MIAMI UNIVERSITY
The Graduate School

Certificate for Approving the Dissertation

We hereby approve the Dissertation

of

Ancilleno Davis

Candidate for the Degree

Doctor of Philosophy

James T. Oris, Co-Advisor

R. Hays Cummins, Co-Advisor

David Russell, Reader

Mary Henry, Reader

Thomas Crist, Committee Member

Jing Zhang, Graduate School Representative

ABSTRACT

CHANGING PERSPECTIVES ON CITIZEN SCIENCE USING EBIRD DATA ON GRAND BAHAMA ISLAND, THE BAHAMAS.

by

Ancilleno O. Davis

Citizen science has broadened the scope of biodiversity monitoring and research. Citizen scientists visit more locations, more often and collect data on more species than any single study can. They have fewer restrictions related to funding, scheduling and political will. They create more data than ever before, especially in remote locations such as Small Island Developing States (SIDS). However, citizen science uses traditional science perspectives in data analysis; acknowledging the perspectives of the citizen scientist is important when making conservation decisions based on citizen science data. I use novel perspectives that make citizen science data more useful/powerful. I focus on 16 native bird species and 20 migrant species of international concern using volunteer observations from the open access eBird database. Using forestry maps and satellite data, I created a new, adaptable, classified habitat map for Grand Bahama and appended the habitat data to eBird observations for the island. Observer effort was significantly higher in beach and grass habitats. I found most of the focal species in this study outside their documented habitat type. Bird species richness and observer richness differed significantly among habitat types. Bird species composition was significantly influenced by habitat type and survey effort. Mantel tests showed significant correlation between geographic locations and both bird species dissimilarity and observer dissimilarity. The Mantel tests also showed significant correlations between observer community differences and species community differences. I used Moran's I to determine spatial autocorrelation of observer effort and recorded species diversity within the dataset. Observer richness and the total number of surveys were negatively spatially autocorrelated in the overall dataset. I found that observer community similarity showed significant effects on recorded survey effort and species diversity in most habitats.

CHANGING PERSPECTIVES ON CITIZEN SCIENCE USING EBIRD DATA ON GRAND BAHAMA ISLAND, THE BAHAMAS.

A DISSERTATION

Presented to the Faculty of

Miami University in partial

fulfillment of the requirements

for the degree of

Doctor of Philosophy

Department of Biology

by

Ancilleno O. Davis

The Graduate School
Miami University
Oxford, Ohio

2018

Dissertation Directors: James T. Oris & R. Hays Cummins

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

DEDICATION

This work is dedicated to my family, Leonardo and Oliver Davis Ramirez who were born during the last years of my PhD program, and Alma, my loving wife who has been beside me throughout. I could not have done it without their presence and support.

ACKNOWLEDGMENTS

I would like to acknowledge the professors on my committee who came through with financial and logistic support in the final years of my program to ensure I completed the degree. The Government of the Bahamas' Forestry Department for granting access to their data to support the research and for trusting me with its use. Thanks to the graduate students in Miami University's Geography and Biology departments that helped me work through lines of R and JavaScript or had a sympathetic ear when I needed to vent, especially, Justin, Lauren, Tony and Meghan. Finally, a special thanks to the support staff throughout the University that helped me as I worked across disciplines and learned to navigate the Miami University culture, in the Statistics Department, Western Program for Individualized Studies, The Office of Diversity Affairs, The Graduate School, International Students and Scholar Services, and the Center for Teaching Excellence.

General Introduction

Citizen science has allowed us to broaden the scope of monitoring activities significantly over the past century (Dickinson, Zuckerberg, & Bonter, 2010). Though increased observer diversity is known to improve data quality (Burgess et al., 2017; Danielsen, Burgess, Jensen, & Pirhofer-Walzl, 2010), few studies have looked at the effect observer behavior or ability has on the collected data (Johnston, Fink, Hochachka, & Kelling, 2018; Kelling et al., 2015). This study focuses on quantifying the variation in observer community composition in eBird data (Davis, 2017) on the island of Grand Bahama and determining the effect observer communities may have on the biodiversity record. In particular, I look at observer habitat use, differences in the time spent surveying and the number of surveys conducted, and differences in species records. I use the observer behavior to determine knowledge gaps in bird distribution; consider new relationships between species occurrences that take into account differences in monitoring; and predict potential species locations based on the data. The work comprises three chapters.

Chapter 1: "Combining citizen science and open source geospatial techniques improves habitat knowledge for Bahamian birds", consists of the creation of an adaptable, open access habitat map using the Google Earth Engine platform (Gorelick et al., 2017a) and appending the newly generated habitat data to bird occurrence data collected by citizen scientists via eBird (Sullivan et al., 2009, 2016). The map combines ecologically meaningful habitat classifications from Bahamian government sources and was used to determine differences in observer effort and bird distribution among habitats. I compared the distribution of effort among habitats to determine if the effort was proportional to the habitat distribution on the island. I used citizen science data in combination with the appended habitat data to determine if species were found within their published habitats, as indicated in popular bird guides (Raffaele, Wiley, Garrido, Keith, & Raffaele, 1998; Wardle, Gape, & Moore, 2014; White, 1998) and scientific literature (Currie, Wunderle, Ewert, Davis, & McKenzie, 2005; Lloyd & Slater, 2010, 2011).

Chapter 2: "*Bird Species Diversity and Dissimilarity and their Relationships to Habitat and Observer Variation on Grand Bahama, The Bahamas*" explores the spatial relationships between bird communities in eBird data on the island and the relationship between the observer communities and the bird communities they detect. First, I examine patterns of bird species richness as it relates to habitat type, numbers of observers, and total survey effort. Variation in bird species composition was examined using multivariate ordination with habitat type and survey effort as predictor variables. I used Mantel tests (Mantel & Valand, 1970, 2018) to determine whether the differences in observer communities and reported bird communities correlated to the distance between locations or to one another. I used Moran's I (Gittleman & Kot, 1990) to determine whether the observer effort and bird diversity data was spatially autocorrelated within the eBird dataset. I discuss how biological processes may affect species distribution, but observer behavior follows different patterns. I discuss how the bird species or amenities may affect observer activity at different locations. I also discuss the importance of observer identities in determining the level of survey activities the species identified.

Chapter 3: "*Here together, there together: species and observer co-occurrence in eBird records on Grand Bahama, The Bahamas"* explores co-occurrence in eBird data. I calculate probabilistic co-occurrence matrices (Griffith, Veech, & Marsh, 2016; Veech, 2013) for the species reported on the island and the observers that reported them. I used co-occurrence matrices to determine species that share locations or habitats within the data set and discuss whether these species co-occurrences may be ecologically significant and useful in monitoring or assessment of native and migrant species of conservation concern. Similarly, I conducted co-occurrence analysis on the species using the observers as sites to determine whether certain species are similarly detectable by observers in eBird. I calculated observer co-occurrence matrices at localities to detect observers with significant overlap in the locations they visit. I discuss the implications of observer co-occurrence for tour operator or community monitoring program managers as these co-occurrences imply observer knowledge of and access to monitoring locations. Similarly, I created observer co-occurrence matrices using bird species as sites to determine if some observers were significantly more likely to detect the same species or groups of species. I discuss co-occurrence for observers at species and how these co-occurrences may indicate similarity between observer experience or interest and how these co-occurrences may impact the reliability of conservation or monitoring decisions. Finally, locality co-occurrences were calculated using species and observers as sites. Localities that co-occur within the record of species indicates that the localities have similar resources to provide for those species. Similarly, localities that co-occur within observer groups may indicate that they are marketed or accessible to particular types of guest such as golf courses, local school grounds, and cruise ship ports. I discuss the implications of these co-occurrences on monitoring program development and conservation planning.

Study Area

Grand Bahama lies approximately 100 km east of Florida, USA and is the northernmost island in the Bahamian archipelago. It is on a migratory pathway (Lincoln, 1879) between North America and South America. Some species, which breed in North America, may overwinter on the island or are year round residents (Raffaele et al., 1998; White, 1998).

Grand Bahama includes several cays to the east, which are considered part of the island, but these were removed from analysis. Analysis focused on the contiguous terrestrial area of Grand Bahama Island because birder access to the cays would require additional logistic considerations such as a boat or ferry.

eBird Data

eBird is an online database, digital mobile app and web-based software platform that allows observers to record birds they have seen along with time and location data (Sullivan et al., 2009). Users can include details such as the number of birds seen within each species, age and sex of the bird as well as details about the trip and other observers that were with them. Website users can search the data by date, location, species or by observer (observers can choose whether or not their identity is public); many active users allow their names to be publicly available. The unique opportunity to view the observer names or unique id in this citizen science database allows us to test the relationships between individual and group behavior and examine biases associated with observers, including site or habitat preferences and thus the reliability of describing avian species habitat preferences and the composition of avian communities by habitat. Observer ability may determine the quality of data in the data set.

Habitat data

The eBird Basic Dataset (EBD) does not include habitat data for the Bahamas. Habitat types were derived from historical forestry maps of Grand Bahama Island. Seven classes are used in this study: water, pine, wetland, sand, urban, grass and high water table communities (hwtc). Habitat assessment and classification was conducted using Landsat 8 satellite images and high spatial resolution satellite images in Google Earth Engine's API.

