp. 1969–1976, 2006.

[8] S. Pau, E. M. Wolkovich, B. I. Cook et al., "Predicting phenology by integrating ecology, evolution and climate science," *Global Change Biology*, vol. 17, pp. 3633–3643, 2011.

[9] R. L. DeLeon, E. E. DeLeon, and G. R. Rising, "Influence of climate change on Avian Migrants' first arrival dates," *Condor*, vol. 113, no. 4, pp. 915–923, 2011.

[10] M. C. Forchhammer, E. Post, and N. C. Stenseth, "North Atlantic Oscillation timing of long- and short-distance migration," *Journal of Animal Ecology*, vol. 71, no. 6, pp. 1002–1014, 2002.

[11] S. B. Smith and P. W. C. Paton, "Long-term shifts in autumn migration by songbirds at a coastal eastern north American stopover site," *Wilson Journal of Ornithology*, vol. 123, no. 3, pp. 557–566, 2011.

[12] T. S. Palmer, "In Memoriam: wells woodbridge cooke," *The Auk*, vol. 34, no. 2, pp. 119–132, 1917.

[13] D. Sam, "Interview with Chandler S. Robbins," in *The North American Bird Phenology Program Website*, USGS Patuxent Wildlife Research Center, Laurel, MD, USA, 2009.

[14] J. A. Allen, "The A.O.U. check-list of North American birds, third edition," *The Auk*, vol. 27, no. 4, pp. 466–472, 1910.

[15] United States Geological Survey, "North American Bird Phenology Program," 2008, http://www.pwrc.usgs.gov/bpp/.

[16] T. R. Robinson, R. R. Sargent, and M. B. Sargent, "Ruby-throated Hummingbird (*Archilochus colubris*)," in *The Birds of North America Online*, A. Poole, Ed., Cornell Lab of Ornithology, Ithaca, NY, USA, 1996.

[17] D. Austin, "Bird flowers in the eastern United States," *Florida Scientist*, vol. 38, pp. 1–12, 1975.

[18] W. H. Wilson Jr., D. Kipervaser, and S. A. Lilley, "Spring arrival dates of Maine migratory breeding birds: 1994–1997 vs. 1899–1911," *Northeastern Naturalist*, vol. 7, pp. 1–6, 2000.

[19] C. J. Butler, "The disproportionate effect of global warming on the arrival dates of short-distance migratory birds in North America," *Ibis*, vol. 145, no. 3, pp. 484–495, 2003.

[20] A. Ledneva, A. J. Miller-Rushing, R. B. Primack, and C. Imbres, "Climate change as reflected in a naturalist's diary, Middleborough, Massachusetts," *Wilson Bulletin*, vol. 116, no. 3, pp. 224–231, 2004.

[21] D. L. Swanson and J. S. Palmer, "Spring migration phenology of birds in the Northern Prairie region is correlated with local climate change," *Journal of Field Ornithology*, vol. 80, no. 4, pp. 351–363, 2009.

[22] T. R. Karl and K. E. Trenberth, "Modern Global Climate Change," *Science*, vol. 302, no. 5651, pp. 1719–1723, 2003.

[23] S. R. Loarie, P. B. Duffy, H. Hamilton, G. P. Asner, C. B. Field, and D. D. Ackerly, "The velocity of climate change," *Nature*, vol. 462, no. 7276, pp. 1052–1055, 2009.

[24] J. A. Foley, R. DeFries, G. P. Asner et al., "Global consequences of land use," *Science*, vol. 309, no. 5734, pp. 570–574, 2005.

[25] E. C. Ellis, K. K. Goldewijk, S. Siebert, D. Lightman, and N. Ramankutty, "Anthropogenic transformation of the biomes, 1700 to 2000," *Global Ecology and Biogeography*, vol. 19, no. 5, pp. 589–606, 2010.

[26] T. H. Sparks and O. Braslavská, "The effects of temperature, altitude and latitude on the arrival and departure dates of the swallow *Hirundo rustica* in the Slovak Republic," *International Journal of Biometeorology*, vol. 45, no. 4, pp. 212–216, 2001.

[27] D. W. Inouye, B. Barr, K. B. Armitage, and B. D. Inouye, "Climate change is affecting altitudinal migrants and hibernating species," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 4, pp. 1630–1633, 2000.

[28] J. C. Fyfe and G. M. Flato, "Enhanced climate change and its detection over the rocky mountains," *Journal of Climate*, vol. 12, no. 1, pp. 230–243, 1999.

[29] C. R. Brown, "Purple Martin (*Progne subis*)," in *The Birds of North America Online*, A. Poole, Ed., Cornell Lab of Ornithology, Ithaca, NY, USA, 1997.

[30] J. A. Litvaitis, "Response of early successional vertebrates to historic changes in land use," *Conservation Biology*, vol. 7, no. 4, pp. 866–873, 1993.

[31] R. W. Allen and M. M. Nice, "A study of the breeding biology of the Purple Martin (*Progne subis*)," *American Midland Naturalist*, vol. 47, pp. 606–665, 1952.

[32] A. J. Miller-Rushing, T. L. Lloyd-Evans, R. B. Primack, and P. Satzinger, "Bird migration times, climate change, and changing population sizes," *Global Change Biology*, vol. 14, no. 9, pp. 1959–1972, 2008.

[33] S. Chib, "Estimation and comparison of multiple change-point models," *Journal of Econometrics*, vol. 86, no. 2, pp. 221–241, 1998.

[34] A. P. Møller, D. Rubolini, and E. Lehikoinen, "Populations of migratory bird species that did not show a phenological response to climate change are declining," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 42, pp. 16195–16200, 2008.

Hindawi Publishing Corporation
International Journal of Zoology
Volume 2012, Article ID 428752, 5 pages
doi:10.1155/2012/428752

Research Article

An Evaluation of Ad Hoc Presence-Only Data in Explaining Patterns of Distribution: Cetacean Sightings from Whale-Watching Vessels

Louisa K. Higby,[1,2] Richard Stafford,[3] and Chiara G. Bertulli[2,4]

[1] *School of Ocean Sciences, Bangor University, Menai Bridge, Anglesey LL59 5AB, UK*
[2] *School of Engineering and Natural Sciences, Faculty of Life and Enviromental Sciences, Elding Whale-Watching, Ægisgata 7, 101 Reykjavik, Iceland*
[3] *Division of Science, Institute of Biomedical and Environmental Science and Technology, University of Bedfordshire, Luton LU1 3JU, UK*
[4] *School of Engineering and Natural Sciences, Faculty of Life and Enviromental Sciences, University of Iceland, Sturlugata 7, 101 Reykjavik, Iceland*

Correspondence should be addressed to Louisa K. Higby, louisahigby@gmail.com

Received 1 March 2012; Accepted 10 April 2012

Academic Editor: Anne Goodenough

The analysis of presence-only data is a problem in determining species distributions and accurately determining population sizes. The collection of such data is common from unequal or nonrandomised effort surveys, such as those surveys conducted by citizen scientists. However, causative regression-based methods have been less well examined using presence-only data. In this study, we examine a range of predictive factors which might influence Cetacean sightings (specifically minke whale sightings) from whale-watching vessels in Faxaflói Bay in Iceland. In this case, environmental variables were collected regularly regardless of whether sightings were recorded. Including absences as well as presence in the analysis resulted in a multiple-generalised linear regression model with significantly more explanatory power than when data were presence only. However, by including extra information on the sightings of the whales, in this case, their observed behaviour when the sighting occurred resulted in a significantly improved model over the presence-only data model. While there are limitations of conducting nonrandomised surveys for the use of predictive models such as regression, presence-only data should not be considered as worthless, and the scope of collection of these data by citizen scientists using modern technology should not be underestimated.

1. Introduction

Presence-only data, data where presence of a species or individual is recorded, but where absences are not, are frequent in many *ad hoc* scientific surveys, such as those datasets collected by volunteers or citizen scientists [1–3]. While presence-only data have been shown to produce good maps of species ranges in some occasions (e.g., [1]), using such data to infer changes in distribution, population sizes, and other ecological parameters can be difficult [2]. Reasons for this largely relate to unequal sampling effort [1–4]. Low- or zero-sampling effort could easily miss the presence of a low-density species in certain areas, but the amount of effort applied is largely unknown, and hence lack of presence of a species could relate to a real absence, or simply a lack of effort. Misidentification of species, or misreporting of locations, can confound such studies, although such issues can also occur in any volunteer programme, regardless of sampling strategy employed [1, 2].

Volunteers, and citizen scientists recruited through "crowd-sourcing" events, however, are a cheap method of collecting data over a wide spatial or temporal scale [1, 5]. As such, presence-only data are becoming common, and an evaluation of their use in scientific research is timely. While this study does not strictly use citizen science data, the sampling regime used is, by necessity, not of equal effort in space or time, and some aspects of the dataset collected are, again by necessity, presence only. However, since absence data were also recorded, the dataset provides an ideal opportunity to test the use of presence-only data in causative regression models.

Multiple linear regression and associated linear model reduction methods are a common tool in addressing habitat or environmental variables related to habitat selection by organisms [6–8]. For example, in a stepwise model reduction approach, a large number of explanatory factors can be used to predict presence (or number) of individuals in each location, and through examination of each factor's relative importance, the most important factors involved in explaining the majority of the variation in the dependent variable can be found [9].

Such a multiple regression approach to predicting sightings of minke whales (*Balaenoptera acutorostrata*) was used in the current study. Data on sightings of these species were collected from whale-watching vessels, and a range of environmental factors, such as weather conditions, sea state, and temperature, were also recorded or obtained from existing public data sources. These factors were used in conjunction with sightings of Cetaceans and observations on their behaviour to identify the role of environmental factors in the occurrence of whale sightings. Since data were collected regularly throughout the cruise, many factors were known when sightings did not occur, and removing these "absence" data points in subsequent analyses allowed for an evaluation of presence-only data in causative regression models.

2. Methods

Boat-based surveys were carried out from Faxaflói Bay (64.20871° N, 22.19869° W), on the south west coast of Iceland during whale-watching trips. Each survey session was broken down into 15 min observation intervals where environmental variables were recorded (Table 1). In addition, minke whale (*Balaenoptera acutorostrata*) encounters were recorded, along with the numbers seen at each sighting, and this figure was used as a dependent variable in the analysis (with zero indicating absence).

A total of 634 data points were used in the analysis, of which 133 recorded sightings of minke whales (15-minute periods when 1 or more minke whales were observed). These surveys were carried out over 104 days, between April and July 2010 (days when the sea state exceeded Beaufort scale 3 were discarded from the analysis—following prior recommendations, e.g., [10, 11]).

A multiple linear model was constructed with all environmental factors listed in Table 1 as possible explanatory factors for the number of minke whales seen in any 15-minute time period (with the exception of behaviour—which could not be included in models where absence data was used). A stepwise model reduction process was then undertaken (using forward and backward processes) as described in [9], using AIC as a model reduction method. Next, any 15-minute period where no minke whales were seen was removed to create a presence-only dataset, and the stepwise model process ran again. Next, the presence-only dataset was analysed again using all the previous environmental factors along with behaviour of the cetacean when it was seen.

