Transforming Citizen Science Point Data into Informative Range Maps

Laura Brianna Gast

PENNSYLVANIA STATE UNIVERSITY – MGIS CAPSTONE PROJECT

Advisor: Nathan Piekielek, PhD

9 January 2017

Transforming Citizen Science Point Data into Informative Range Maps

Laura Brianna Gast

Introduction

Birdwatching has become increasingly more popular in recent years. In 1982, for instance, only about 12 percent of Americans (21.2 million people) reported birding as a hobby. In 2001, that number rose to 33 percent (70.4 million people) and has been steadily rising since. Furthermore, the demographics of those that birdwatch as a hobby are more diverse in terms of age, education, race, and income level. [1] Birdwatchers are also typically enthusiastic participants in conservation efforts. A number of studies have shown that birders make substantial contributions to conservation initiatives that increase with birding experience. These contributions are not only monetary, but non-monetary, in the form of volunteer hours or as participation in citizen science events. [2]

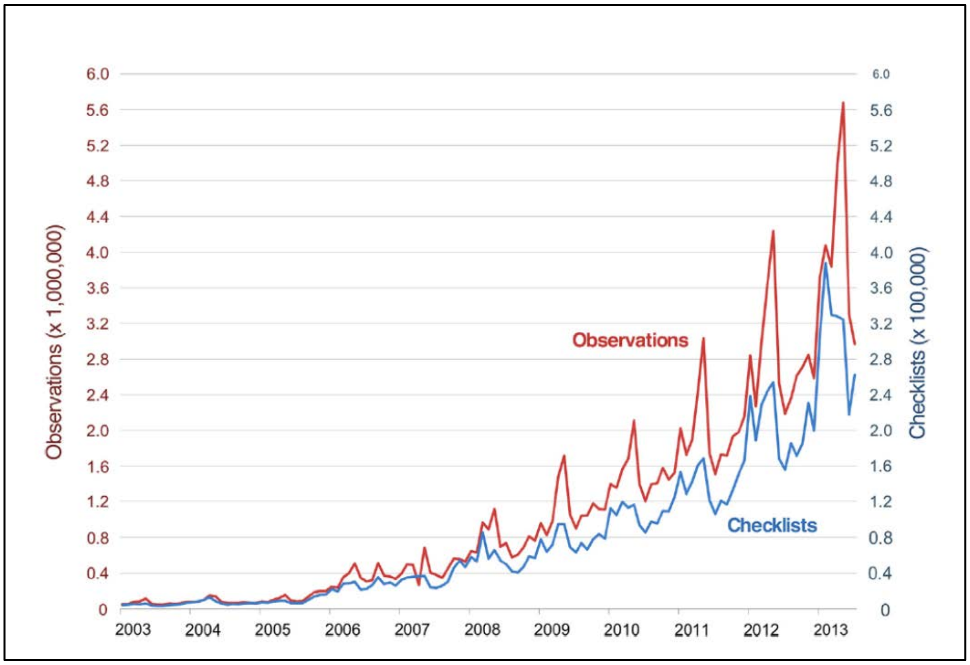
Citizen Science

Broadly speaking, citizen science refers to research collaborations between scientists and volunteers, particularly (but not exclusively) to expand opportunities for scientific data collection and to provide access to scientific information for community members. There are projects across a range of scientific pursuits, form astronomy to oceanography, ecology, and ornithology.

One of the oldest citizen science programs is Audubon's Christmas Bird Count (CBC). First launched in 1900 as an alternative to the Christmas Hunt, it has served as the largest and longest continuously running animal census. Volunteers gather to identify and count every bird seen or heard within a 15-mile radius circle on a set date sometime between December 14 and January 5 each year. The CBC serves as a basis for the identification and demarcation of the wintering range for many avian species. The data from the census is used in making models to predict how climate change (and other habitat changes like recovering or establishing new preserves) might affect avian habitats. [3] However, their models have the base assumption that the CBC data is consistent and with non-systematic error. While efforts have been made to statistically control for these errors – for instance, by controlling for effort hours and weather conditions during each count period - if historical CBC data is incorrect, it follows that those models relying on that data might be imperfect.