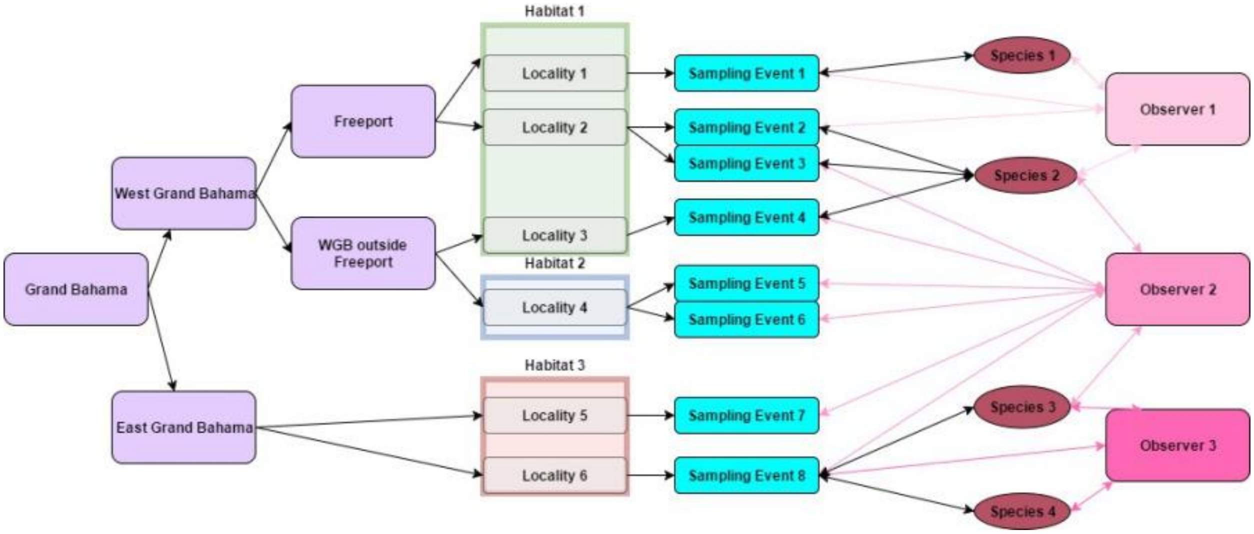


Figure 1: Basic data hierarchy within eBird data with added habitat information.

Open-Source, Open-access commitment

This project utilized free and open source data and software to classify the habitat and conduct all data analysis. All required code to create the products are included both in the appendices and online at https://github.com/Ancilleno. I use open access data, software and code to facilitate application of this research in the Bahamas and insular Caribbean or other Small Island Developing States where access to proprietary software and other materials may be limited.

Chapter 1: Combining citizen science and open source geospatial techniques improves habitat knowledge for Bahamian birds[1]

Ancilleno O. Davis[1], Jessica L. McCarty[2]

[1] Ecology, Evolution, and Environmental Biology, Miami University, Oxford, OH 45056 USA, email: ancillenodavis@gmail.com;

[2] Department of Geography, Miami University, Oxford, OH 45056 USA, email: jmccarty@miamioh.edu

Abstract (250-300 words)

Bird biodiversity contributes to tourism economies and reflects ecosystem health at multiple scales. Monitoring and conservation of avian communities requires knowledge of the habitats that birds are associated with, moving beyond single habitat observations. Many formal surveys of avian communities focus on narrow spatial or temporal scales. Formal ornithological studies often ignore broader communities to focus on a few species or populations, especially in the Caribbean. Landscape-scale habitat maps would be useful to identify bird and habitat relationships but global land cover maps are too coarse to represent habitat accurately within Small Island Developing States (SIDS), such as The Bahamas, and the categories used may not be locally relevant. Using citizen science data from sources like eBird and Audubon Christmas Bird Counts broadens our knowledge taxonomically with data on more bird species than any one single study. Volunteer observers visit more habitats and locations than are visited in published studies, with more repeat visits over longer periods. Satellite imagery and semi-automated habitat classification by local experts combined with citizen science may improve knowledge of habitat use for bird species. This study provides one method for identifying species habitat use, by combining satellite imagery and citizen science data on bird occurrence in predominantly open source geospatial and statistical softwares and platforms. Adding habitat classes to citizen science data will allow stakeholders to assess monitoring gaps and better understand habitat use for birds of conservation concern. Educators and conservation agencies can also use habitat data to guide outreach efforts and encourage outreach activities that reduce habitat bias in citizen science data sets.

Keywords: Bahamas, citizen science, eBird, habitat classification, geospatial, open source, remote sensing

---

[1] Prepared for and submitted to Journal of Caribbean Ornithology

Abstracto (Español)

La biodiversidad de las aves contribuye a las economías turísticas y refleja la salud de los ecosistemas en múltiples niveles. El monitoreo y la conservación de las comunidades avícolas requieren el conocimiento de los hábitats que las aves usan. Muchas encuestas formales de comunidades aviares se enfocan en escalas espaciales o temporales estrechas. Los estudios ornitológicos formales a menudo ignoran las comunidades más amplias para centrarse en unas pocas especies o poblaciones, especialmente en el Caribe. Los mapas de hábitats a escala de paisaje serían útiles para identificar las relaciones de aves y hábitat, pero los mapas de cobertura terrestre global son demasiado toscos para representar el hábitat con precisión dentro de los Pequeños Estados Insulares en Desarrollo (Small Island Developing States) como las Bahamas, y las categorías utilizadas no son relevantes a nivel local. El uso de datos de la ciencia ciudadana de fuentes como eBird y Audubon Christmas Bird Counts amplía nuestro conocimiento taxonómicamente con datos sobre más especies de aves que cualquier otro estudio. Los observadores voluntarios visitan más hábitats y lugares que los visitados en estudios publicados, con más visitas repetidas durante períodos más largos. Las imágenes satelitales y la clasificación de hábitats asistidos por computadora por expertos locales combinados con la ciencia ciudadana pueden mejorar el conocimiento del uso del hábitat de las especies de aves. Este estudio proporciona un método para identificar el uso del hábitat de las especies, mediante la combinación de imágenes satelitales y datos de ciencia ciudadana sobre la ocurrencia de aves. Agregar clases de hábitat a datos de ciencia ciudadana permitirá a las partes interesadas evaluar vacíos de monitoreo y comprender mejor el uso del hábitat para aves de interés para la conservación. Los educadores y las agencias de conservación también pueden usar los datos de hábitat para guiar los esfuerzos de divulgación y alentar las actividades de divulgación que reducen el sesgo del hábitat en los conjuntos de datos de ciencia ciudadana.

Keywords: Bahamas, ciencia ciudadana, clasificación del hábitat, código abierto, detección remota, eBird, geoespacial

Introduction

Habitat type influences bird biodiversity and bird observations. Bird biodiversity is an important contributor to tourism as well as a sign of ecosystem functioning (Puhakka *et al.* 2011). The Commonwealth of the Bahamas (hereafter referred to as The Bahamas) lies at the beginning of a major migration route for North American species of conservation concern (Lincoln *et al.* 1998; Raffaele *et al.* 1998). Several species also remain in the country over the winter (Raffaele et al. 1998). In the past decade or more, there has been a countrywide push for conservation of natural areas, and *The Forestry Act, 2010* (BHS) (Government of the Bahamas, 2010) charters a path to managing native forests for commercial logging. High levels of endemism and seasonal occurrence of migrant birds make the Bahamas an important biodiversity hotspot for the region (Wege *et al.* 2010).