Given that data are count data and residuals from standard linear models were not normally distributed, generalised linear models were used for analysis (as per [12, 13]). The full dataset showed variance far exceeded the mean for the counts of both minke whales and white-beaked dolphins, and GLMs based on the negative binomial distribution were used [13]. When zero counts were excluded, data were still not normally distributed, but followed the assumptions of Poisson distributions (mean∼variance), and these GLMs were based on this distribution [13]. Given GLM does not provide a goodness-of-fit statistic (such as R^2), we use (1) an evaluation of the sum of the residuals as a measure of goodness of fit to compare presence and absence data models, where lower values indicate better fit, (2) manually calculated R^2 as a model comparison mechanism, but not as a true measure of the predictive power of the model, by subtracting the quotient of the residual and model deviance from one, and (3) ran the analysis using traditional linear models, despite limitations applied to count data, to provide a more comparative study of the proportion of variability explained. In this case, all data for dependent variables (minke whale sighting number) were $\log_{10} + 1$ transformed, since this normalised the residuals of the presence-only dataset.

3. Results

Analysis of the full dataset, excluding behaviour but including absences of minke whales, produced a reduced model with five significant ($P < 0.05$) explanatory factors (Table 2), a mean squared residual value of 0.528, and an estimated $R^2 = 0.394$. In comparison, traditional linear modelling (despite nonnormal residuals) resulted in a highly significant final model ($F_{9,\,624} = 16.22$; $P < 0.001$) and an adjusted $R^2 = 0.178$.

The reduced dataset—only using data when minke whales were present—produced a model with considerably the worst explanatory power (although significance testing was not possible due to the different sizes of datasets), with two significant explanatory factors (Table 2), with a mean-squared residual value of 0.907 and an estimated $R^2 = 0.140$. Similar decreases in fit were obtained by traditional linear models ($F_{5,\,127} = 2.91$; $P = 0.005$; adjusted $R^2 = 0.086$).

Inclusion of the behaviour of the minke whale as an observation using the presence-only dataset gave an improved fit regression with two significant explanatory factors (Table 2), and mean squared residual value of 0.804 and an R^2 of 0.240. In this case, feeding behaviour occurred during significantly more sightings than the other behaviours. Again, similar trends were found with traditional linear models ($F_{7,\,125} = 4.67$; $P < 0.001$; adjusted $R^2 = 0.163$); this model was significantly better than the presence-only model not including behaviour (ANOVA test on two fitted models, $P = 0.0015$).

From an analysis of explanatory variables in the reduced models (Table 2), it can be seen that cetacean sightings were affected by similar factors in most models—sea and

Table 1: Explanatory factors initially included in the linear model. Use of continuous and discrete (category) variables follows suggestions of [14], where category variables are categorised in as few as possible meaningful categories, and ordinal variables are assumed to be continuous. Logistic variables are categorised as discrete with $n = 2$ levels.

Factor	Variable type	Notes
Date	Continuous	From day 1 to day 158
Boat	Category ($n = 2$)	Differences in height of sighting platform; possible differences in acoustics of boat engines
Time of day	Continuous	That is, 4:30 pm = 16.5
Behaviour of cetacean	Category ($n = 3$)	Surfacing, feeding, and other
Number of humpback whales	Continuous	
Number of killer whale	Continuous	
Number of white-beaked dolphins	Continuous	
Sea state	Continuous	
Percentage of cloud cover	Continuous	
Weather conditions	Category ($n = 3$)	Sun, Cloud, and rain
Wind direction	Category ($n = 5$)	N, E, S, W, and no wind
Tidal conditions	Category ($n = 2$)	Flood or ebb
Swell height	Continuous	
Visibility	Continuous	
Sea surface temperature	Continuous	
Observer	Category ($n = 3$)	Three different observers recorded results

weather conditions were important. Sea temperature was also important in some cases, and the behaviour being performed by the cetacean was important, when included, for explaining the observed sightings of minke whales.

4. Discussion

In general, we demonstrate that presence-only data limit the explanatory properties of models such as multiple linear regression. However, the inclusion of explanatory factors, which can only be included in presence-only models (i.e., they relate to the actual sightings), can increase the power of such modelling approaches on presence-only data.

One important consideration is that datasets containing multiple zero counts require more complex models than those which do not [12, 13]. As such, presence-only data, by its nature, excludes all zero counts, and means models built using standard linear model techniques can be used for analysis. Such models give a much better and intuitive understanding of model fit using the familiar R^2 variable. While in this study, presence-only data were analysed using generalised linear models for comparison with absence data, a simple log transform of the dependent variable normalised the residuals of the linear model approach.

Given that the current dataset was not collected by citizen scientists, we need to consider how the results might apply to citizen-science-collected data, and whether the technique is valuable if applied to the collection of data using citizen science methods. Firstly, what is clearly important is the amount of data available. A successful citizen science programme could greatly increase the number of records and may result in high-quality predictive models being built on presence-only data. Furthermore, it should also be noted that if larger numbers of participants are taking part in such surveys, the chance of missing an actual sighting of a cetacean on a whale-watching trip will be reduced. As such, if the number of "returns" or submission of data is high for each whale-watching trip, then it could be considered that most actual sightings will have been recorded, hence it is more likely that where no data are present, there were no Cetaceans, rather than this being a false absence.

Explanatory models frequently showed factors such as cloud cover, visibility, and sea state to be important. While these factors could influence actual distribution of Cetaceans, it is more likely that they influence the observer's ability to detect them [10, 15]. While such issues may cause some concern for conversion to citizen-science-collected data, the extra volumes of data collected may be able to be standardised for "detection" conditions, by subsetting the data prior to analysis (i.e., into rough or calm conditions, or into sunny versus overcast conditions).

When behaviour was included in the predictive model, there was a significant increase in its explanatory power. In this study, behaviour helped to explain the number of minke whales present at a particular sighting, with more minke whales present when feeding was occurring than for the other behaviours, likely as minke whales may be more likely to be or remain in the presence of a boat when there is significant food available in the area. However, the use of such an approach is only possible when using presence data, since behaviour cannot be recorded if no individual is seen. However, recent research suggests that untrained volunteers are not good at recording behaviour accurately [16]. Despite this, given that data submission could be via photograph or video, this recognition of behaviour could be verified by researchers (e.g., [1]), as could other accuracy-of-data issues,

Table 2: Factors present in the stepwise-reduced model for each dataset where significant multiple linear regressions were obtained. P indicates that the factor was present in the reduced model, and + or − indicates whether there was a positive or negative relationship of the factor compared to the number of Cetaceans. Where no + or − is present, the factor was a category variable. N/A indicates that this variable was not used as an explanatory factor in the analysis.

Factor	Full dataSet	Presence only	Presence only and behaviour
Date	P+		
Boat	P		
Time of day			
Behaviour of cetacean	N/A	N/A	P
Presence of humpback whales			
Presence of killer whale			
Presence of white-beaked dolphin			
Sea state		P−	
Percentage of cloud cover	P+		
Weather conditions			
Wind direction			
Tidal conditions			
Swell height			
Visibility	P+	P+	
Sea surface temperature	P−		P−
Observer			

such as location and time of the record, which can all be verified using modern digital technology [1, 5].

Results relating to observation of surface sightings of minke whales are clearly dependent on the whales being at the surface. Therefore, factors which influence the dive time of the whales will also be apparent in terms of the surface sightings. For example, surface intervals of minke whales are known to vary throughout the year and throughout the day [17], and this seems to be correlated to the type of food the whales are foraging on and the local bathymetry of the foraging site [18]. In this study, both date and sea surface temperature had opposite effects (with a positive relationship between sightings and date, and a negative relationship with sea surface temperature). Such a finding is consistent with previous findings, indicating that surface intervals may be longer in the spring, since whales may feed on plankton blooms, but may also associate themselves with areas of upwelling (cooler water) where these blooms, or other food, may be more abundant [19].

Overall, presence-only data do provide some useful, biological information regarding the sightings of Cetaceans, and the ability to collect a greater volume of data should offset concerns over its value. However, whether data are collected by citizen scientists or trained scientists, there are limitations of the use of tourist vessels in this type of explanatory variable approach. A first step in many analyses of data using model-reduction approaches is the determination of whether sightings or distributions of a species can be explained by chance, by testing the distribution data against a Poisson distribution [20]. Given that tourist vessels automatically head to previous sightings and communicate with each other as to the location of recently seen whales, any approximation to a regular grid, or random sampling, cannot be assumed, and the fact that distribution of whale sightings could be entirely random cannot be ruled out. Furthermore, although for minke whales, much published research has been conducted from whale-watching vessels (e.g., [18, 19]), the behaviour of other cetacean species to whale-watching boats can be variable, with some species, such as humpback whales, avoiding vessels by increasing dive time [21] and others actively approaching boats [22]. In particular, whale-watching vessels appear to record more "active" behaviour, such a leaping out of the water, and also from younger individuals [23]. However, the use of presence-only data, with recordings and understanding of the implications of the behaviour (avoiding, approaching, jumping, etc.), may still allow useful data on factors such as habitat selection to be collected from whale-watching vessels.

As considered elsewhere, the use of digital technology and internet storage facilities can both increase uptake and accuracy of citizen science work [1, 5], and while collection of presence-only data has some disadvantages, some limitations cannot be improved by the use of trained personnel and require dedicated random surveys. In some cases, presence-only data even have some advantages and are worthy of consideration given the current resource cuts to science budgets and the need to greater engage the public with scientific research.

Acknowledgments

The authors would like to thank the Elding Whale Watching Company, with special thanks to G. Vignir Sigursveinsson and Rannveig Grétarsdóttir without whose support, providing a platform for all survey activities, this research would not have been possible. They are grateful to CSI for funding the data collection in the year 2010, and also to the Faxaflói Cetacean Research volunteer Mirjam Held, who helped with data collection during the 2010 field season in Faxaflói Bay. They would also like to thank the anonymous reviewer for helpful suggestions to improve the paper.

References

[1] R. Stafford, A. G. Hart, L. Collins et al., "Eu-social science: the role of internet social networks in the collection of bee biodiversity data," *PloS one*, vol. 5, no. 12, p. e14381, 2010.

[2] J. Franklin, *Mapping Species Distributions: Spatial Inference and Prediction*, Cambridge University Press, Cambridge, UK, 2009.

[3] M. W. Tingley and S. R. Beissinger, "Detecting range shifts from historical species occurrences: new perspectives on old data," *Trends in Ecology and Evolution*, vol. 24, no. 11, pp. 625–633, 2009.

[4] C. Hassall and D. J. Thompson, "Accounting for recorder effort in the detection of range shifts from historical data," *Methods in Ecology and Evolution*, vol. 1, no. 4, pp. 343–350, 2010.

[5] C. L. Catlin-Groves, "Submitted to this special issue. The citizen science landscape: from volunteers to citizen sensors and beyond," *International Journal of Zoological Research*. In press.

[6] A. E. Goodenough, A. G. Hart, and R. Stafford, "Regression with empirical variable selection: description of a new method and application to ecological datasets," *PLoS One*, vol. 7, no. 3, Article ID e34338, 2012.