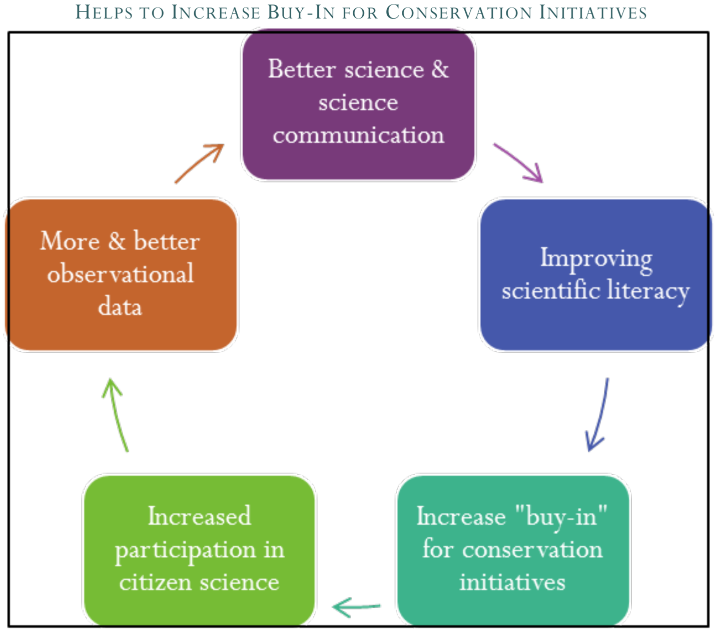
A different citizen science program called eBird was launched in 2002 with the specific goal to maximize the utility and accessibility of the vast numbers of bird observations made each year by birdwatchers of all abilities. For birders, eBird acts as a real-time online checklist program, allowing users to summarize and explore their personal bird checklist data. These observations are then joined together and shared with the global community of educators, ornithologists and conservation biologists. It would be impossible for ornithologists to collect this much data for their research. For example, in May 2015 alone, eBird participants recorded over 9.5 million bird observations. When Trying to understand species distribution, abundance and movement, these observations are extremely useful. [4]

FIGURE 1: EBIRD SUBMISSIONS OVER TIME, AS TOTAL CHECKLISTS SUBMITTED (RIGHT), AND AS THE NUMBER OF INDIVIDUAL BIRDS REPORTED WITHIN THOSE CHECKLISTS (LEFT) [4]



As more and better observational data is available, we are able to produce better science. Effective communication of these results, along with improved scientific literacy, allows scientists to better share their knowledge and concerns about environmental issues with others, increasing buy-in from citizens for conservation initiatives. Engaging more citizens with conservation efforts has been shown to lead to increases in participation in related citizen science programs (Figure 2). [5]

FIGURE 2: SUMMARY OF HOW BETTER SCIENCE COMMUNICATION HELPS TO INCREASE BUY-IN FOR CONSERVATION INITIATIVES



In this white paper, we provide a case exploration of published range maps and citizen science reported observations for single species, the Black-bellied Whistling Duck. We hope that the points mentioned in this article might allow for a discussion about improving guide maps to aid citizen scientists in the field.

Field Guide Books

Field guides are important pieces of science communication, as they allow for interested individuals to venture out into the field, seek out, and identify flora, fauna, or other objects (e.g., minerals). These guides are an indirect way to make the knowledge of specialists and experts available to interested amateurs. Modern birding field guides exist to teach readers how to identify birds by showing both an image of the bird alongside a short text description. Within the image, most guides provide arrows or text indicating key identification points; descriptions tend to include behavioral characteristics, song or call descriptions, as well as maps showing the expected distribution and migration patterns of the specific species. [6] These maps are a critical component of a well-designed field guide: for a person to identify a single bird, when many species look very similar, knowing that only one of those similarly appearing species should be present in that area is immensely helpful. On the other hand, if a birder in the field truly believes they saw a certain bird, but the guide book shows that the species isn't supposed to be anywhere near that particular region at that time of year (or at all), it can be incredibly disheartening and discouraging. This situation occurs often for amateur birders, or even experienced birders in new countries or regions, and this is what the following case study seeks to explore.

Exploratory Case Study: The Black-Bellied Whistling Duck in the Central USA

The black-bellied whistling duck (BBWD) is a mostly non-migratory bird, although some of the most northern flocks may move slightly south in the peak of winter. Within the US, it can be found in three general populations, grouped around Florida, the Texas-Louisiana Gulf coast, and in southeast Arizona. [7] The central population of BBWD began to appear in the late 1960s, and have been increasing in abundance and distribution since that time. [8] The duck itself is quite noisy, with a clear whistling call, and is quite tame even in the wild. As such, they are quite obvious when they are present in their habitat of quiet, shallow, freshwater ponds, lakes, marshes, and cultivated reservoirs. They have taken nicely to man-made lakes in subdivisions in south-central Texas.