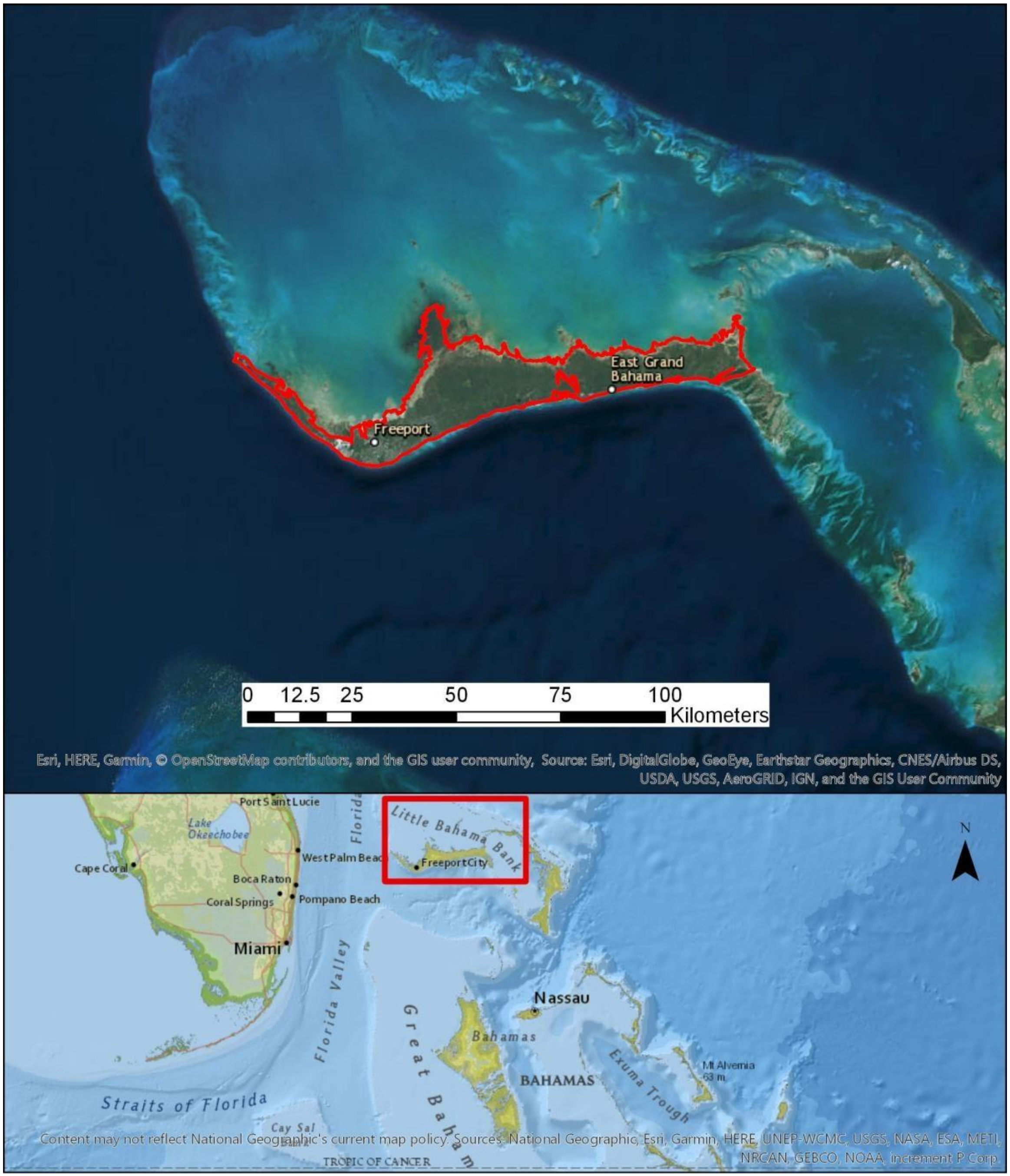
The island of Grand Bahama, the study site, is especially important as a gathering site for migrant birds returning northward (White, 1998). Birders frequent the island and enter data in eBird, an online and mobile platform for citizen scientists to report bird observations (Sullivan *et al.* 2009; https://ebird.org/home). Understanding biodiversity distribution in local forests will be important for successful management required by several national partnerships (Government of the Bahamas 2015). Climate change is an imminent threat for the people and wildlife in the (UNFCCC 2015). It is important to establish current habitat spatial distributions to determine how these future impacts will affect biodiversity and bird-related tourism activity. No current scientific research relating the distribution of pine forest habitat on Grand Bahama to birding and species distribution exists. This work addresses that gap by focusing on the methods to create dynamic maps of species and habitats. Our habitat and biodiversity mapping goals directly supports the targets outlined in the *Bahamas Spatial Data Infrastructure Act, 2014* (BHS) (Government of the Bahamas, 2014) and *Forestry Act, 2010* (BHS) (Government of the Bahamas 2010).

Grand Bahama contains all habitat types present throughout The Bahamas: sandy and rocky shorelines; tidal creeks and wetlands; coppice, mangrove and pine forests; and human dominated landscapes, including built-up areas and agriculture. Grand Bahama also has high bird species diversity recorded in eBird and high diversity of human observers who use the eBird platform to enter data. Official land use and land cover maps for Grand Bahama developed during the 1970s refer to data from as far back as 1950s. These maps provide partial coverage of the island but more than is available for other islands. Grand Bahama also has groundwater and elevation maps derived from in situ data and inference (Government of the Bahamas 1974).

This article explores the relationships among habitats of species occurrence via crowdsourced citizen science birding activity in eBird. Here I examine the relationships of habitat type, habitat distribution, and the spatiotemporal properties of eBird for quantifying species presence and diversity on Grand Bahama. I use open source geospatial data and tools to develop large-scale habitat maps, quantify habitat-species relationships, and present methods that may be applicable throughout the country and the Caribbean region.

Study Area

The study area covers the contiguous terrestrial area of Grand Bahama (26.659447° N, 78.52065° W) (*Figure 1*). Grand Bahama is the nearest Bahamian island to Florida, USA (103 km). The study area extends from West End, Grand Bahama (26.691962° N, 78.999593° W) to McLean’s Town (26.642316° N, 77.934377°W), where the Grand Bahama Highway ends and a permanent creek separates the mainland from the eastern cays. Those cays were excluded, because it is not possible for an observer on the main island to visit those locations without a boat or airplane. The maximum elevation on the island is 25.6 m above sea level with large expanses of shallow wetlands and high water table communities on the north side of the island (U.S. Army Corps of Engineers 2004). White sand beaches rise into low dunes along inland wetlands and native plant communities on the southern edge of the island. Human infrastructure and communities are concentrated west of the city of Freeport, though several towns exist along the southern coast of the island. Large expanses of nonnative grasses are present on golf courses on the island, near schools, and within residential developments, while native grasses and herbaceous growth are prevalent throughout the island where open canopies and exposed soil allow growth. Commercial agriculture is present on the island but does not represent significant land areas. Sandy beaches along the southern coast usually extend a few meters from the Grand Bahama Highway, but sand flats may extend for up to half a kilometer on the northern side.



Esri, HERE, Garmin, © OpenStreetMap contributors, and the GIS user community, Source: Esri, DigitalGlobe, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, USGS, AeroGRID, IGN, and the GIS User Community

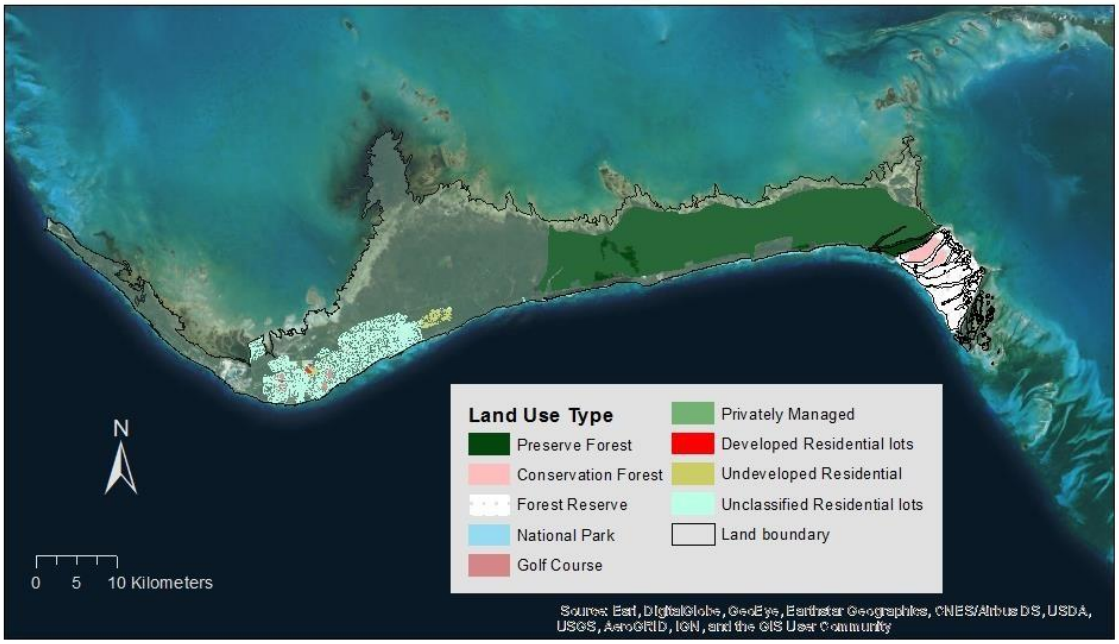
Content may not reflect National Geographic's current map policy. Sources: National Geographic, Esri, Garmin, HERE, UNEP-WCMC, USGS, NASA, ESA, METI, NRCAN, GEBCO, NOAA, increment P Corp.

Figure 2: Study area of Grand Bahama Island

Logging and Habitat Protection

Logging and the establishment of protected areas on each island has also influenced the habitat configuration and extent. On Grand Bahama, logging for Caribbean Pine (*Pinus* *caribaea*) was introduced to the island in the 1970s, and maps of the historic distribution of pine forests on the island can be inferred from the maps of lumber and pulpwood harvest. The broad expanse of time between data collection and map creation makes it difficult to assess accuracy or distribution of habitats at that time. However, the area described as pine forest before the logging activities has generally returned to pine forest after more than 40 years of growth.

After logging, the pine forest was left to regenerate while protected in an established forestry preserve (The East Grand Bahama Protected Area; *Figure 2*) along with conservation forests and preserve forest designated areas. In 2008, the Government of the Bahamas issued the Caribbean Challenge to Caribbean Nations to protect 20% of nearshore marine resources by the year 2020. Much of this protected land falls into that nearshore area and the goal is to have it actively managed to attain full protection. Part of this active management goal includes the establishment of the Department of Forestry within the Ministry of the Environment and Housing. The Bahamas National Trust currently manages the forest preserve and they are actively working to append more area onto the protected forest and expand their personnel to assist in managing the area. Improving knowledge of avian species distribution and habitat use on the island will be useful to support these goals.



Source: Esri, DigitalGlobe, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, USGS, AeroGRID, IGN, and the GIS User Community

Figure 3 Official land use and protected areas on Grand Bahama, Island.

Habitats in this study

The habitat classification for this study included six terrestrial habitat types plus water.

These habitat types were derived from classes used on maps supplied by the Department of Forestry, which reflect habitats used by that agency for classification, management, and protection. These habitat types may also influence accessibility to human observers and bird species use and diversity. The habitats included were water, pine, wetland, sand, urban, grass, and high water table communities (hwtc).