[7] M. J. Whittingham, P. A. Stephens, R. B. Bradbury, and R. P. Freckleton, "Why do we still use stepwise modelling in ecology and behaviour?" *Journal of Animal Ecology*, vol. 75, no. 5, pp. 1182–1189, 2006.

[8] J. Fan and J. Lv, "A selective overview of variable selection in high dimensional feature space," *Statistica Sinica*, vol. 20, no. 1, pp. 101–148, 2010.

[9] M. J. Crawley, *Statistics: An Introduction Using R*, Wiley, Chichester, UK, 2005.

[10] P. G. H. Evans and P. S. Hammond, "Monitoring cetaceans in European waters," *Mammal Review*, vol. 34, no. 1-2, pp. 131–156, 2004.

[11] D. Palka, "Effects of Beaufort sea state on the sightability of harbor porpoises in the Gulf of Maine," *Forty-Sixth Report of the International Whaling Commission*, pp. 575–582, 1996.

[12] M. Ridout and C. Demetrio, "Generalized linear models for positive count data," *Revista de Matematica e Estatstica*, vol. 10, pp. 139–148, 1992.

[13] M. J. Crawley, *The R Book*, Wiley, Chichester, UK, 2007.

[14] D. J. Pasta, "Learning when to be discrete: continuous vs. categorical predictors," Paper 248, 2009, SAS Global Forum 2009.

[15] S. Dawson, P. Wade, E. Slooten, and J. Barlow, "Design and field methods for sighting surveys of cetaceans in coastal and riverine habitats," *Mammal Review*, vol. 38, no. 1, pp. 19–49, 2008.

[16] R. L. Williams, S. Porter, A. G. Hart, and A. E. Goodenough, "Submitted to this special issue. The accuracy of behavioural data collected by visitors in a zoo environment Can visitors collect meaningful data?" *International Journal of Zoological Research*. In press.

[17] K. A. Stockin, R. S. Fairbairns, E. C. M. Parsons, and D. W. Sims, "Effects of diel and seasonal cycles on the dive duration of the minke whale (Balaenoptera acutorostrata)," *Journal of the Marine Biological Association of the United Kingdom*, vol. 81, no. 1, pp. 189–190, 2001.

[18] K. Macleod, R. Fairbairns, A. Gill et al., "Seasonal distribution of minke whales Balaenoptera acutorostrata in relation to physiography and prey off the Isle of Mull, Scotland," *Marine Ecology Progress Series*, vol. 277, pp. 263–274, 2004.

[19] P. C. Gill, M. G. Morrice, P. Brad, P. Rebecca, A. H. Levings, and C. Michael, "Blue whale habitat selection and within-season distribution in a regional upwelling system off southern Australia," *Marine Ecology Progress Series*, vol. 421, pp. 243–263, 2011.

[20] A. E. Goodenough, S. L. Elliot, and A. G. Hart, "Are nest sites actively chosen? Testing a common assumption for three non-resource limited birds," *Acta Oecologica*, vol. 35, no. 5, pp. 598–602, 2009.

[21] A. Schaffar, B. Madon, V. Garrigue, and R. Constantine, "Avoidance of whale watching boats by humpback whales in their main breeding ground in New Caledonia," Paper SC/61/WW/6, International Whaling Commission, Cambridge, UK, 2009.

[22] F. Ritter, *Interactions of Cetaceans with Whale Watching Boats—Implications for the Management of Whale Watching Tourism*, M.E.E.R.e.V., Berlin, Germany, 2003.

[23] M. Weinrich, "Are behavioral data from whalewatch boats biased?" Paper SC/61/WW/3, International Whaling Commission, Cambridge, UK, 2009.

Hindawi Publishing Corporation
International Journal of Zoology
Volume 2012, Article ID 149026, 9 pages
doi:10.1155/2012/149026

Research Article

Identifying Large- and Small-Scale Habitat Characteristics of Monarch Butterfly Migratory Roost Sites with Citizen Science Observations

Andrew K. Davis,[1] Nathan P. Nibbelink,[2] and Elizabeth Howard[3]

[1] *Odum School of Ecology, The University of Georgia, Athens, GA 30602, USA*
[2] *D.B. Warnell School of Forestry and Natural Resources, The University of Georgia, Athens, GA 30602, USA*
[3] *Journey North, 1321 Bragg Hill Road, Norwich, VT 05055, USA*

Correspondence should be addressed to Andrew K. Davis, akdavis@uga.edu

Received 29 January 2012; Revised 15 March 2012; Accepted 20 March 2012

Academic Editor: Anne Goodenough

Monarch butterflies (*Danaus plexippus*) in eastern North America must make frequent stops to rest and refuel during their annual migration. During these stopovers, monarchs form communal roosts, which are often observed by laypersons. Journey North is a citizen science program that compiles roost observations, and we examined these data in an attempt to identify habitat characteristics of roosts. From each observation we extracted information on the type of vegetation used, and we used GIS and a national landcover data set to determine land cover characteristics within a 10 km radius of the roost. Ninety-seven percent of roosts were reported on trees; most were in pines and conifers, maples, oaks, pecans and willows. Conifers and maples were used most often in northern flyway regions, while pecans and oaks were more-frequently used in southern regions. No one landcover type was directly associated with roost sites, although there was more open water near roost sites than around random sites. Roosts in southern Texas were associated primarily with grasslands, but this was not the case elsewhere. Considering the large variety of tree types used and the diversity of landcover types around roost sites, monarchs appear highly-adaptable in terms of roost site selection.

1. Introduction

Research on one of the world's most famous insects, the monarch butterfly (*Danaus plexippus*, Figure 1), has benefitted greatly from numerous citizen science programs in North America devoted to tracking this species at various life stages. The attention given to this insect no doubt stems from its large size, easily identifiable orange and black colors (Figure 1), and its well-known and spectacular migrations, which are unique among butterflies. All of these factors make this butterfly extremely charismatic, and this helps to promote public participation in various citizen science programs. For example, the larval stages of this insect are monitored each summer by volunteers of the Monarch Larval Monitoring Project (http://www.mlmp.org/), and these data have been used to document geographic and temporal variation in population recruitment [1, 2]. In the western North American population, volunteers count numbers of adult monarchs that overwinter in clusters along the California coast (Western Monarch Thanksgiving Count), and a recent analysis of these data showed the importance of climatic conditions at the natal sites for predicting overwintering numbers [3]. There is another citizen science program whereby volunteers submit samples of a monarch-specific protozoan parasite (MonarchHealth; http://www.monarchparasites.org/), which has led to the identification of trends in disease prevalence during the summer and fall [4]. Finally, numerous scientific investigations have made use of data from a citizen science program called Journey North (http://www.learner.org/jnorth/), which asks volunteers in North America to report sightings of adult monarchs during the winter [5], during the spring migration [6–8], and during the fall migration when monarchs from

Figure 1: Photograph of an adult monarch butterfly (*Danaus plexippus*), nectaring on milkweed (*Asclepias* sp.). Photo taken by Pat Davis in New York City, NY.

the eastern population are travelling to their Mexican overwintering site [9]. The primary fall sightings are of nocturnal roosts, which monarchs form during their southward migration [10], and that are easily recognized by laypersons, since they often consist of hundreds or thousands of monarchs (Figure 2).

Monarch roosts can be considered stopover sites, which are essentially places where migratory animals pause during their journey to rest and/or refuel. Like most migratory organisms, monarchs utilize stopover sites to feed and deposit fat reserves [11] and to rest at night. Moreover for monarchs, depositing fat reserves during the migration not only provides fuel for the flight, but is essential to their overwintering survival [12]. As such, determining where stopover sites are for monarchs is an important issue in conserving their migration [9]. Further, while there is a wealth of research into stopover ecology of migrating birds (e.g., [13–18]), there are comparatively few studies examining the nature of stopover behavior in monarchs [19–21]. Moreover, there are no published studies where monarch stopover habitat is documented, other than anecdotal observations of roost trees [10]. In fact, it is not known even if monarchs select specific large- or small-scale habitat features at all when they stop or if roost site selection is completely random. Prior examination of roost observations indicated that few locations are utilized by monarchs for roosting year after year [9], which argues for the latter scenario, although more thorough investigation on this idea is warranted. Furthermore, like most migratory animals, monarchs must face continually changing landscapes throughout the entire flyway, including prairies and farmland in the American Midwest, deciduous forests in the eastern seaboard, and dry scrublands in Texas and northern Mexico. Given these changing landscapes they encounter, how then would their stopover habitat preferences (if there are any) change as they progress southward?

The Journey North roost observation database is uniquely positioned to offer insights into this question. When volunteers observe a migratory roost, they not only report the location and date, but are also encouraged to submit general observations on the roost, such as the type of tree or vegetation in which the roost was observed. In this study, we screened four years of Journey North's migratory roost sightings (from eastern North America only) and, from these records, we recorded the type of tree (or other vegetation) in which the roost was observed. We also examined the landscape-level features of the roost site using a GIS approach; here we compared the land use surrounding each roost location to those of randomly selected locations at similar stages of the migration, which we arbitrarily divided into five flyway regions. Our goals for this study were to (1) document the large- and small-scale habitat preferences of monarchs at roosting sites and (2) determine if monarchs display a uniform preference for specific stopover habitats throughout the migration flyway or does their preference change as the migration progresses. Results from this study will not only further scientific understanding of monarch butterfly migration, but should also be relevant to the science of animal migration in general. In fact, to our knowledge this study is the first to examine how stopover habitat preferences of a single migratory animal vary throughout an entire migration flyway.

2. Methods

2.1. Roost Observations. We examined roost observations from the Journey North program between 2005 and 2008 (Figure 3), which are accessible online in the archives section of the program (http://www.learner.org/jnorth/maps/archives.html). For the purposes of this study, we focused on the primary flyway only (the central flyway) and did not consider observations from the Atlantic flyway [9], since very few tagged monarchs from that region are ever recovered in Mexico [20, 22, 23]. Each roost observation in the database is associated with a date (of the first night of observation), and latitude and longitude (of the zip code of the observer's mailing address, see below). While all roost observations have at least these components, observers are also encouraged to record notes about the roost and even take pictures, which are also archived with the sightings. For this study we screened these written notes and recorded what the monarchs were reported roosting on (i.e., tree, shrub, etc.). Moreover, since the aim of this study was to compare roost characteristics along the migratory flyway, we arbitrarily created 5 "flyway regions" of 4° latitude blocks that encompassed the majority of the flyway and roost observations in Canada and the United States (Figure 3). We then categorized the roost observation data (type of vegetation, etc.) into these regions based on the latitude of the observation.

2.2. Landscape Features of Roost Sites. The latitude and longitude associated with roost observations were imported into ArcGIS for analyses of land use surrounding roost sites. We point out that the coordinates of roosts in the Journey North database are not necessarily for the roost tree itself; when new participants sign up, they are asked to report their home address, and from this information, coordinates

(a) (b) (c) (d)

Figure 2: Photographs of monarch migration roosts on various tree types submitted by Journey North citizen scientists. Photograph credits: (a) Iris Tower, Youngstown, NY; (c) Emily McCormick, Mount Cory, OH; (b) Ron & Bobbie Streible, Fort Morgan, AL; (d) Bruce Morrison, Hartley, IA.