This population of ducks was chosen for this project as they are broadly known within birding communities to be rapidly expanding their wintering range in Texas and into neighboring states, yet guide books do not yet reflect this expansion. [9] [7] [10] As such, this population of BBWD serves as an obvious example of THE THING

Guidebook Ranges for BBWD

An individual looking at any of the three of the most widely-used guides for bird identification might come away with different understandings of the distribution and abundance of the central population of BBWD. The Cornell Laboratory (Figure 3), Sibley Guide (Figure 4), and Audubon Guide (Figure 5) range maps show distinctly different wintering ranges for this group.

Figure 3: Cornell Lab Online map for BBWD Distribution [8]

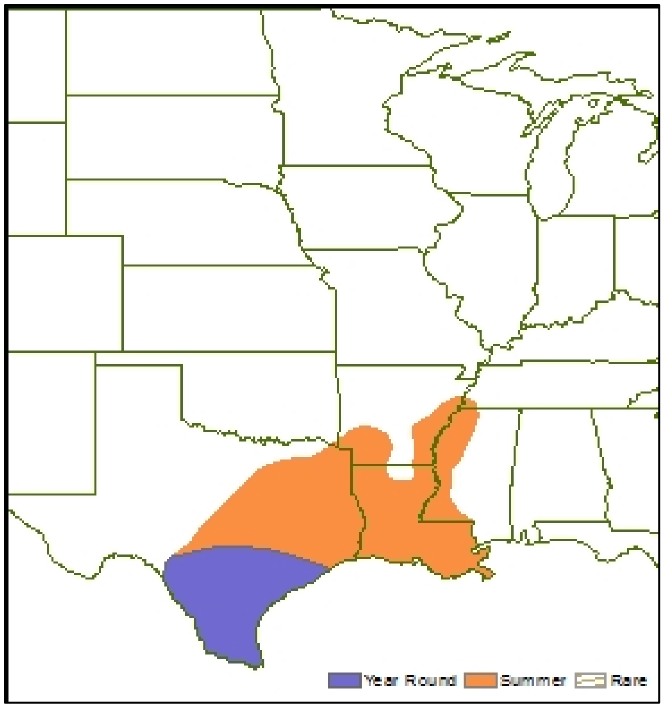


Figure 4: Sibley Guide (iPhone App) for BBWD Distribution

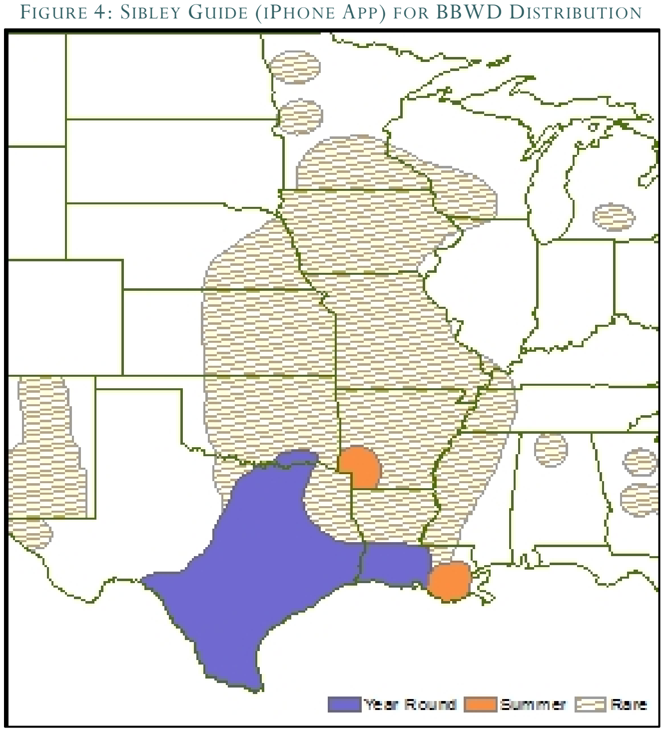
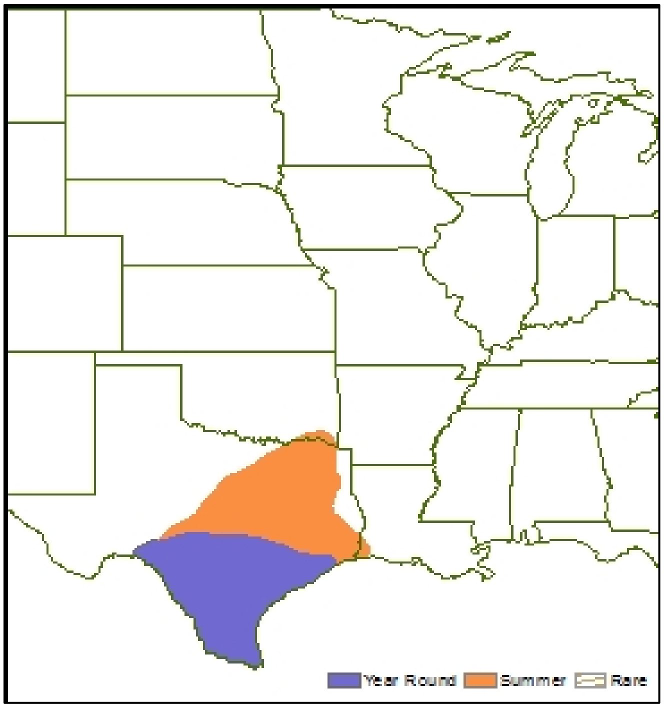
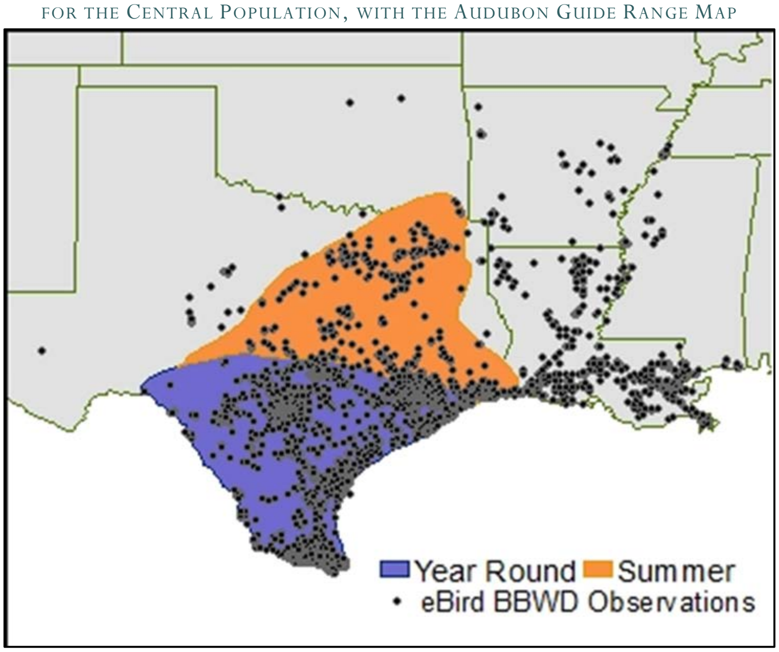


Figure 5: Audubon Guide (iPhone App) for BBWD Distribution



These maps are correct from a scientific angle; they represent the outcome of statistically significant predictive niche modeling, based upon observational data, expert knowledge and other environmental or habitat data. However, when compared with observed records from eBird, the matter of fact is that what observers are seeing in the field, is not the same as what is shown in these guide book range maps (Figure 6).

Figure 6: eBird BBWD Observations reported in December and January for the Central Population, with the Audubon Guide Range Map