Water in this study includes all areas of naturally occurring water and waterways in direct contact with the ocean. Tidal creeks and channels enter the island from both the north and south while manmade marinas are generally located on the southern side of the island allowing some penetration of ocean water to the center of the island. Though not a terrestrial habitat, the boundaries of water are important in delineating other habitat types and accessibility for human observers (Sullivan *et al.* 2009). Some waterways restrict the passage of vehicles, and therefore reduce human visitation and potentially human observation of birds. Water also complicates remote sensing-based habitat maps for The Bahamas. Specifically, clarity of the ocean water allows light penetration to the sea bed for several meters surrounding the islands of The Bahamas, which can cause overestimation of the extent of satellite-derived terrestrial habitats due to optically shallow waters (Barnes *et al.* 2018).

Pine forests occupy most of the inland area on Grand Bahama. Pine forests or pine barrens in the Bahamas are made primarily of Caribbean Pine (*Pinus caribaea*;Sanchez *et al.* 2014), forming the canopy of this habitat with various other understory plants, vines, and grasses. The pine barrens are only present in the four northern islands of Grand Bahama, Andros, New Providence, and Abaco. The islands were harvested for lumber and pulp beginning in the early 20th century through the 1970's (Sanchez *et al.* 2014). At present, the pines are still recovering and the government is considering long-term sustainable forestry practices to manage the habitat and the native species that use the pines (Government of the Bahamas 2015). The habitat is naturally fire dependent, however, large scale logging, and changes in frequency and intensity of hurricane activity in The Bahamas may have altered the fire regime (Robbins *et al.* 2008). Logging roads have been made throughout most of the pine habitat on Grand Bahama (*Figure 2*) but the habitat was later protected by the Government of The Bahamas (Government of the Bahamas 1987). This is the primary habitat of concern for the island because resident bird species breed and nest there and the Bahamian government intends to develop a sustainable forestry industry for The Bahamas (Macgregor 1996).

Wetlands throughout The Bahamas include permanently wet shallow shorelines of waterways and freshwater and brackish water bodies. These water bodies may include: blue holes and ponds that contain fresh and brackish water; saltwater inlets, estuaries, and creeks; tidal sand flats; and ephemeral wetlands that form after heavy rains or extreme weather events (U.S. Army Corps of Engineers 2004). Mangrove trees typically line the edges of saltwater wetlands (Villiger 2008) and various types of algae may grow on or in the water body on the sediment or substrate. The type of algae influences the overall color of the wetland and those algae in turn can be influenced by the flow of water and salinity (U.S. Army Corps of Engineers 2004; Lirman *et al.* 2007). This study focuses on shallow wetlands with saturated soils and water less than a few feet deep above the surface. Wading birds and birds that probe the muddy substrate use this habitat (Raffaele *et al.* 1998; White 1998). Taller tree species such as pines and native hardwoods are unlikely to be found (Villiger 2008).

Sand and limestone rock appears white or near white because of the calcium carbonate foundation. Limestone rock is temporarily visible until the surface darkens through natural processes after land is cleared. High-energy shorelines also have white sand made of ooliths or pink sand made from the weathered shells of mollusks. Shore birds like laughing gulls and terns may treat large expanses of exposed limestone and areas of sand similarly by congregating there, especially after heavy rains (A.O. Davis pers. obs.).

Urban habitat includes all developed or built areas with homes, businesses, impermeable man-made surfaces, and small manicured lawns. Roads that extend through native vegetation indicate large areas of planned residential development on Grand Bahama. These “cut roads” may enter pine forests, mangroves, and other high water table communities. Marina-style developments may extend out into water bodies or create new waterbodies inland. Areas without any existing infrastructure are not included in the urban habitat type. Planned development areas were classified into the predominant habitat because “cut roads” are often less than one pixelwidth and tree canopies often grow over these narrow roads. The bird species that use urban areas are also distinct, with significant amounts of generalists and invasive species.

Grass lawns are included as a distinct habitat type of grass because of the association with large-scale changes in the native habitat, such as golf courses, hotels, schools, and community parks. These areas have higher walkability and visibility, and may increase the amount of human beings that visit or frequent a site. Lawn maintenance represents high levels of water consumption and may increase water availability for some wild animals, but human presence may also affect species diversity. Bahamian schools often restrict or prevent adult visitation on campus outside of school-sanctioned activities; hotels and golf courses may prevent local residents from accessing the property. Some hotel properties also restrict access to those that work at the location or are paying clients.

High Water Table Communities (hwtcs) include areas where the freshwater comes to or near the soil surface. These areas support plant communities that are more tolerant of high soil moisture, flooding, and salt-water influx. Palms, mangroves, and buttonwood species may be more prevalent. Though pine trees may be visible, they are prone to die-offs after heavy rains flood the area or storm surge drives salt water into the area.

Coppice forests are broadleaf evergreen forests made primarily of native species adapted to the arid climate and the powerful hurricane winds. These forest often have a very dense canopy and shallow soils above limestone rock but may exhibit more mature soils still only a meter deep at most. There is relatively little coppice forest (i.e., dry broadleaf forests with some evergreen species) on the island of Grand Bahama. Carey *et al.* (2014) note that the pinelands of Grand Bahama are managed to prevent a transition to coppice. This habitat type was not included in our classification.

Habitat use of avian focal species from the literature

The 36 focal bird species treated here (*Table 2*) are primarily woodland species and generalists that use multiple habitats. I focus on 20 birds listed as occurring in the Bahamas or Caribbean and the Detroit River Valley by the Detroit Audubon Society (unpubl.) as well as 16 more species that are Bahamian endemic species or known to breed on Grand Bahama Island (Lloyd & Slater 2010, 2011). Historical habitat use for the focal species was gathered from three primary references: Brudenell-Bruce's "The Birds of New Providence and the Bahama Islands" (1975); Anthony White's "A Birder's Guide to The Bahama Islands (Including Turks and Caicos)" (1998); and Raffaele et al's "Birds of the West Indies" (1998). "Beautiful Bahama Birds" (Wardle *et al.* 2014) was also available during this process, but references Raffaele et al. (1998) for the relevant species and did not provide new habitat information. Raffaele et al. (1998) also cite both Brudenell-Bruce (1975) and White (1998). Review of the literature was conducted to confirm which of the focal species occurred on the Island of Grand Bahama and which habitats they occur in.

Citizen science bird data

Bird biodiversity data were initially derived from three crowd-sourced, citizen science, and publicly available datasets for the years of 1988 to 2016: the Audubon Christmas Bird Count (CBC) data (National Audubon Society 2010), the Bird Banding Laboratory (BBL;Smith 2013), and eBird (Sullivan *et al.* 2009). Given the data limitations of the Audubon CBC and the BBL, only eBird was used in the final habitat analysis and citizen science bird reports due to its data density for quantifying species presence and diversity on Grand Bahama island, Geographic specificity of locations, unique Observer IDs, and volume of data.

eBird Basic Dataset (EBD; https://ebird.org/science/download-ebird-data-products) was accessed via its web portal. Bird observations for The Bahamas from 24 April 1901 to 30 November 2017 were retrieved. data wereThe data were imported and analyzed using the programming language R (Venables & Smith 1990; R Core Team 2018) version 3.3.2 64-bit employed via RStudio version 1.1.442. The EBD data were geographically clipped to Grand

Bahama and the contiguous terrestrial area and a temporal subset performed for the dates of 1 January 1988 to 31 December 2016. The R code used to process the bird and observation data, as well as to assign and analyze habitat type, for this analysis can be found at https://github.com/Ancilleno.

Habitat descriptions in EBD data

Along with location data in eBird EBD products, observers have the option of adding additional details regarding the site, the species seen, and other observers. Occasionally, eBird users will include significant detail about the site, but this information is not standardized. eBird data allows the option to download land cover data as determined from the Global Land Cover

Facility (http://www.landcover.org/). However, the resolution, 250 m pixels, is too coarse for The Bahamas and does not reflect changes in land cover across the landscape.

Satellite-derived habitat map

The satellite-based habitat classification was performed using Google Earth Engine (GEE; Gorelick *et al.* 2017). The Random Forest classification model (Breiman 2001), an ensemble approach for producing and merging many decision trees for accurate and stable predictions, available in GEE was used. Only cloud-free, 30 m Landsat 8 Operational Land Imager (OLI) images from 1 January 2017 to 31 December 2017 were considered. A total of 88 images were composited together across four Landsat 8 Path/Row coordinates to create a 2017 Habitat Classification Map of Grand Bahama: P13/R41 (20 scenes), P13/R42 (22 scenes), P14/R41(21 scenes), and P14/R42 (21 scenes). Six reflectance bands were included in the classification, specifically B2 (blue spectra on the electromagenetic spectrum), B3 (green), B4 (red), B5 (near infrared), B6 (shortwave infrared), and B7 (shortwave infrared). Bands 10 and 11 were not included due to coarse resolution.