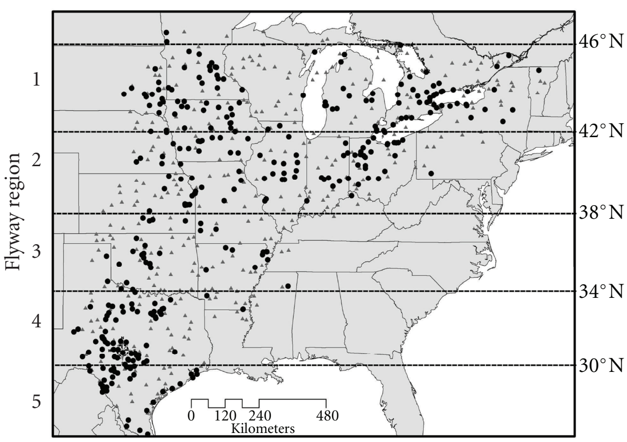


Figure 3: Map of roost observations (circles) reported to Journey North from 2005–2008 ($n = 310$) with arbitrarily created flyway regions used in this study indicated. Triangles indicate randomly selected locations ($n = 352$) for land use analysis.

are generated by Journey North personnel using a database of coordinates for North American zip (postal) codes. This practice was started for ease of overlaying points on an online, continent-scale map and since most observers do not know their latitude and longitude. For our purposes this means that the coordinates for any given observation could be centered on a point several kilometers distant from the roost (the center of the zip code). However, we attempted to minimize this problem by creating a buffer around each point with a 10 km radius (314 km^2), and evaluating the land cover within this area, which should be large enough to encompass the roost itself. The average area of zip codes in the United States is 222.7 km^2 [24], and for urban areas that have multiple zip codes in the same city this number is likely to be much smaller, which only improves the chance that the buffer encompasses the roost. To minimize spatial autocorrelation, we eliminated all duplicate coordinates of roosts that were reported in the same city (which happens when two separate observers reported the same roost, from a roost being spotted in multiple years, or from two roosts sighted near one another). This left 310 spatially independent roost observations for analysis. In addition, we randomly selected a series of points ($n = 352$) throughout each flyway region for comparison to the monarch-selected locations. For this we generated a minimum convex polygon around the entire flyway and, within that area, randomly generated points within ArcGIS, preventing points from occurring within 20 km of one another.

To evaluate land cover characteristics of both the monarch-selected and random locations we overlaid a national land cover map [25] where land cover had been digitally classified into 19 categories (though for the purposes of this project we only considered 7 of the largest categories—deciduous forest, coniferous forest, cropland, grassland, urban, open water, and wetland). Then, we calculated the percent cover of each category within the buffered area surrounding each point.

2.3. Data Analyses. There were 217 observations where the type of vegetation was specified. Using these data we compared the frequency of the most commonly reported tree species across flyway regions using chi-square statistics. Using the large-scale land cover data set containing both monarch-selected sites ($n = 310$) and randomly selected locations ($n = 352$), we used logistic regression to simultaneously examine the effects of each land use category and flyway region (predictor variables) on whether a location was monarch-selected or random (response variable). We also included two-way interaction terms between each land use category and flyway region to determine if habitat preferences vary throughout the flyway. The full model with all main and interaction effects was simplified using likelihood ratio tests (Δ deviance) to evaluate the importance of nonsignificant terms following Crawley [26]. Nested models without interaction terms were compared against the full model prior to the removal of any main effects. Significance of terms remaining in the final model are reported based on Wald χ^2. All analyses were conducted using Statistica 6.1 software [27].

3. Results

3.1. Small-Scale Roost Habitat Characteristics. Of all roost observations where the type of vegetation was specified ($n = 217$), 97.7% of the roosts were reported on trees, with the remainder being on herbaceous vegetation, including two observations of monarchs roosting on seaside goldenrod (*Solidago sempervirens*), and one each of common groundsel (*Senecio vulgaris*), beach grass (*Ammophila* sp.), and golden crownbeard (*Verbesina encelioides*). There was no clear preference for one tree type; there were a total of 38 tree species reported overall (as hosting roosts) in the four years examined. The 10 most common tree species reported are listed in Table 1, broken down by flyway region. The most frequently reported trees included pines or other conifers (21.8%), maple species (20.7%), followed by oaks (15.6%), pecans (14.5%), and willows (7.8%). Collectively, these made up 80.4% of the observations (where the tree type was specified). The frequency of these 5 tree types (i.e., their use as roost sites) appeared to vary across the flyway regions (Table 1); a 5×5 contingency table analysis based on these top five rows revealed that these frequencies differed significantly (df = 16, $\chi^2 = 108$, $P < 0.001$). In general, pines/conifers and maples were used most often in the northern areas of the flyway, while pecans and oaks were more frequently used in the southern regions.

3.2. Landscape Characteristics of Roost Sites. Of the 7 land use categories evaluated in both monarch-selected locations and random ones, the majority (~50%) of the landscapes across most flyway regions were composed of crops (Figure 4), which makes sense given that much of the central flyway traverses the agricultural region of the American Midwest (Figure 3). Following that were the broadleaf (deciduous) forest and grasslands categories. Visual comparison of the breakdown of all land use categories at monarch-selected sites (Figure 4(a)) versus random sites in the same region (Figure 4(b)) gives the impression that land use at random sites is fairly uniform throughout the flyway while that of actual roost sites varies to some degree. In particular, there appeared to be a distinct shift in the relative proportions of land use in the two southernmost regions (northern and southern Texas). In region 4, most of the land around selected roost sites was cropland (67%), while in the last region, 61% of the land around roost sites was grassland, compared to 15% around random locations in that region.

In the logistic regression model examining large-scale land use at monarch-selected sites versus random ones the results were complex. The probability of a monarch roost appeared to depend on the amount of deciduous forest, urban area, open water area, and wetland area around the site (all significant main effects; Table 2). In direct comparison of land use between roost sites and random sites, it appears that monarch-selected sites had less overall deciduous forest cover than random sites, more urban area, a higher percentage of open water nearby, and less wetland cover than random locations (Figure 5). Further, there were significant interaction effects (i.e., meaning that the strength of the main effect depended on the flyway region) in the percent deciduous forest, grassland, and urban area (Table 2).

4. Discussion

Places where monarch butterflies stop during their migration represent important links between breeding and overwintering areas, and identifying habitat requirements of roosting monarchs is therefore a key component to our understanding of this phenomenon. The ephemeral nature of migratory roosts [9], plus their broad geographic scope, makes them difficult to study using conventional scientific methodology. However, by using observations made by this nationwide network of citizen scientists, we hope to have made the first steps in addressing this question. For example, while monarch roosts were nearly exclusively on trees, we found no overwhelming preference for a tree species or type (i.e., conifer versus deciduous), other than a general tendency for maples and conifers in the north and pecans and oaks in the south (Table 1). A tendency to use males was also casually noted by F. A. and N. R. Urquhart [32] who were located in the northernmost region. When one considers the entire flyway however, given the diverse branch and leaf morphology of the various trees reported as used, it appears that monarchs are highly adaptable in terms of their roost tree use. This is also evidenced by the pictures submitted

Table 1: Summary of 10 most commonly reported tree types used for monarch roosts from 2005 to 2008, in all 5 flyway zones. Only observations where the roost tree type was specified are included.

Tree species	Flyway region 1	2	3	4	5	Total (%)
Pine/conifer (multiple species)	12	21	0	5	1	39 (21.8)
Maple (*Acer* sp.)	26	10	1	0	0	37 (20.7)
Oak (*Quercus* sp.)	4	4	1	9	10	28 (15.6)
Pecan (*Carya illinoinensis*)	0	0	4	17	5	26 (14.5)
Willow (*Salix* sp.)	4	2	1	5	2	14 (7.8)
Walnut (*Juglans* sp.)	5	3	1	0	0	9 (5.0)
Ash (*Fraxinus* sp.)	3	4	0	0	0	7 (3.9)
Elm (*Ulmus* sp.)	2	0	1	3	0	6 (3.4)
Hackberry (*Celtis occidentalis*)	0	1	0	4	0	5 (2.8)
Palm (type not specified)	0	0	0	0	4	4 (2.2)

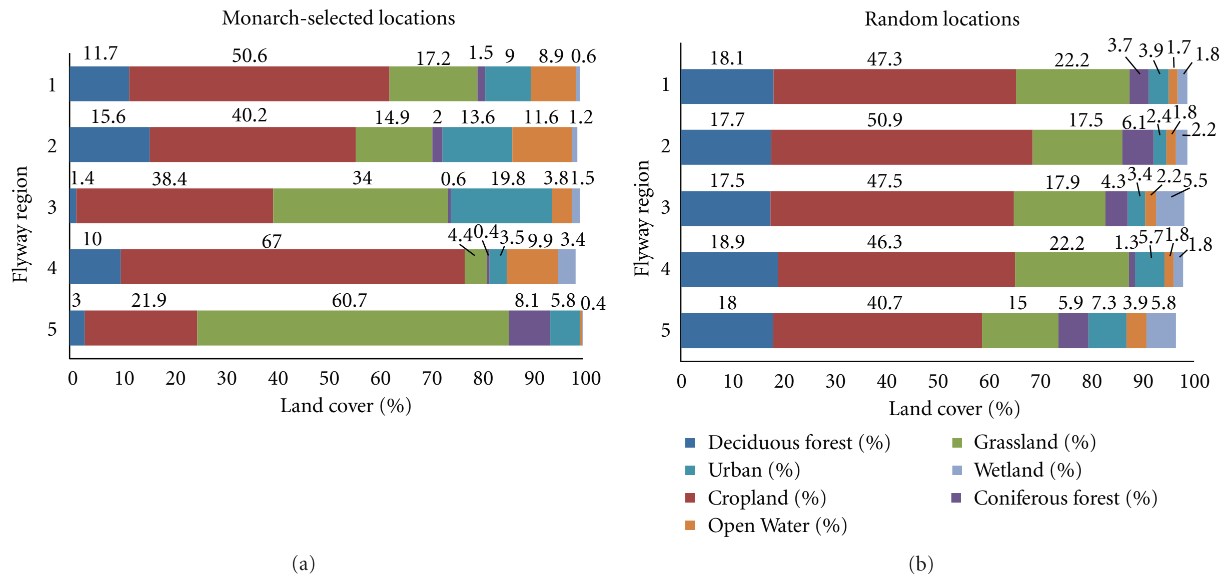


Figure 4: Relative proportions of all land-use categories in areas where roosts were observed ("monarch-selected", (a)) and in randomly selected locations (b) across all flyway regions in this study.

by Journey North participants (Figure 2); one can see that monarchs are capable of settling on a wide variety of branch structures (from needles to small-leafed trees to large-leafed trees).

Similar to the small-scale patterns obtained, from a large-scale habitat perspective, there was no one land cover type that best predicted the location of monarch roosts throughout the flyway; there was statistically significant difference in the proportion of multiple land cover types between monarch-selected and random sites (Table 2), and there were multiple interactions with flyway region. These complex results make interpreting these data difficult. One clear pattern in the land cover analyses was that, in nearly all flyway regions, there was a greater proportion of open water in the ($314\,km^2$) roost area than around random locations (Figures 4 and 5). This land cover category would include large rivers, ponds, and lakes. It may be that monarchs use these land features as beacons while searching from the air for potential roost sites, as these areas would tend to be lush with vegetation and possibly support a variety of nectaring plants. Conversely, monarchs roosted in areas that had significantly less wetland land cover than random sites did (Figure 5). Here we can only speculate as to the reason for this dichotomy; it may be that such areas are not as visible from the air as are open water bodies.