In order to explore these discrepancies, we performed a case study comparing citizen science observational data to published range maps. We have distilled these potential sources of difference into three main groups, (a) model choices related to the data sources, (b) model choices related to timing of data sources, and (c) cartographic choices, namely the methods by which scientists display their predictive model outputs.

For the citizen science observational studies, we obtained the entire eBird dataset for all BBWD observations in the United States, [11] as well as the Audubon Christmas Bird Count data for all BBWD observations [12]. For published range maps, we digitized and georeferenced images of publish range maps from the Sibley iPhone App [7], Cornell Laboratory online [8], and the Audubon guide book iPhone App [10]. Data manipulation was performed in R (Version 3.3), and all mapping was performed in ArcGIS (Version 10.4).

Difference in Data Source

As is always the fact in statistical analyses, small changes in the base dataset can change the outcome of any predictive model. When examining the Audubon guide book wintering ranges, the CBC dataset appears to align slightly more than the eBird data from the corresponding time periods (December and January) (Figure 7 & Figure 8). Unfortunately, it is apparent that the guide book does not reflect what is actually being observed and reported. For instance, the guide book indicates that the BBWD is only present in central and northern Texas during the summer, yet both observational datasets clearly show that this is not the case. Furthermore, both observational datasets show reliable sightings of BBWD along the Louisiana coast, something that is unaccounted for in the published range map.