Training and validation data for the Random Forest classification was created from manual delineation (i.e., "heads-up" digitization) of commercial satellite data by a regional expert: A.O. Davis, a Bahamian citizen, with 10+ years work experience in conservation and environmental education with non-profit organizations and government agencies throughout the islands. Ancillary data, used to augment to visual delineation, was assembled from two primary sources: print habitat maps scanned to digital files provided by The Bahamas Government and ESRI shapefiles of land use type and residential property delineations provided by the Forestry Department of The Bahamas. The Department of Lands and Surveys provided Land Capability Maps, outlining frequently flooded areas which can be described as wetlands (Bahamas Government - DLS 1974.) Forest Map No. 03/87 shows the borders and protection levels of forests on Grand Bahama, such as "Forest Reserved", "Protected Forest", and "Forest Conservation" as well as the forest type and regeneration maps (Bahamas Government - Forestry 1973). These maps indicated type of forest and degree of regeneration post-fire at the location, with notes on density of pine, and how well stocked the stands were for commercial harvest. The Grand Bahama Groundwater Land Ownership maps (Bahamas Government - Forestry 1974) were used to confirm high water table communities. A minimum of 100 point locations for each habitat type were identified by A.O. Davis using very high resolution (VHR) natural color satellite imagery from ~ 1.5 m CNES/AirBus data. Each habitat type was given a class label and a unique classified pixel value from zero to six (*Table 1*). Finally, visually classified habitat point locations were then split into training and validation data at a 7:3 ratio, in order to prevent spatially autocorrelated results (Millard & Richardson 2015) within GEE Random Forest classification. The resulting GeoTIFF habitat classification map was imported into R for analysis.

Methods

Assigning habitat type to eBird observations

Habitat type at the eBird survey locations (i.e., latitude/longitude) were extracted from the satellite-derived habitat map in R using the extract function in the "raster" package (Hijmans *et al.* 2017). Locations outside of the contiguous terrestrial area returned "NA" for habitat. They were then excluded from further analysis.

Calculating differences in diversity

R was used to calculate observer and species richness, and Shannon's diversity based on the number of citizen science surveys at locations within each habitat. Overall observer and species richness per habitat was compared using Jaccard's dissimilarity index (Goodall 1966) and Bray-Curtis dissimilarity indices (Beals, 1984), used here to compute the shared overlap of bird species and observers (presence/absence) across habitats (Magurran 2004). The Bray-Curtis index is semimetric while the Jaccard index is metric, but they are rank order similar and produced similar ordination plots. The Bray-Curtis ordination plot was chosen for its tighter configuration though the overall relationship between groups did not change. Shannon's diversity was calculated using the diversity function in the R package "vegan" (Oksanen *et al.* 2018). The Jaccard index was calculated using the clujaccard function in the "fpc" package (Hennig & Hennig 2018).

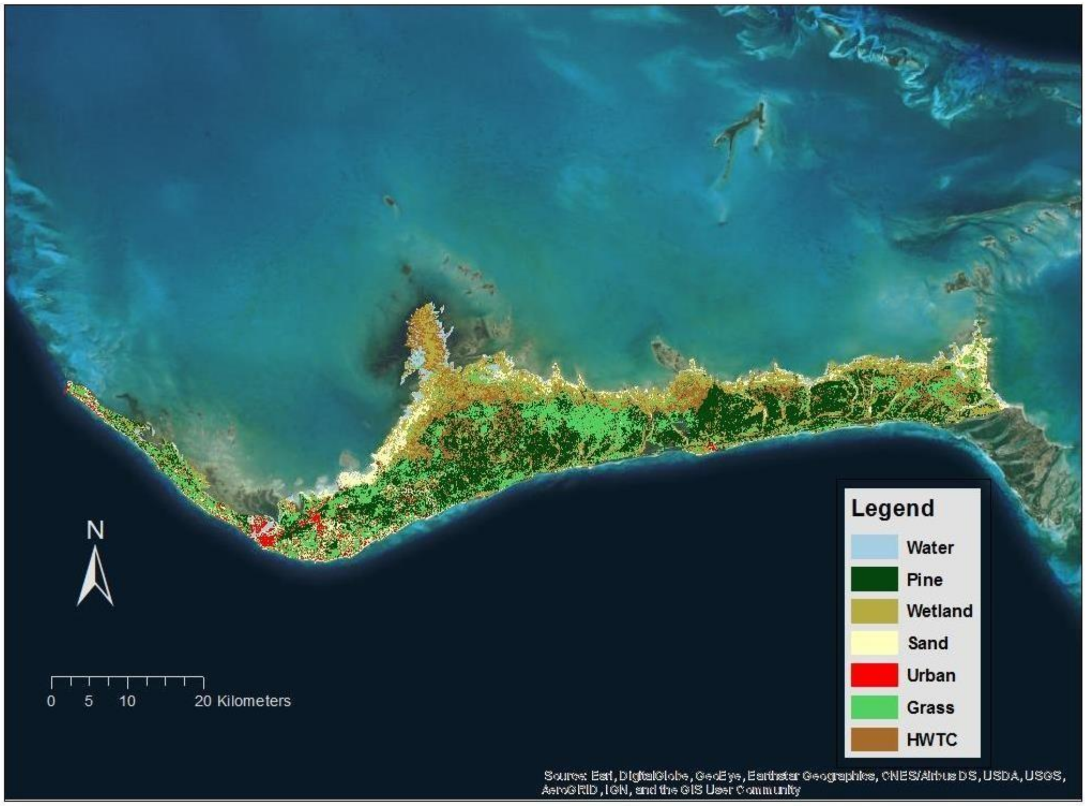
Results

Literature-based habitat use of focal species

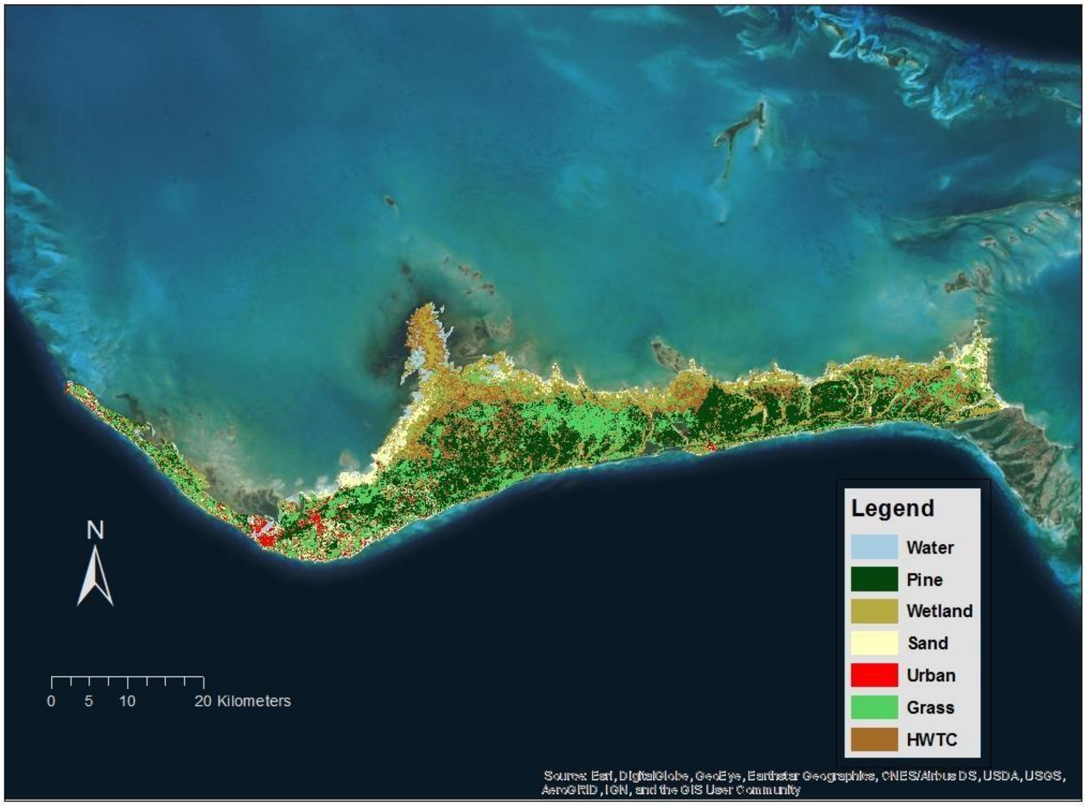
Aggregated habitat use tables (*Table 3*, *Table 4*) were compiled for the focal species based on available literature (Brudenell-Bruce 1975; Raffaele *et al.* 1998; White 1998). These synthesis tables were separated into two tables to indicate Detroit Audubon Society (DAS) species of international concern (*Table 3*) and Bahamian species considered to be resident (*Table 4*). Habitat use differed between the DAS and Bahamian species. Most Bahamian species were reported to use the pine forest habitat with all habitats being used by at least one species (*Table 4*). Most DAS species were found to utilize the wetland habitat (*Table 3*). The Common Ground-Dove (*Columbina passerina*) is the only species described as using all habitat types. Coppice and agriculture habitats are reported because they are referenced in the literature, however, these habitats were not distinguished in the satellite-derived habitat classification; coppice forests were combined with pine classes and agriculture was combined with grass class.

Satellite-derived habitat map

On Grand Bahama, the majority habitat classes are pine forest (11.25 km^2 or 28.51% of total land cover) and wetlands (11.25 km^2 = 21.95%). Pine forests dominate land cover on the eastern side of the island and hwtc/wetlands and human-altered landscapes dominate the west (*Figure 3*, *Figure 4*). The Random Forest habitat classification model was unable to distinguish between residential buildings with large grassy lawns, other grassy areas such as athletic fields, and golf courses or areas with invasive grasses.