There were certain land cover patterns uncovered that may have resulted from inherent biases in observer distribution; most Journey North participants tend to live in urban areas (E. Howard, *unpublished data*). For example, monarch roosts had a higher proportion of urban land use around them than did random sites (Figure 5). This was probably an artifact of the tendency for most observers to live in or near cities (i.e., fewer observers in rural areas means fewer roost sightings). Similarly, in nearly all flyway regions, roost sites tended to have less deciduous forest area around them than did random sites (Figure 5), which could indicate either

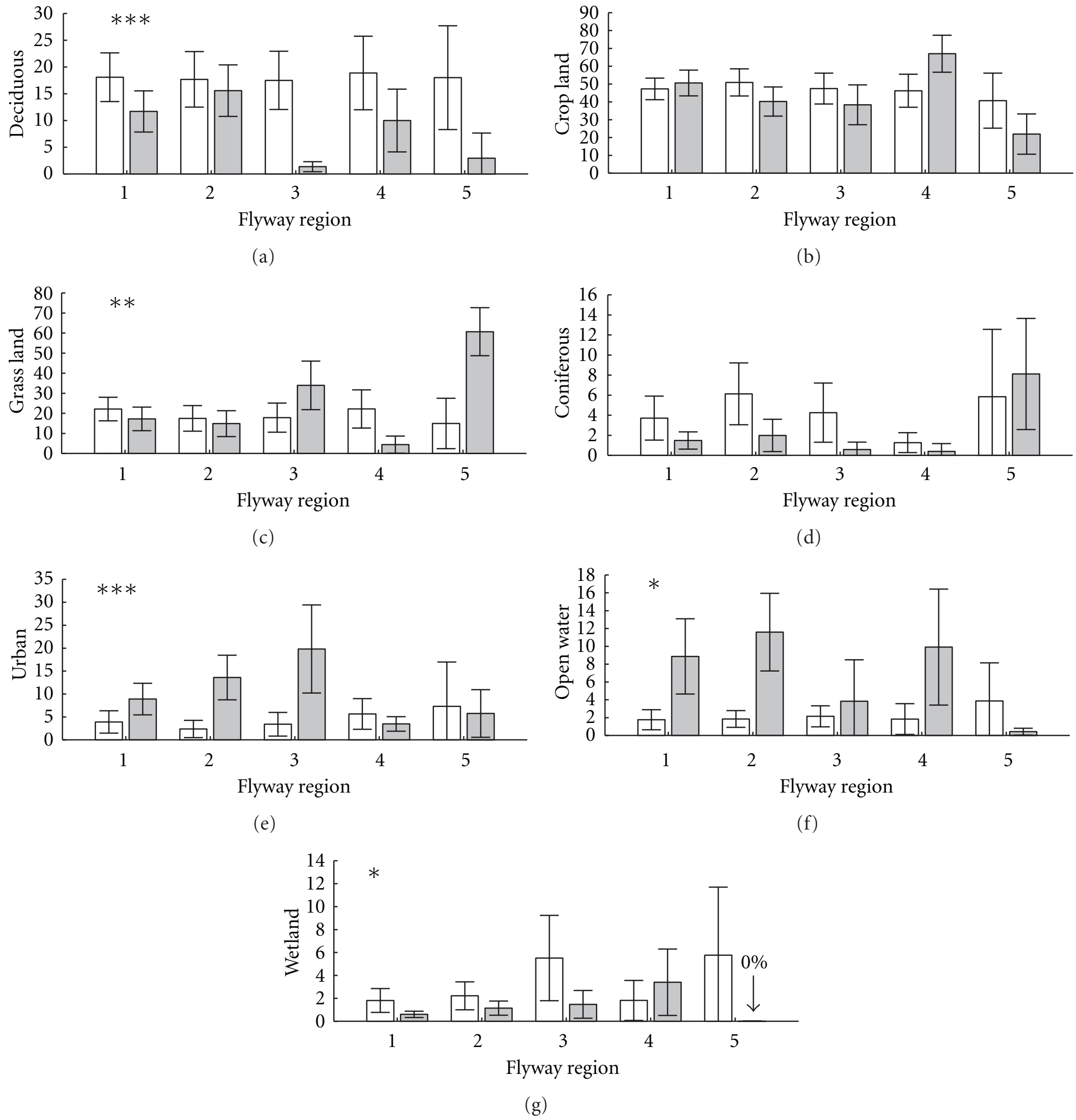


Figure 5: Percentage of land use categories in all monarch-selected (grey bars) and randomly selected (open bars) roost locations in all flyway regions. Mean percentages shown with 95% confidence intervals. Asterisks indicate significance of the main effect (*), the interaction with flyway region (**), or both (***) in the logistic regression model (Table 2).

a general avoidance of these land types for roosting purposes, or (more likely) that roosts in these habitats are not often spotted by laypeople. A roost of several hundred monarchs in a forest would not stand out as readily as one in a lone tree in the middle of a cornfield.

In addition to the significant main effects of land cover types, the logistic regression model revealed several significant interaction effects with flyway region and certain land cover types (Table 2, Figure 5), which indicates the effect of the land cover in question varied depending on the flyway region. In other words, there was some degree of change to the land cover preference (or avoidance) throughout the flyway. For example, the tendency for roosts to be associated with greater urban area was stronger in the northern regions than the southern (Figure 5), and the avoidance of deciduous forest was most pronounced in the southern regions. Further, there was a statistical effect of grassland, but it depended on the flyway region; in the southernmost region in particular, monarch roosts were in areas with 61% grassland, but this habitat was not widespread throughout that region (random sites had 15% grassland; Figure 4). In this region (southern Texas), monarchs appear to be drawn to this type of large-scale habitat.

The collective results from both the small- and large-scale analyses in this study should have conservation implications, but not in the manner we anticipated. Conserving migratory habitats is an important issue for all migrant species [28]. With this issue in mind, we had attempted to ascertain if there were certain types of small- or large-scale habitat features that could be identified as being important to

Table 2: Summary of logistic regression model examining effects of land use categories and flyway regions (predictors) on whether a location is monarch-selected or random (response). Full model with all main effects and interactions was simplified using likelihood ratio tests (Δ deviance) to evaluate the importance of nonsignificant terms following Crawley [26]. Nested models without interaction terms were compared against the full model prior to the removal of any main effects. Significance of terms remaining in the final model are reported based on Wald χ^2. The effects of cropland and coniferous forest were not significant and removed from the final model.

Variable	df	Wald	P
Flyway region	4	17.88	0.0013
Deciduous forest	1	16.04	0.0001
Grassland	1	0.027	0.8690
Urban	1	11.89	0.0006
Open water	1	21.45	0.0000
Wetland	1	10.05	0.0015
Flyway region ∗ grassland	4	23.45	0.0001
Flyway region ∗ urban	4	12.97	0.0114
Flyway region ∗ deciduous	4	13.18	0.0104

monarchs for at least one part of the stopover, their overnight roosting. However, the data we gathered in this effort did not point to a select few landscape features or roost tree types, but instead indicate that monarchs are capable of using a wide variety of habitats for roosting. Knowing this, it then becomes challenging, from a management perspective, to pinpoint what could be done to conserve important stopover areas. It may be that conservation efforts need to be targeted at the areas where it is most clear what (primary) habitats monarchs are using, such as in southern Texas, and less so in areas where there are less "habitat" preferences that can be identified, such as in the northern flyway regions. Texas is also an area of high importance for monarch migration, since, here, monarchs deposit considerable amounts of fat [29], which they will use to sustain themselves over the winter [12].

Throughout this paper we have stressed that we would not have been able to study this phenomenon of migratory roosting by any other means than by using Journey North's nationwide citizen science network. We recognize, however, that this data set was not ideal for answering our original questions regarding habitat characteristics of roosts and that there are areas where this program could be strengthened. Moreover, by highlighting these limitations (below), managers of other citizen science projects may be able to learn from these problematic issues, which may be common to many programs. Perhaps the largest drawback of the Journey North program is the fact that the coordinates of all sightings, including those of "roosts," are of the geographic center of the zip code from the observer's address, which could be several kilometers away from the actual roost. There is no remedy for this problem, unless observers take GPS readings of roost trees, which would certainly be difficult to implement into the Journey North protocol. It would also have been helpful if the protocol for reporting roosts included providing information about the surrounding habitat, such as how many trees are nearby (and not being used by monarchs) and what species they are. This would have allowed for more direct evaluation of roost tree "preference" at the sites where monarchs stop over (i.e., by comparing trees that were used to those that were not). Furthermore, in addition to habitat data, one area where Journey North could strengthen its protocol is in the reporting of the size of the roosts (i.e., number of monarchs). Many people state in their notes that they saw "hundreds" or (very often) "thousands" of monarchs in the roost observed. Estimating numbers of clustering monarchs is notoriously difficult, even for trained scientists (e.g., [30]), so this would also be difficult to implement in the current protocol. However, if actual numbers were associated with roost observations (and if we were confident in their accuracy), it would theoretically allow for annual estimates of the size of the entire migratory generation. Estimating long-term trends in abundance for this population is something that has been attempted with other data sets, but with inconsistent results [31, 32]. With the vast number of observers in the Journey North program, such data would undoubtedly be of value in this regard.

Finally, this study may well represent the first-ever examination of habitat requirements of a single migratory organism across an entire migration flyway. Such an approach allows us to identify any changes in habitat requirements at different stages of the migration. And indeed, although the monarchs' "habitat preferences" during migration appear to be broad, we did see certain changes in large- and small-scale habitat preferences as the migration advanced southward. The next step may be to assess the availability of nectaring sources at all stages of the migration, especially in the latter stages of the migration where monarchs are accumulating the most fat [29]. Additional questions regarding roosting or stopover behavior could also be addressed in the future, using citizen science observational data or direct study at specific stopover sites [20, 21]. Thanks to the efforts of hundreds of dedicated and observant people in North America who participate in citizen science programs, the answers to these and other questions are now within reach.

Acknowledgments

This project could not have been completed without the contributions of the thousands of Journey North participants who watch the skies each fall and faithfully submit roost observations. Lincoln Brower has provided expert advice to the Journey North program over the years. Funding for Journey North was provided by the Annenberg Foundation. The authors thank the members of the MonarchNet working group (Karen Oberhauser, Sonia Altizer, Leslie Ries, Dennis Frey, Becky Bartel, Elise Zipkin, James Battin, and Rebecca Batalden) for helpful discussion about the project, as well as two anonymous reviewers for suggestions for improvement on the paper.

References

[1] M. D. Prysby and K. Oberhauser, "Temporal and geographical variation in monarch densities: citizen scientists document monarch population patterns," in *The Monarch Butterfly: Biology & Conservation*, K. S. Oberhauser and M. J. Solensky, Eds., pp. 9–20, Cornell University Press, Ithaca, NY, USA, 2004.