The guidebook states that range maps are made "using three different modeling methods that fit complex, non-linear relationships between the species occurrence data and environmental data. [They] then validate each of the models with independent data from historical time periods, compare their predictive ability, and chose the one that performed best overall" [13]. For prediction models to be accurate, delineation of current and historical ranges must be accurate. As National Audubon Society has described in their major report detailing the predicted effects of climate change on the habitats of birds, "predictions about the future require the development of models, and all models entail uncertainty. In the case of climate change, our best hope for making sound conservation decisions is to account for uncertainty to the degree possible." [14] Any details or data sources used for these predictive models are kept secret, perhaps to keep these complex models as proprietary information.

FIGURE 7: CHRISTMAS BIRD COUNT OBSERVATIONS AS COMPARED TO AUDUBON GUIDE RANGE MAPS

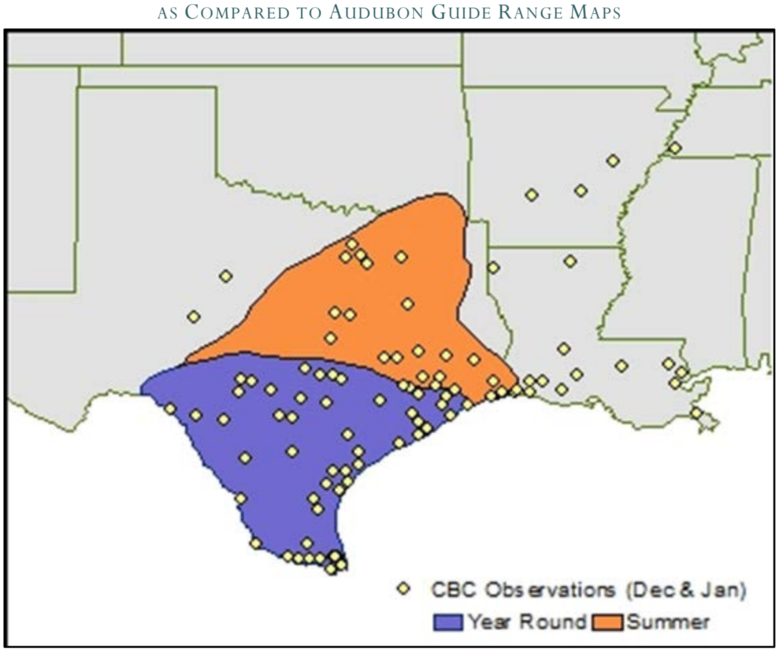
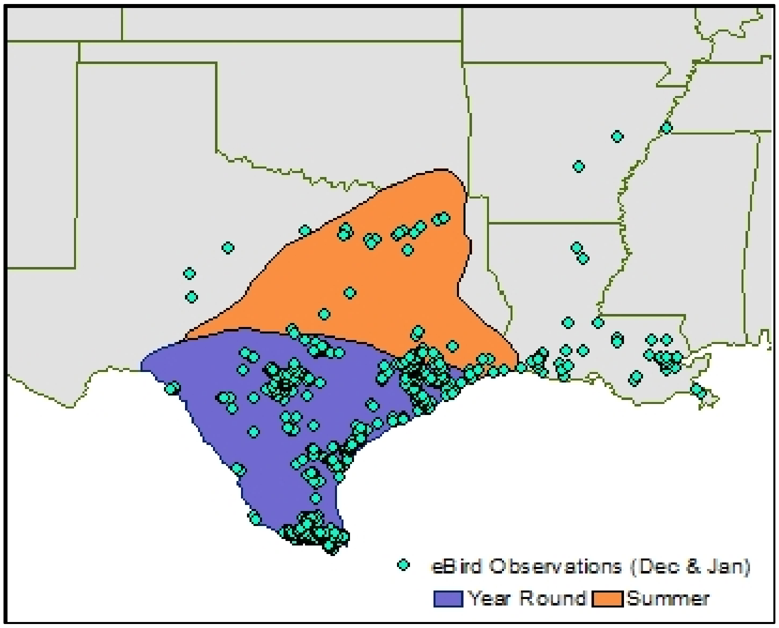


FIGURE 8: EBIRD OBSERVATIONS (DECEMBER & JANUARY)
AS COMPARED TO AUDUBON GUIDE RANGE MAPS



Differences in Time of Update or Publication

Another probable reason for the discrepancies between actual observations and published range maps is that our observational data is available in near-real time through the near-instantaneous publication of eBird observations and the rapid publication of annual CBC datasets. When examining eBird observations for December and January, recorded with dates between 1969 and 2006 (eBird allows for retroactive recording and reporting datasets copied from personal datasets), these records align reasonably well with the published Audubon guide book ranges (Figure 9). However, the Cornell Guidebook range better aligns with eBird observations for December and January recorded between 2007 and 2012 (Figure 10).