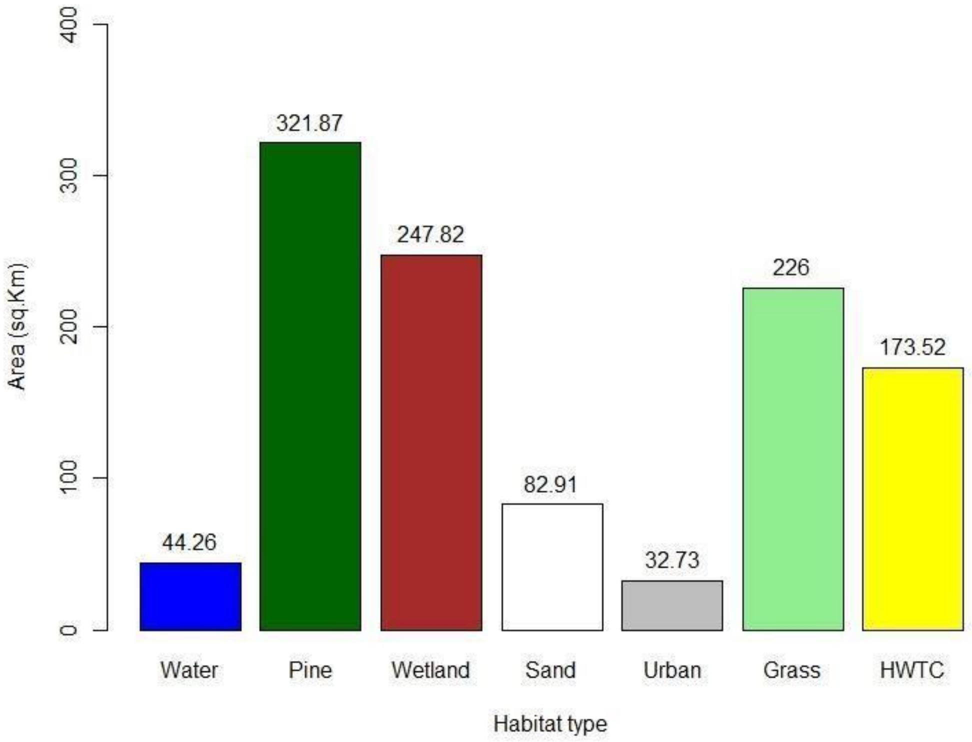


Source: Esri, DigitalGlobe, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, USGS, AeroGRID, IGN, and the GIS User Community



Source: Esri, DigitalGlobe, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, USGS, AeroGRID, IGN, and the GIS User Community

Figure 4. 30 m resolution Landsat 8 classified habitat map of Grand Bahama, The Bahamas showing 6 terrestrial habitat types and water superimposed on satellite image.


Area (sq.Km)
0
100
200
300
400
44.26
321.87
247.82
82.91
32.73
226
173.52
Water
Pine
Wetland
Sand
Urban
Grass
HWTC
Habitat type

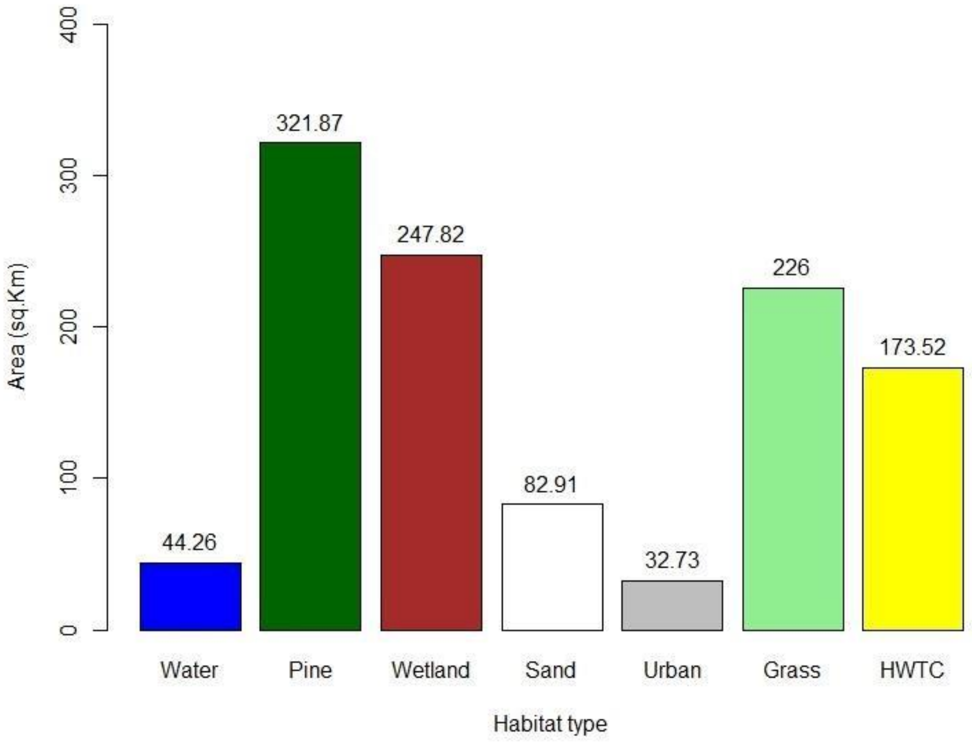


Figure 5.Figure 4. Terrestrial area of each classified habitat type on Grand Bahama; area of each class reported by numbers at top of bars (km2), based on classified habitat map using cloud-cleared Landsat 8 imagery from 2016.

GEE was also used to generate a confusion matrix with user created validation data (Foody, 2002; *Table 5*). The Cohen's Kappa test found "substantial agreement" between the habitat classification and validation data, with a Kappa coefficient of 0.78. The overall accuracy of the habitat map using the Random Forest classification method in GEE was 81.43% (Kappa = 0.78 at 95% Confidence Interval; *Table 5*). The code to complete the habitat classification of 2017 Landsat data of Grand Bahama can be found in Appendix I. Optimization of pine forest habitat classification is necessary to compare with previous studies focused on the pine habitat. Therefore, the accuracy of the wetland classification (user's accuracy of ~ 49%) is not as pertinent for the overall analysis as the user's accuracy of other classes. With a user's accuracy of ~ 76% and a producer's accuracy of 100%, the pine forest class was adequately mapped in our approach.

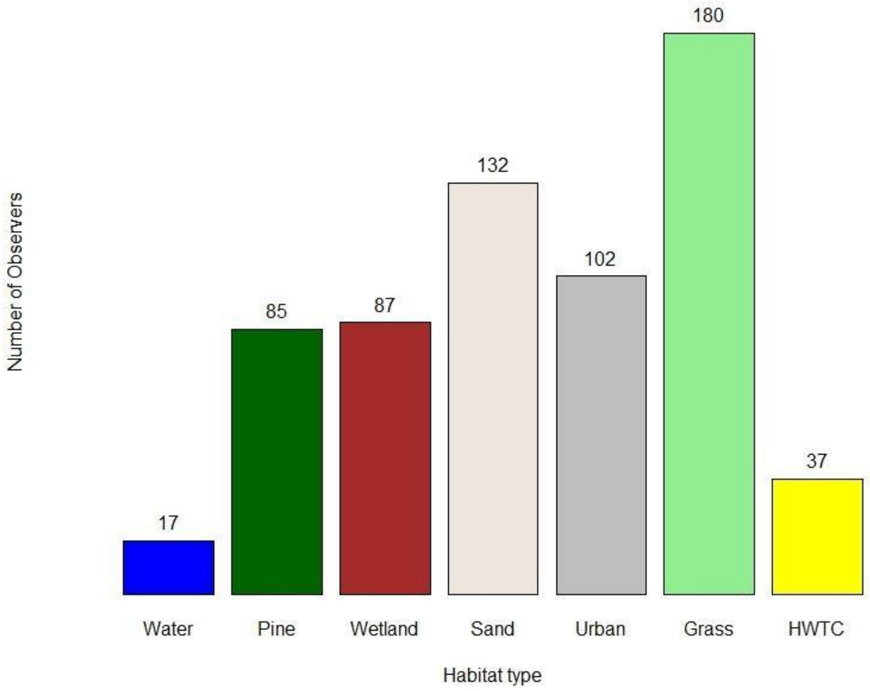
Citizen science bird data

On Grand Bahama, 4,915 surveys were conducted at 660 locations by 346 different observers between 1 January 1988 and 31 December 2016. Nearly one third of all surveys

(31%), one quarter of all observers (25%), and one fifth of all locations surveyed (19%) and recorded in eBird for The Bahamas were on Grand Bahama. The majority of the species recorded in The Bahamas (75%, n = 302) were recorded on Grand Bahama between 1988 and 2016. Grand Bahama also had the second highest recorded cumulative bird species and observer diversity in eBird in The Bahamas for 1980-2016 and the highest number of surveys conducted. Observers attempted to identify all birds seen in 4,242 of the surveys conducted on Grand Bahama. There were significant biases toward early morning hours and winter months, but the eBird data includes records for every month of the year and at various times of day. eBird records on Grand Bahama include 34 of the focal species listed in *Table 3* and *Table 4*.

Canvasback (*Aythya valisineria*) have not been recorded in The Bahamas and is of "Least Concern" according to the IUCN Red List assessment by BirdLife International (BirdLife International 2016). The Kirtland's Warbler (*Setophaga kirtlandii*) is critically endangered (BirdLife International 2012) and has been recorded on Grand Bahama, but not during a survey where all species were recorded.

Wetland and pine habitats are the dominant classes in total land area, followed by grass, sand, hwtc, and lastly, urban areas (*Figure 4*) based on the satellite-derived map. However, within eBird, the grass and sand classes are the most commonly surveyed, followed by urban, then wetlands and pine, with hwtc being the least visited and surveyed (*Figure 5* & *Figure 6*).