[2] E. Lindsey, M. Mehta, V. Dhulipala, K. Oberhauser, and S. Altizer, "Crowding and disease: effects of host density on response to infection in a butterfly-parasite interaction," *Ecological Entomology*, vol. 34, no. 5, pp. 551–561, 2009.

[3] S. R. Stevens and D. Frey, "Host plant pattern and variation in climate predict the location of natal grounds for migratory monarch butterflies in western North America," *Journal of Insect Conservation*, vol. 14, no. 6, pp. 731–744, 2010.

[4] R. A. Bartel, K. S. Oberhauser, J. C. de Roode, and S. M. Altizer, "Monarch butterfly migration and parasite transmission in eastern North America," *Ecology*, vol. 92, no. 2, pp. 342–351, 2011.

[5] E. Howard, H. Aschen, and A. K. Davis, "Citizen science observations of monarch butterfly overwintering in the southern United States," *Psyche*, vol. 2010, Article ID 689301, 6 pages, 2010.

[6] E. Howard and A. K. Davis, "Documenting the spring movements of monarch butterflies with Journey North, a citizen science program," in *The Monarch Butterfly: Biology & Conservation*, K. Oberhauser and M. Solensky, Eds., pp. 105–114, Cornell University Press, Ithaca, NY, USA, 2004.

[7] A. K. Davis and E. Howard, "Spring recolonization rate of monarch butterflies in eastern North America: new estimates from citizen-science data," *Journal of the Lepidopterists' Society*, vol. 59, no. 1, pp. 1–5, 2005.

[8] E. Howard and A. K. Davis, "A simple numerical index for assessing the spring migration of monarch butterflies using data from Journey North, a citizen-science program," *Journal of the Lepidopterists' Society*, vol. 65, pp. 267–270, 2011.

[9] E. Howard and A. K. Davis, "The fall migration flyways of monarch butterflies in eastern North America revealed by citizen scientists," *Journal of Insect Conservation*, vol. 13, no. 3, pp. 279–286, 2009.

[10] F. A. Urquhart and N. R. Urquhart, "Breeding areas and overnight roosting locations in the northern range of the monarch butterfly (*Danaus plexippus plexippus*) with a summary of associated migratory routes," *Canadian Field Naturalist*, vol. 93, pp. 41–47, 1979.

[11] L. P. Brower, L. S. Fink, and P. Walford, "Fueling the fall migration of the monarch butterfly," *Integrative and Comparative Biology*, vol. 46, no. 6, pp. 1123–1142, 2006.

[12] A. Alonso-Mejía, E. Rendon-Salinas, E. Montesinos-Patiño, and L. P. Brower, "Use of lipid reserves by monarch butterflies overwintering in Mexico: implications for conservation," *Ecological Applications*, vol. 7, no. 3, pp. 934–947, 1997.

[13] S. R. Morris, M. E. Richmond, and D. W. Holmes, "Patterns of stopover by warblers during spring and fall migration on Appledore Island, Maine," *Wilson Bulletin*, vol. 106, no. 4, pp. 703–718, 1994.

[14] K. Winker, "Autumn stopover on the isthmus of tehuantepec by woodland nearctic-neotropic migrants," *Auk*, vol. 112, no. 3, pp. 690–700, 1995.

[15] C. E. Gellin and S. R. Morris, "Patterns of movement during passerine migration on an island stopover site," *Northeastern Naturalist*, vol. 8, no. 3, pp. 253–266, 2001.

[16] J. Jones, C. M. Francis, M. Drew, S. Fuller, and M. W. S. Ng., "Age-related differences in body mass and rates of mass gain of passerines during autumn migratory stopover," *Condor*, vol. 104, no. 1, pp. 49–58, 2002.

[17] J. D. Carlisle, G. S. Kaltenecker, and D. L. Swanson, "Stopover ecology of autumn landbird migrants in the boise foothills of southwestern Idaho," *Condor*, vol. 107, no. 2, pp. 244–258, 2005.

[18] D. W. Mehlman, S. E. Mabey, D. N. Ewert et al., "Conserving stopover sites for forest-dwelling migratory landbirds," *Auk*, vol. 122, no. 4, pp. 1281–1290, 2005.

[19] A. K. Davis and M. S. Garland, "Stopover ecology of monarchs in coastal Virginia: using ornithological methods to study monarch migration," in *The Monarch Butterfly: Biology & Conservation*, K. Oberhauser and M. Solensky, Eds., pp. 89–96, Cornell University Press, Ithaca, NY, USA, 2004.

[20] J. W. McCord and A. K. Davis, "Biological observations of monarch butterfly behavior at a migratory stopover site: results from a long-term tagging study in coastal South Carolina," *Journal of Insect Behavior*, vol. 23, no. 6, pp. 405–418, 2010.

[21] J. W. McCord and A. K. Davis, "Characteristics of monarch butterflies (*Danaus plexippus*) that stopover at a site in coastal South Carolina during fall migration," *Journal of Research on the Lepidoptera*, vol. 45, pp. 1–8, 2012.

[22] M. S. Garland and A. K. Davis, "An examination of monarch butterfly (*Danaus plexippus*) autumn migration in coastal Virginia," *American Midland Naturalist*, vol. 147, no. 1, pp. 170–174, 2002.

[23] L. Brindza, L. P. Brower, A. K. Davis, and T. Van Hook, "Comparative success of monarch butterfly migration to overwintering sites in Mexico from inland and coastal sites in Virginia," *Journal of the Lepidopterists' Society*, vol. 62, no. 4, pp. 189–200, 2008.

[24] Geography Division, United States Census Bureau, and U.S. Gazetteer Files, 2010, http://www.census.gov/geo/www/gazetteer/gazetteer2010.html.

[25] United States Geological Survey, Land Cover of North America at 250m, and Sioux Falls SD, 2005, http://www.cec.org/naatlas/.

[26] M. J. Crawley, *Statistical Computing: An Introduction to Data Analysis Using S-Plus*, John Wiley & Sons, Chichester, UK, 2002.

[27] "Statistica," Statistica version 6.1, Statsoft Inc., 2003.

[28] F. R. Moore, J. Sidney, A. Gauthreaux, P. Kerlinger, and T. R. Simons, "Habitat requirements during migration: important link in conservation," in *Ecology and Management of Neotropical Migratory Birds*, T. E. Martin and D. M. Finch, Eds., pp. 121–144, Oxford University Press, New York, NY, USA, 1995.

[29] L. P. Brower, "New perspectives on the migration biology of the monarch butterfly, *Danaus plexippus*, L.," in *Migration: Mechanisms and Adaptive Significance*, M. A. Rankin, Ed., pp. 748–785, The University of Texas, Austin, Tex, USA, 1985.

[30] W. H. Calvert, "Two methods of estimating overwintering monarch population size in Mexico," in *The Monarch Butterfly: Biology & Conservation*, K. Oberhauser and M. Solensky, Eds., pp. 121–127, Cornell University Press, Ithaca, NY, USA, 2004.

[31] L. P. Brower, O. R. Taylor, E. H. Williams, D. A. Slayback, R. R. Zubieta, and M. I. Ramirez, "Decline of monarch butterflies overwintering in Mexico: is the migratory phenomenon at

risk?" *Insect Conservation and Diversity*, vol. 5, no. 2, pp. 95–100, 2012.

[32] A. K. Davis, "Are migratory monarchs really declining in eastern North America? Examining evidence from two fall census programs," *Insect Conservation and Diversity*, vol. 5, no. 2, pp. 101–105, 2012.

Hindawi Publishing Corporation
International Journal of Zoology
Volume 2012, Article ID 102350, 6 pages
doi:10.1155/2012/102350

Research Article

From Citizen Science to Policy Development on the Coral Reefs of Jamaica

M. James C. Crabbe

Institute of Biomedical and Environmental Science Technology and Faculty of Creative Arts, Technologies and Science, University of Bedfordshire, Park Square, Luton LU1 3JU, UK

Correspondence should be addressed to M. James C. Crabbe, james.crabbe@beds.ac.uk

Received 2 November 2011; Revised 30 November 2011; Accepted 6 December 2011

Academic Editor: Richard Stafford

This paper explores the application of citizen science to help generation of scientific data and capacity-building, and so underpin scientific ideas and policy development in the area of coral reef management, on the coral reefs of Jamaica. From 2000 to 2008, ninety Earthwatch volunteers were trained in coral reef data acquisition and analysis and made over 6,000 measurements on fringing reef sites along the north coast of Jamaica. Their work showed that while recruitment of small corals is returning after the major bleaching event of 2005, larger corals are not necessarily so resilient and so need careful management if the reefs are to survive such major extreme events. These findings were used in the development of an action plan for Jamaican coral reefs, presented to the Jamaican National Environmental Protection Agency. It was agreed that a number of themes and tactics need to be implemented in order to facilitate coral reef conservation in the Caribbean. The use of volunteers and citizen scientists from both developed and developing countries can help in forging links which can assist in data collection and analysis and, ultimately, in ecosystem management and policy development.

1. Introduction

Coral reefs throughout the world are under severe challenges from a variety of anthropogenic and environmental factors including overfishing, destructive fishing practices, coral bleaching, ocean acidification, sea-level rise, algal blooms, agricultural run-off, coastal and resort development, marine pollution, increasing coral diseases, invasive species, and hurricane/cyclone damage [1–3]. It is the application of citizen science to help generation of scientific data and capacity-building, and so underpin scientific ideas and policy development in the area of coral reef management, that are explored in this paper, concentrating on Jamaican coral reefs.

The "compulsive" appetite for increasing mobility [4] allied to a social desire for extraordinary "peak experiences" [5] has led to the modern "ethical consumer" for tourism services [4, 6] derived from the "experiential" and "existential" tourist of the 1970s [7]. Several organisations have taken the concept of ecotourism further to embracing tourism with citizen science, whereby the tourist gets to work on research projects under the supervision of recognised researchers. Several organisations worldwide have developed citizen science programmes. The drivers behind these activities vary significantly between scientific studies, education, and/or getting the public more engaged and raising awareness of the natural environment. The overall driver cannot only determine the type, quality, and quantity of data required but also the level of volunteer expertise needed. Three organisations that have developed citizen science with tourism are the Earthwatch Institute, (http://www.earthwatch.org/), Operation Wallacea (http://www.opwall.com/), and Coral Cay Conservation (http://www.coralcay.org/). All are international environmental charities, working with a wide range of partners, from individuals who work as conservation volunteers on research teams through to corporate partners (such as HSBC with Earthwatch), governments, and institutions. Research volunteers work with scientists and social scientists around the world to help gather data needed to address environmental and social issues. It is the long-term strategy of these organisations combined with their citizen science funding models that underpins their successes; they are "in for the long haul" and can effect conservation in a different way to a standard 3-year research grant. Key elements are developing projects that can be used by volunteers and verifying the scientific information in a statistically significant way. This paper shows how an Earthwatch programme using volunteers on coral reefs generated scientific information which was used to inform management strategies in Jamaica.