It is important to note, that while each guide book or guide app has a recorded date of publication for the entire book, there is no information regarding when each individual map was updated or when (or if) the model for each range map was re-evaluated to compare to more recent observational data. Modern technology should be able to remedy this issue. In this examination, two of the guide book sources were collected from smartphone applications that are readily and inexpensively available. While printed guide books are static, these applications could be updated whenever new information is available, at little to no cost to either the publisher or the user. Furthermore, with advances in computing power, it would be possible to automate the creation of new maps after the inclusion of new information. For instance, it is not inconceivable that in the near future, observational datasets from a source such as eBird could be automatically downloaded, prediction algorithms could be re-run with the new dataset, and the resulting maps could be sent to users in the form of an update.

FIGURE 9: EBIRD OBSERVATIONS FOR 1969 - 2006 (DECEMBER & JANUARY) AS COMPARED TO THE AUDUBON GUIDE BOOK RANGE

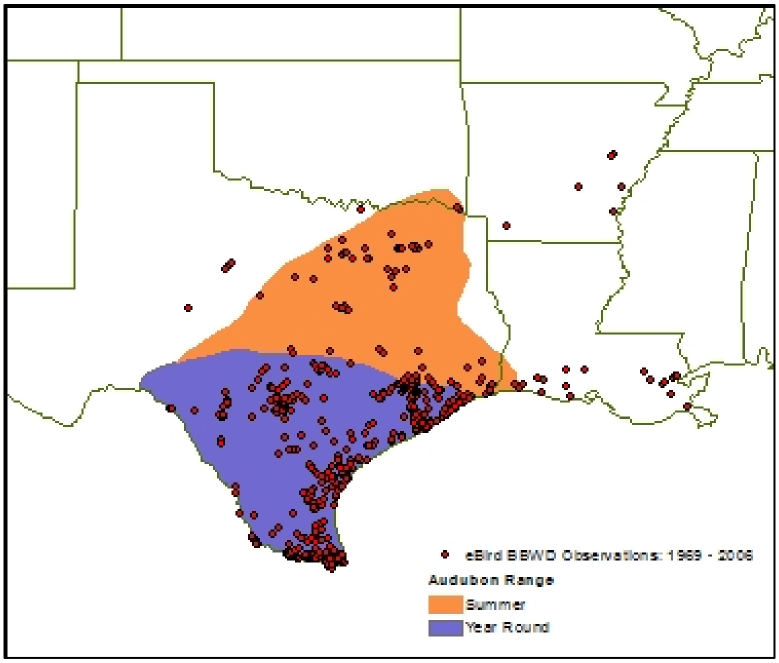
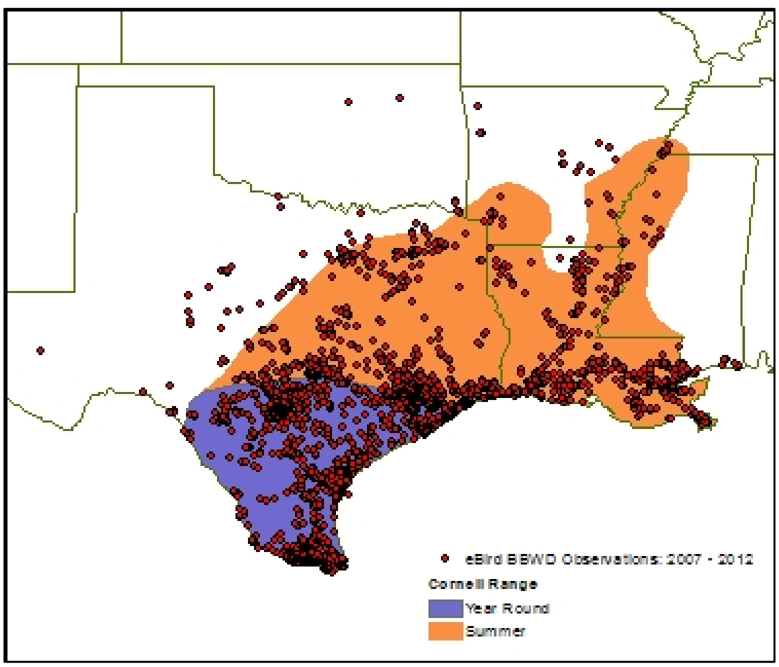