Number of Observers
17
85
87
132
102
180
37
Water
Pine
Wetland
Sand
Urban
Grass
HWTC
Habitat type

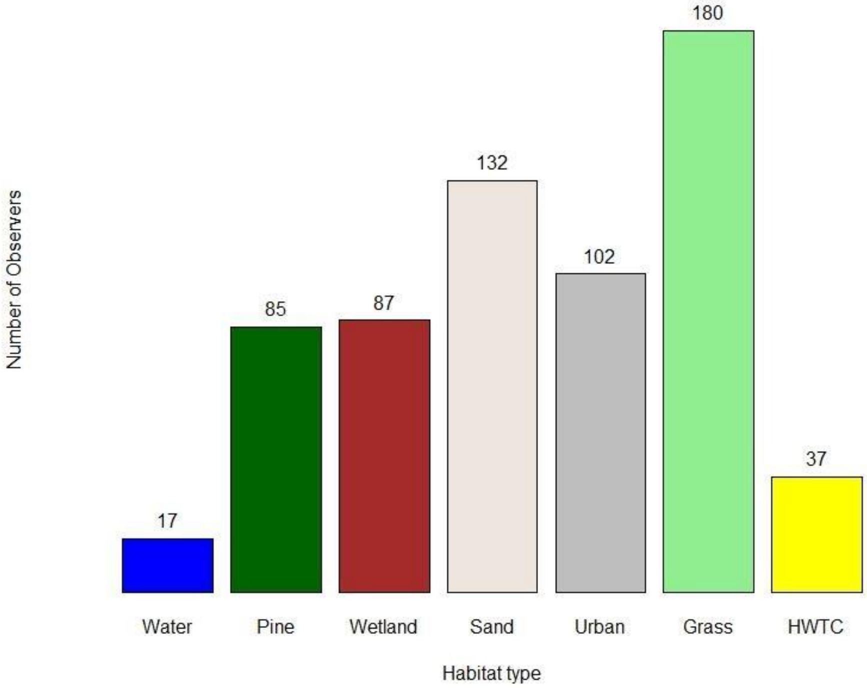


Figure 6. Total unique observers that reported survey data in eBird on Grand Bahama between 1 January 1988 and 31 December 2016 in 6 terrestrial habitats inland water bodies. Habitat classification based on cloud-cleared Landsat satellite images collected during 2016.

In total, 283 unique observers participated in surveys on Grand Bahama between 1 January 1988 and 31 December 2016. The observers visited at least two habitat types, on average. The number of visitors that participated in surveys within each habitat type varied among habitats. Grass habitats were visited by the most unique observers (180; 64%) that conducted surveys within that habitat and attempted to identify and report all species seen followed by sand (132; 47%). Water and hwtc communities had the fewest observers (17 and 37 or 6% and 13%, respectively). About one third of observers conducted surveys in each of the pine, wetland, and urban habitats (30%, 31%, and 36%, respectively; *Figure 6*). hwtc, wetland, and water had the fewest unique localities (n=18, 28, and 44, respectively) while sand (n=107) and grass habitats (n=100) have the most localities. Pine and urban habitats held similar numbers of survey locations (pine = 74, urban = 79). Most observers that collect data on all species detected also visit multiple locations and habitats (*Figure 6*).

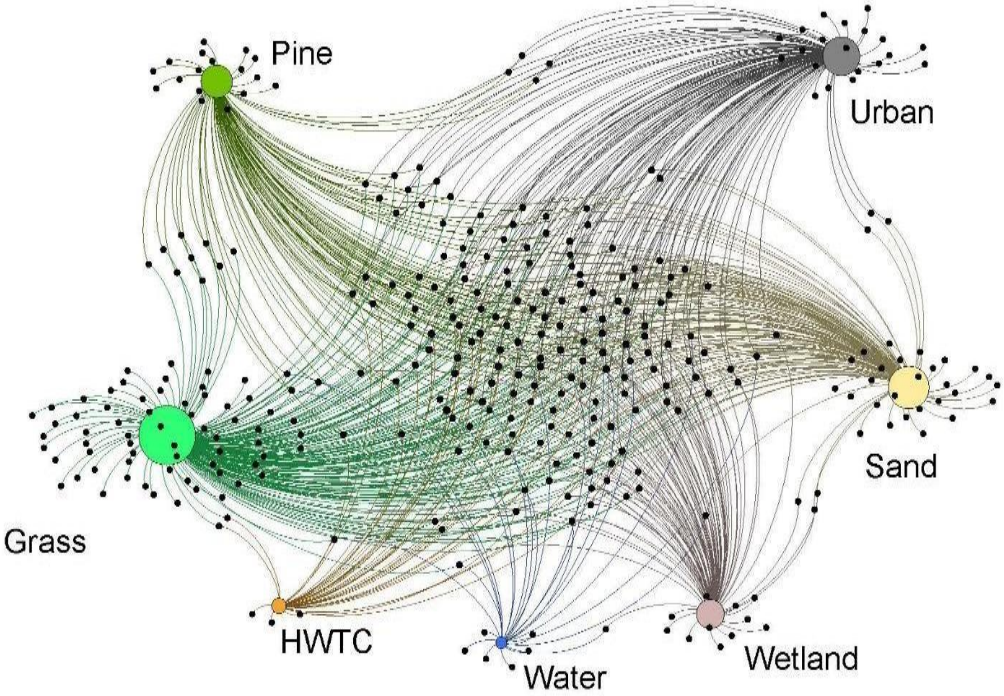


Figure 7. Network map of observer habitat use on Grand Bahama. Node size indicates relative proportion of connections.

Grass habitats account for 39.5% of all surveys conducted followed by sand (31%). The remaining 29.5% of surveys were conducted in urban, wetland, pine, hwtc, or water habitats (*Figure 7*). Furthermore, the average time spent surveying at locations within each habitat differed significantly between and among habitats.

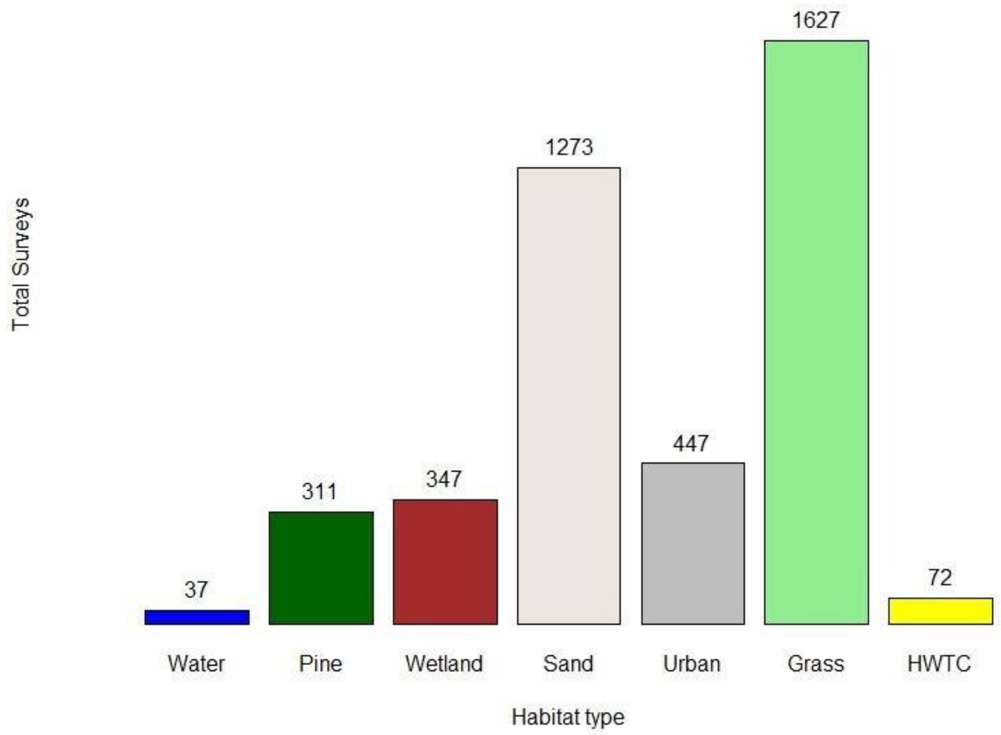


Figure 8. Total surveys entered in eBird for Grand Bahama between 1 January 1988 and 31 December 2016, by habitat.

In all habitats both BAH and DAS focal species were detected (*Table 6*). There was no significant difference in the proportion of surveys with DAS and BAH species within water, urban, grass, or hwtc habitats. However, a significantly higher proportion ($p=0.0021$) of pine surveys reported BAH species compared to those that reported DAS species. The opposite was true for wetland and sand habitats, where a significantly higher proportion of surveys reported DAS species when compared to BAH species ($p<0.0001$).

Species communities had little difference between habitat pairs, though the greatest differences were seen between water and wetland or sand and grass communities (Table *7*). These differences are greater than 0.5 and indicate that the difference in community members is greater than random. hwtc also differed from grass and sand. Urban and pine communities did not differ greatly from any of the other communities (JD<0.43). The Jaccard distance index shows that the human observer communities in wetland, sand, urban, and grass communities are most similar while human observers in the hwtc, water, and pine communities are more distinct. For all habitat pairs, the Jaccard distance was greater than 0.5 indicating that the two habitats have less than half of their observers in common.