2. Materials and Methods

2.1. Training of Volunteers. All volunteers were SCUBA divers of at least PADI Open Water standard. Training took place at the Discovery Bay Marine Laboratory, Jamaica, and consisted of lectures and interactive discussions covering scleractinian coral biology and taxonomy, coral recognition, data measurements and analysis, and health and safety. Volunteers all had to accomplish open water diving tests, and coral recognition tests in the field, after studying coral taxonomy books and passing land-based tests.

2.2. Reef Sites. Four randomly located transects, each 15 m long and separated by at least 5 m, were laid at 5–8.5 m depth at each of five sites on the North coast of Jamaica near Discovery Bay: Rio Bueno (18° 28.805′ N; 77° 21.625′ W), M1 (18° 28.337′ N; 77° 24.525′ W), Dancing Ladies (18° 28.369′ N; 77° 24.802′ W), Dairy Bull (18° 28.083′ N; 77° 23.302′ W), and Pear Tree Bottom (18° 27.829′ N; 77° 21.403′ W). These sites were chosen as being workable by volunteers, as they were with 20 min boat ride from the Discovery Bay Marine Laboratory, where all volunteers stayed. Sites had been studied before over a number of years by marine scientists from many countries. GPS coordinates were determined using a hand-held GPS receiver (Garmin Ltd., UK).

2.3. Citizen Science Data Collection. Corals 2 m either side of the transect lines were photographed for archive information, and surface areas were measured with flexible tape as described previously using SCUBA [8–10]. Depth of samples was between 5 and 8.5 m, to minimise variation in growth rates due to depth [11]. To increase accuracy, surface areas rather than diameters of live nonbranching corals were measured [8, 9]. Sampling was over as wide a range of sizes as possible. Colonies that were close together (<50 mm) or touching were avoided to minimise age discontinuities through fission and altered growth rates [12–14]. In this study *Montastrea annularis* colonies were ignored, because their surface area does not reflect their age [12], and because hurricanes can increase their asexual reproduction through physical damage [13]. Overall, over 6,000 measurements were made on over 1,000 coral colonies, equally distributed between the sites for species and numbers of colonies.

This work was conducted at Discovery Bay during July 15–31 and December 19–30 in 2000, March 26–April 19 in 2002, March 18–April 10 in 2003, July 23–August 21 in 2004, July 18–August 13 in 2005, April 11–18 in 2006, December 30 in 2006–January 6 in 2007, and July 30–August 16 in 2008. Surveys were made at the same locations at the same sites each year. Data from ninety volunteers was used over this period.

2.4. Storm Severity. Data on storm severity as it impacted the island was obtained from UNISYS (http://weather.unisys.com/hurricane/atlantic/), the NOAA hurricane site (http://www.nhc.noaa.gov/pastall.shtml). Information on bleaching was obtained from the NOAA coral reef watch site (http://coralreefwatch.noaa.gov/satellite/current/sst_series_24reefs.html).

2.5. Data Analysis. Data analysis on corals was using ANOVA. Skewness (sk, [15]) was used to estimate the distribution of small and large colonies in the coral populations around Discovery Bay in Jamaica. In a normal distribution, approximately 68% of the values lie within one standard deviation of the mean. If there are extreme values towards the positive end of a distribution, the distribution is positively skewed, where the mean is greater than the mode (the mode is the value that occurs the most frequently in a data set) (right tail is longer). The opposite is true for a negatively skewed distribution, where the mean is less than the mode (left tail is longer). With regard to coral populations, negative skewness implies more large colonies than small colonies, while positive skewness implies more small colonies than large colonies.

3. Results

3.1. Coral Sizes and Growth. All the Jamaican sites showed some similarities in distribution of the size classes for the species studied between 2002 and 2008. However, there were differences between the different sites and between the different species studied at the sites. Skewness values (sk) were used to compare the distribution of the data between 2002 and 2008. For *S. siderea*, all sk values were positive, with more small colonies than in a normal distribution for 2002 and 2008, with little change between the dates (all sk values between 0.5 and 1.6). With *D. labyrinthiformis* colonies, there was a change from negative skewness in 2002 at Dairy Bull and Pear Tree Bottom, with more large colonies than in a normal distribution (sk values −0.25 and −0.006, resp.) to smaller colonies than in a normal distribution in 2008 (sk values of 0.20 and 0.97, resp.). There were no significant changes from 2002 to 2008 at the other sites, with positive sk values from 0.1 to 0.89. *M. meandrites* colonies at Rio Bueno and Dairy Bull showed a relative decrease in the distribution of larger colonies from 2002 to 2008, with changes in sk values from −0.03 in 02 to 0.78 in 08, and from −0.05 to 0.03, respectively; the other sites all exhibited slightly positive sk values in both years from 0.1 to 0.5. For *Agaricia* species, there was very little change between the years at all the sites, with sk values from 0.4 to 1.6. For *P. astreoides*, all values were positive for both years, with an increase in skewness at Rio Bueno from 0.2 to 2.6, showing a marked change in distribution towards the smaller colony sizes. At the other sites there were only small increases in sk values from 2002 to 2008, with Pear Tree Bottom showing a decrease in skewness from 0.9 to 0.6. *D. strigosa* colonies showed similar results to *P. astreoides*, all sk values being positive for 2002 and 2008, with an increase at Rio Bueno from 0.2 to 2.2 and at Pear Tree Bottom from 0.4 to 2.4; other sites showed similar sk values for 2002 and 2008 from 0.6 to 1.6. *C. natans* skewness changed from −0.07 to 0.68 at Rio Bueno from 2002 to 2008 (a decrease in larger colonies relative to a normal distribution) and at Dancing Ladies from −0.31 to 0.38. Other sites showed similar skewness in 2002 and 2008 (sk values between 0.5 and 0.6), except Pear Tree Bottom, which exhibited near normal distribution of colonies about the mean for both 2002 and 2008 (sk values <0.01). Interestingly, in 2005, the year after hurricane Ivan, the most severe storm to impact the reef sites over the study period, there was a slight reduction in the numbers of the smallest size classes, particularly notable at Dairy Bull.

In addition, our volunteer studies showed that radial growth rates (mm/yr) of non-branching corals calculated on an annual basis from 2000 to 2008 showed few significant differences either spatially or temporally along the North coast, although growth rates tended to be higher on reefs of higher rugosity and lower macroalgal cover [16].

3.2. Extreme Climate Events. The only extreme climate event that significantly impacted the Jamaican reef sites during the study period was the mass Caribbean bleaching event of 2005 [17]. Analysis of satellite data showed that there were 6 degree heating weeks (dhw) for sea surface temperatures in September and October 2005 near Discovery Bay, data which was mirrored by data loggers on the reefs.

3.3. Development of Coral Reef Action Plan. The coral size and growth data collected by the citizen scientists show that corals of above average size for their species at the sites studied lack resilience, particularly after the major bleaching event of 2005. Because of this, there is a need for different zones to have different levels of protection. To this end, the data was used in the development of an action plan for Jamaican coral reefs, presented to the Jamaican National Environmental Protection Agency, and described in Table 1.

Table 1: Seven-point action plan for Jamaican coral reefs.

(1) The reefs around Jamaica could be designated as the Jamaican Coral Reef Marine Park. This could include all the fringing reefs, seagrass beds, and mangroves from Negril to all along the north coast to the eastern tip of the island. On the south coast it could include Port Royal and Portland Bight. The advantage of this is that one can then consider protection of the Jamaican reefs as a whole. Another advantage is that climate change effects can be considered in a more holistic way
(2) There could be a single body, possibly the National Environment Protection Agency (NEPA), or a subset of NEPA, given authority to manage the Park
(3) There could be a statement drawn up on "protection and wise use" of the Park. Drawing up that statement should include all stakeholders, from fishermen through Industry and tourism to policy makers
(4) The Park could be managed using a "zoning" system. This has been valuable in a number of areas, not least the Great Barrier Reef. This will allow some areas to have greater restrictions (e.g., fishing, resort pollution, ship pollution) than others. Such zoning should help avoid the "tragedy of the commons". Zoning Plans define what activities can occur in which locations, both to protect the marine environment and to separate potentially conflicting activities
(5) Divisions into zones could be General Use, Conservation Park, Habitat Protection, Marine National Park, Another zone might be a Buffer Zone, next to a Marine National Park
(6) Each zone should have at least one of the following: (i) Community Partnerships, (ii) Local Marine Advisory Committees, and (iii) Reef Advisory Committees. These bodies should be responsible for regulating their own area and should be responsible to the overall Marine Park Management body. They would also be responsible for community involvement and information
(7) Permissions within the zones (e.g., for tourism, fishing, etc.) would be given by the Jamaican Government, through NEPA

4. Discussion

4.1. Citizen Science and Use of Volunteer Data. Citizen science and use of data measured by volunteers has been very helpful in a number of zoological areas, including amphibian population and biodiversity studies [18, 19], reporting invasive species [20], environmental monitoring [21], evolutionary change [22], marine species abundance and monitoring [23–25], dryland mapping [26], and conservation planning [27, 28].

This study used self-selected as "Earthwatch volunteers", and all were SCUBA divers. Motivation was high in all the volunteers, as was the validity of the data presented by volunteers. A key element in citizen science is good training of volunteers. In the area of coral reef research described in this study, training was given in species recognition, quantitative measurement techniques and validation, and data analysis. Independent validation of volunteer data, once training had been given, was consistent with previous findings by other groups [29]. The validation of the data produced by the volunteers indicated that with appropriate training, data collection by citizen scientists is appropriate for scientific applications in marine biology.

4.2. Coral Health and Resilience. What is apparent from our studies is that despite the chronic and acute disturbances between 2002 and 2008, demographic studies indicate good levels of coral resilience on the fringing reefs around Discovery Bay in Jamaica (see also [30]). The bleaching event of 2005 resulted in mass bleaching but relatively low levels of mortality unlike corals in the US Virgin islands and Tobago where there was extensive mortality [17, 31], probably because of their greater degree heating week values.

This data shows that while recruitment of small corals is returning after the major bleaching event of 2005 [32], larger corals are not necessarily so resilient and so need careful management if the reefs are to survive such major extreme events.

4.3. From Information to Policy Development: Themes and Tactics. Marine reserves are an important tool in the sustainable management of many coral reefs [33]. However, it is important that the reef ecosystems share regulatory guidelines, enforcement practices and resources, and conservation initiatives and management, underpinned by scientific research. An example of a single marine reserve is the Great Barrier Reef in Australia operated and managed solely by the Great Barrier Reef Marine Park Authority (GBRMPA). In contrast, the second largest barrier reef in the world, the MesoAmerican Barrier Reef, is bounded by four countries (Mexico, Belize, Guatemala, and Honduras), each with its own laws and policies. Here, a number of single and separated marine reserves exist along the barrier reef. In Belize we have successfully transferred scientific expertise in Belize to local volunteers to generate scientific evidence to underpin future management and conservation decisions, as judged, for example, by scientific findings on the impact of hurricanes on reefs in Belize, which showed that hurricanes and severe storms limited the recruitment and survival of nonbranching corals of the Mesoamerican barrier reef [10].

For Jamaica, the Action Plan developed (Table 1) was well received by managers of the National Environment Protection Agency (NEPA). It was felt by managers that this approach could link together the environment with tourism and business, so that environmental issues are seen as part of the way forward, not part of the problem, as has been all too evident in the past. Even if smaller Marine Protected Areas (MPAs) were developed around the island, the adoption of shared ownership of reef ecosystems was felt to be useful way to proceed.