FIGURE 10: EBIRD OBSERVATIONS FOR 2007 - 2012 (DECEMBER & JANUARY) AS COMPARED TO THE CORNELL GUIDE BOOK RANGE



Definition of Edges

A further opportunity to better communicate with guide book maps lies in how cartographers define the 'presence' or 'absence' of a particular species. eBird has one method of addressing this, where all observations are recorded in a raster image, with the probability of seeing that particular species (for a user-defined time period or season) is displayed on a color scale (Figure 11). [11] The Sibley guide book uses a different method, delineating a "rare" category alongside their defined seasonal ranges (Figure 12). [7] These are simple ways for scientists and cartographers to provide their audience with a bit more information about the likelihood of seeing a particular bird in that particular area at a particular time, in a clear and concise manner.

FIGURE 11: EBIRD RASTER OF PROBABILITIES OF A BBWD SIGHTING (DECEMBER AND JANUARY), BASED UPON REPORTED OBSERVATIONS

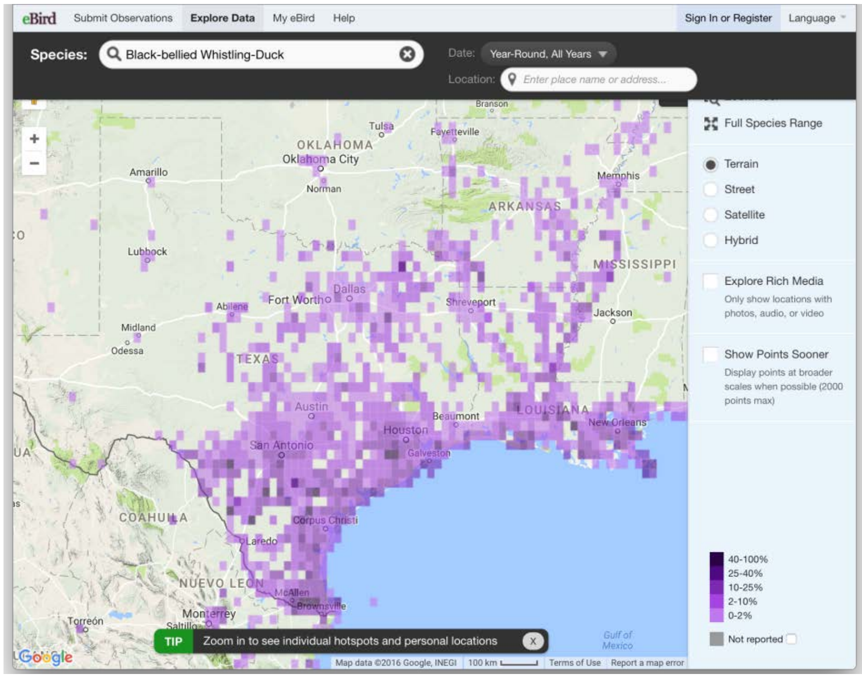
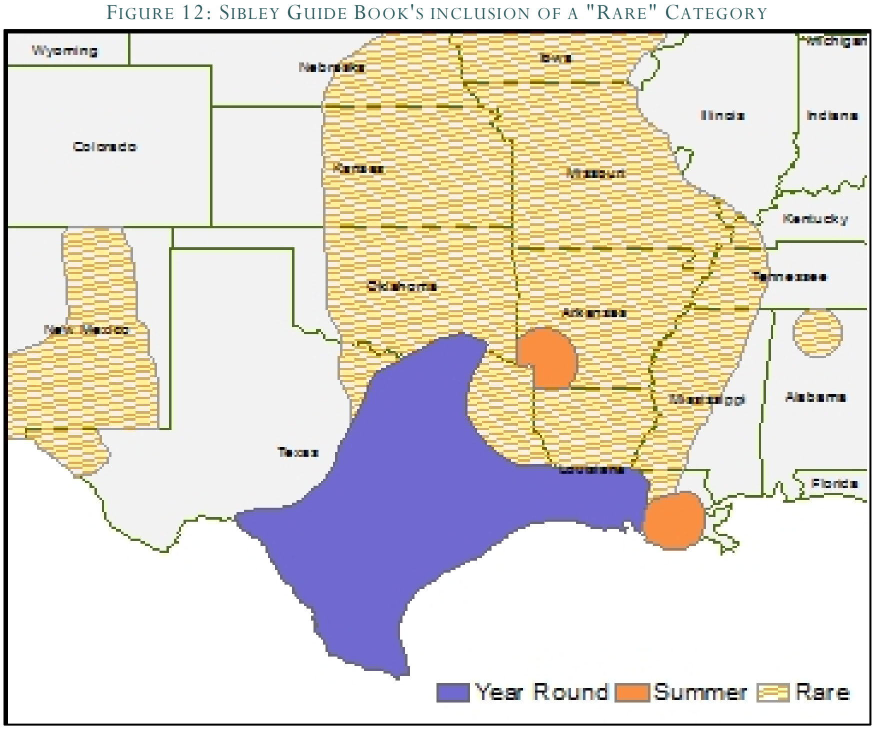