In order to take this forward, it was felt necessary to develop a number of themes and tactics. In a separate capacity building exercise [34], for the MesoAmerican Barrier Reef in Sothern Belize, one officer from the Belize Fisheries Department, three senior officers from NGOs involved in managing Belize MPAs (TIDE, the Toledo Institute for Development and Environment; TASTE, the Toledo Association for Sustainable Tourism and Empowerment; and Friends of Nature), and a Facilitator (the author) from the UK developed six-month Personal/Professional Action Plans which involved

(a) tactics for leading, educating, and supporting issues regarding sustainable development of coral reefs;

(b) tactics for collaboration with other stakeholders to collectively influence policy decisions for coral reef conservation.

Discussion among the participants and facilitator resulted in the generation of a series of generic tactics to be adopted around a number of themes. These are enumerated in Table 2. Such themes and tactics may be useful in development of coral reef policies in the Caribbean and elsewhere.

Table 2: Themes and tactics to facilitate conservation of coral reefs.

Organisation and Management
Tactic number 1: establish a key leader in the Organization/Department to effectively manage the Marine reserves on a day-to-day basis
Tactic number 2: have a selected key leader provide general Terms of Reference of what is expected of staff and immediate/major stakeholders in order to easily facilitate the process of decision making
Education
Tactic number 1: financial resources need to be allocated for an education program. The program should focus on both broad and specific issues that may create friction among stakeholders in the process
Tactic number 2: a group consisting of community leaders and key/immediate stakeholders should be established to create ways and methods of educating different levels of stakeholders in the effectiveness of sustainable development in the marine parks
Tactic number 3: surveys need to be conducted to evaluate level of success and failure. Too often programmes have been formed and implemented but end results have not been evaluated. Surveys should be carried back to stakeholders for a presentation to establish further steps
Support
Tactic number 1: a well-put together presentation needs to be developed and be presented to the key authority that will have over-all say in the marine park(s). This will stress on the support needed to accomplish both the mission and vision statements and will have positive effects in sustainable development
Tactic number 2: nonmonetary incentives need to be established in order to have full support of stakeholders who would otherwise deter progress in sustainable development
Policies
Tactic number 1: establish a set of policies that is considered necessary for proper management of the marine reserves. Such policies will be established by all stakeholders involved
Tactic number 2: create an influencing program for stakeholders to adhere to such policies through an education/retreat program
Tactic number 3: establish exchanges with other organizations in capacity building in policy creation and effective implementation

5. Conclusion

The use of volunteers and citizen scientists from both developed and developing countries can help in forging links which can assist in data collection and analysis and, ultimately, in ecosystem management and policy development. There is much progress internationally in involving organisations to utilize citizen science effectively and efficiently (e.g., [35]).

A number of questions remain for the future, for example, assessing how citizen science could be used to better effect, for example, identifying the potential for citizen science to fill known data gaps, for example, gaps in marine and terrestrial taxonomies. In addition, we need greater understanding of where and how technology (software, statistics) can transform the quality and quantity of data from nonexperts, and how scientists can make best use of technology, for example, in using smart phone apps to identify and/or record species and measurements.

Acknowledgments

The author thanks the Earthwatch Institute, the Royal Society, and the Oak Foundation (USA) for funding, Mr. Anthony Downes, Mr. Peter Gayle, and the staff of the Discovery Bay Marine Laboratory for their invaluable help and assistance, to the two anonymous referees for their valuable comments which improved the manuscript, to E. Martinez, C. Garcia, J. Chub, L. Castro, J. Guy, and Earthwatch colleagues in Belize, and to many volunteers for their considerable help underwater during this project.

References

[1] T. A. Gardner, I. M. Côté, J. A. Gill, A. Grant, and A. R. Watkinson, "Long-term region-wide declines in Caribbean corals," *Science*, vol. 301, no. 5635, pp. 958–960, 2003.

[2] D. R. Bellwood, T. P. Hughes, C. Folke, and M. Nyström, "Confronting the coral reef crisis," *Nature*, vol. 429, no. 6994, pp. 827–833, 2004.

[3] M. J. C. Crabbe, E. E. L. Walker, and D. B. Stephenson, "The impact of weather and climate extremes on coral growth," in *Climate Extremes and Society*, H. F. Diaz and R. J. Murnane, Eds., pp. 165–188, Cambridge University Press, New York, NY, USA, 2008.

[4] P. Burns, "Tribal tourism-cannibal tours: tribal tourism to hidden places," in *Niche Tourism: Contemporary Issues, Trends, Cases*, pp. 101–110, Elsevier, Oxford, UK, 2005.

[5] A. Holden, *Environment and Tourism*, Routledge, London, UK, 2nd edition, 2008.

[6] D. Fennell, *Ethical Tourism*, Routledge, London, UK, 2006.

[7] E. Cohen, "A phenomenology of tourist experiences," *Sociology*, vol. 13, pp. 179–201, 1979.

[8] M. J. C. Crabbe, J. M. Mendes, and G. F. Warner, "Lack of recruitment of non-branching corals in Discovery Bay is linked to severe storms," *Bulletin of Marine Science*, vol. 70, no. 3, pp. 939–945, 2002.

[9] M. J. C. Crabbe and D. J. Smith, "Sediment impacts on growth rates of Acropora and Porites corals from fringing reefs of Sulawesi, Indonesia," *Coral Reefs*, vol. 24, no. 3, pp. 437–441, 2005.

[10] M. J. C. Crabbe, E. Martinez, C. Garcia, J. Chub, L. Castro, and J. Guy, "Growth modelling indicates hurricanes and severe storms are linked to low coral recruitment in the Caribbean," *Marine Environmental Research*, vol. 65, no. 4, pp. 364–368, 2008.

[11] M. Huston, "Variation in coral growth rates with depth at Discovery Bay, Jamaica," *Coral Reefs*, vol. 4, no. 1, pp. 19–25, 1985.

[12] T. P. Hughes and J. B. C. Jackson, "Do corals lie about their age? Some demographic consequences of partial mortality, fission, and fusion," *Science*, vol. 209, no. 4457, pp. 713–715, 1980.

[13] N. L. Foster, I. B. Baums, and P. J. Mumby, "Sexual vs. asexual reproduction in an ecosystem engineer: the massive coral Montastraea annularis," *Journal of Animal Ecology*, vol. 76, no. 2, pp. 384–391, 2007.

[14] R. Elahi and P. J. Edmunds, "Consequences of fission in the coral Siderastrea siderea: growth rates of small colonies and clonal input to population structure," *Coral Reefs*, vol. 26, no. 2, pp. 271–276, 2007.

[15] J. H. Zar, *Biostatistical Analysis*, Prentice-Hall, Upper Saddle River, NJ, USA, 4th edition, 1999.

[16] M. J. C. Crabbe, "Climate change and tropical marine agriculture," *Journal of Experimental Botany*, vol. 60, no. 10, pp. 2839–2844, 2009.

[17] C. M. Eakin, J. A. Morgan, S. F. Heron et al., "Caribbean corals in crisis: record thermal stress, bleaching, and mortality in 2005," *PLoS One*, vol. 5, no. 11, article e13969, 2010.

[18] A. Bonardi, R. Manenti, A. Corbetta et al., "Usefulness of volunteer data to measure the large scale decline of "common" toad populations," *Biological Conservation*, vol. 144, no. 9, pp. 2328–2334, 2011.

[19] D. Sewell, T. J. C. Beebee, and R. A. Griffiths, "Optimising biodiversity assessments by volunteers: the application of occupancy modelling to large-scale amphibian surveys," *Biological Conservation*, vol. 143, no. 9, pp. 2102–2110, 2010.

[20] T. Gallo and D. Waitt, "Creating a successful citizen science model to detect and report invasive species," *BioScience*, vol. 61, no. 6, pp. 459–465, 2011.

[21] C. C. Conrad and K. G. Hilchey, "A review of citizen science and community-based environmental monitoring: issues and opportunities," *Environmental Monitoring and Assessment*, vol. 176, pp. 273–291, 2011.

[22] J. Silvertown, L. Cook, R. Cameron et al., "Citizen science reveals unexpected continental-scale evolutionary change in a model organism," *PLoS One*, vol. 6, no. 4, article e18927, 2011.

[23] C. A. Ward-Paige, C. Pattengill-Semmens, R. A. Myers, and H. K. Lotze, "Spatial and temporal trends in yellow stingray abundance: evidence from diver surveys," *Environmental Biology of Fishes*, vol. 90, pp. 263–276, 2010.

[24] S. Goffredo, F. Pensa, P. Neri et al., "Unite research with what citizens do for fun: recreational monitoring of marine biodiversity," *Ecological Applications*, vol. 20, no. 8, pp. 2170–2187, 2010.

[25] P. G. Finn, N. S. Udy, S. J. Baltais, K. Price, and L. Coles, "Assessing the quality of seagrass data collected by community volunteers in Moreton Bay Marine Park, Australia," *Environmental Conservation*, vol. 37, no. 1, pp. 83–89, 2010.

[26] D. S. Turner and H. E. Richter, "Wet/dry mapping: using citizen scientists to monitor the extent of perennial surface flow in dryland regions," *Environmental Management*, vol. 47, no. 3, pp. 497–505, 2011.

[27] P. De Ornellas, E. J. Milner-Gulland, and E. Nicholson, "The impact of data realities on conservation planning," *Biological Conservation*, vol. 144, no. 7, pp. 1980–1988, 2011.

[28] D. B. Oscarson and A. J. K. Calhoun, "Developing vernal pool conservation plans at the local level using citizen-scientists," *Wetlands*, vol. 27, no. 1, pp. 80–95, 2007.

[29] P. J. Mumby, A. R. Harborne, P. S. Raines, and J. M. Ridley, "A critical assessment of data derived from coral cay conservation volunteers," *Bulletin of Marine Science*, vol. 56, no. 3, pp. 737–751, 1995.

[30] M. James and C. Crabbe, "Coral resilience on the reefs of Jamaica," *Underwater Technology*, vol. 30, no. 2, pp. 65–70, 2011.

[31] J. Mallela and M. J. C. Crabbe, "Hurricanes and coral bleaching linked to changes in coral recruitment in Tobago," *Marine Environmental Research*, vol. 68, no. 4, pp. 158–162, 2009.

[32] M. J. C. Crabbe, "Environmental effects on coral growth and recruitment in the Caribbean," *Journal of the Marine Biological Association of the UK*. In press.

[33] L. Cho, "Marine protected areas: a tool for integrated coastal management in Belize," *Ocean and Coastal Management*, vol. 48, no. 11-12, pp. 932–947, 2005.

[34] M. J. C. Crabbe, E. Martinez, C. Garcia, J. Chub, L. Castro, and J. Guy, "Is capacity building important in policy development for sustainability? A case study using action plans for sustainable marine protected areas in Belize," *Society and Natural Resources*, vol. 23, no. 2, pp. 181–190, 2010.

[35] "UK—Environmental Observation Framework, Workshop Report," pp. 45, 2011, http://www.ukeof.org.uk/documents/.ukeof-citizen-science-workshop-report.pdf.