FIGURE 12: SIBLEY GUIDE BOOK'S INCLUSION OF A "RARE" CATEGORY



Conclusions

Science communication is critical to good and effective policy, as well as to develop community buy-in for conservation efforts. As we move from a data poor environment to a data rich environment with the ever increasing participation in birdwatching and submissions to citizen science projects like eBird and the Christmas Bird Count, the prediction models that guide book range maps are built upon should improve as well. Furthermore, with the improvements in automation and computing power, along with app-based and digital maps and near-real-time observational data, we have the opportunity to leave behind the more static and unchanging range maps toward better representations of current trends. The primary function of maps is to communicate, and this functionality depends on the appearance of the map; the explicit and implicit decisions made by the mapmakers, including how to define the 'presence' or 'absence' of a species in a given location in a given season. The best data, analyses, and conclusions in the world will be completely lost if they are poorly communicated. Improving guide books by better tailoring maps to aid citizen scientists in identifying birds in the field will help to build on the growing enthusiasm for birding in particular and conservation efforts in general.

References

[1] H. K. Cordell and N. G. Herbert, "The Popularity of Birding is Still Growing," *Birding,* pp. 54-61, February 2002.

[2] B. L. McFarlane and P. C. Boxall, "Participation in Wildlife Conservation by Birdwatchers," *Human Dimensions of Wildlife,* vol. 1, no. 3, pp. 1-14, 1996.

[3] National Audubon Society, "Conservation: Audubon Science," [Online]. Available: http://www.audubon.org/conservation/science. [Accessed 10 August 2016].

[4] B. L. Sullivan, J. L. Aycrigg, J. H. Barry, R. E. Bonney, N. Bruns, C. B. Cooper, T. Damoulas, A. A. Dhondt, T. Dietterich, A. Farnsworth, D. Fink, J. W. Fitzpatrick, T. Fredericks, J. Gerbracht, C. Gomes, W. M. Hochachaka, M. Iliff, C. Lagoze, F. A. La Sorte, M. Merrifield, W. Morris, T. B. Phillips, M. Reynolds, A. D. Rodewald, K. V. Rosenberg, N. M. Trautmann and A. Wiggins, "The eBird enterprise: An integrated approach to development and application of citizen science.," *Biological Conservation,* vol. 169, pp. 31-40, January 2014.

[5] S. K. Jacobson, Communcation Skills for Conservation Professionals, Washington, D.C.: Island Press, 2009.

[6] N. Lund, "The High-Tech Future of Staring at Birds in the Woods," Slate.Com, December 2013. [Online]. Available: http://www.slate.com/articles/arts/books/2013/12/birding_field_guides_and_apps_technology_changes_bird_watching_and_bird.html.

[7] Sibley Guides, "Black Bellied Whistling Duck".

[8] "The Birds of North America Online," Cornell Laboratory of Ornithology, 2015. [Online]. Available: http:bna.birds.cornell.edu/BNA..

[9] R. Gibbons, Interviewee, *Conservation Director, Houston Audubon Society.* [Interview]. 22 December 2015.

[10] National Audubon Society, "Black Bellied Whistling Duck," 2016.

[11] Cornell Laboratory of Ornithology, "eBird," Ithica.

[12] National Audubon Society, "The Christmas Bird Count Historical Results [Online].".

[13] K. Kaufman, "Kaufmman Field Guides," 2015. [Online]. Available: http://www.kaufmanfieldguides.com/.

[14] National Audubon Society, "Audubon's Birds and Climate Change Report," National Audubon Society, Washington, DC, 2015.

[15] R. Burden, "Illuminative Evaluation," in *Frameworks for Educational Psychology*, London, Jessica Kingsley, 2008, pp. 218-234.