Habitat use tables (*Table 8* and *Table 9*) and network maps (*Figure 8* and *Figure 9*) were generated for both the DAS and BAH species groups. Observers reported all species in habitats other than the habitats indicated in the literature. All species except the Sharp-shinned Hawk (*Accipiter striatus*) and Common Tern (*Sterna hirundo*) were detected in the habitats they are known to occur from the literature, but the number of records for these two species was extremely small (Sharp-shinned Hawk n=3; Common Tern n=5; *Table 9*).

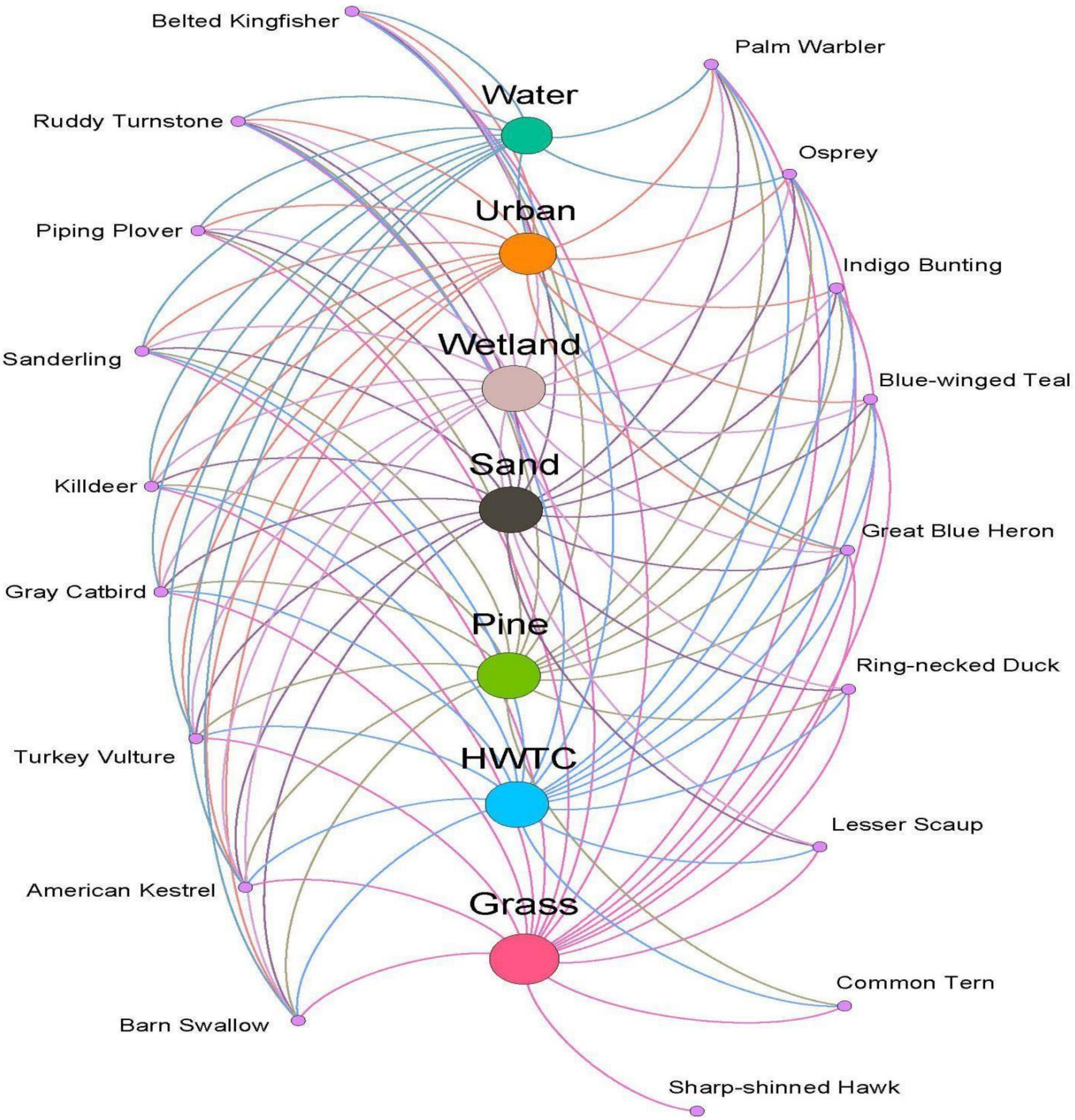


Figure 9. Network map of DAS species and the habitats in which they were reported.

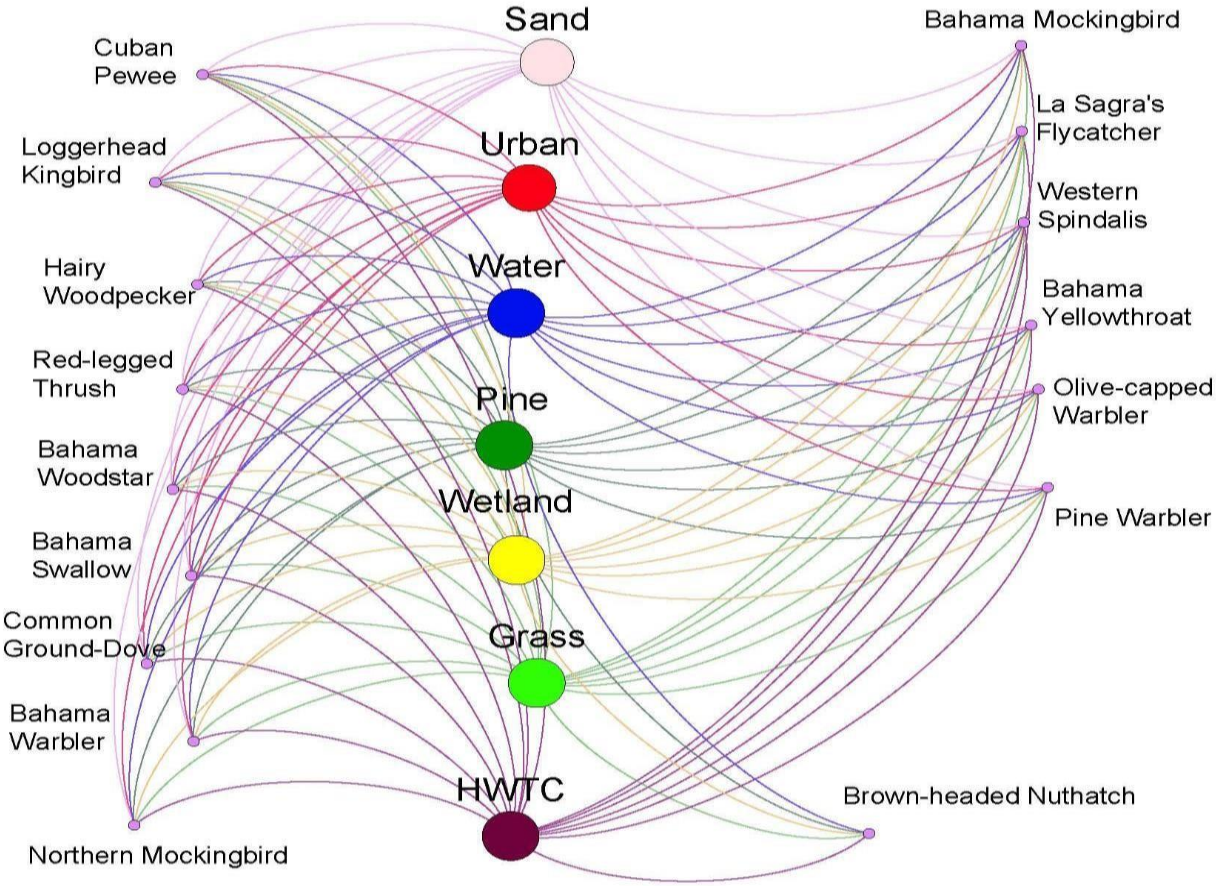


Figure 10. Network map of BAH species and the habitats in which they were reported.

Discussion

Habitat classification accuracy and potential improvements

Habitat map accuracy depends on accurate knowledge of the habitat in situ, spatiotemporal resolution of the satellite data, environmental factors such as cloud cover, season, and incidence of the sunlight on the terrain. For example, Landsat 8 OLI bands 1 and 9 were not included due to potential for strong atmospheric absorption (Shelestov *et al.* 2017), though Band 1 is coastal aerosol and may be useful for delineation of coastal habitats. Bands 10 and 11 can be used to derive soil moisture but at a coarser resolution are not effective in this small island study area. The habitat map generated in GEE was accurate and qualitative assessment coincided with expectations. The authors recommend this methodology for rapid habitat classification, given limitations, because it is low cost, adaptable to other habitat classes or locations, and easily replicated and updated using open source data and code.

Using the satellite-derived habitat map for conservation and management requires recognizing the errors along habitat edges and dynamic land covers. Habitat types along moisture gradients may be more difficult to distinguish or easily confused in satellite data. Generally, these maps are useful for approximating the amount of each habitat in an area and planning for conservation or management, but less useful in drawing boundaries between habitats. Management agencies can use these broader habitat outlines to identify boundaries between habitats. Those boundaries should be assessed, in situ, to determine habitat boundaries, buffers and gradients between adjacent habitats.

Wetland and hwtc community boundaries may shift with rainfall or prolonged drought, explaining the lower accuracy of wetlands. Wetland areas in the Landsat-based habitat map accurately depicted the former path of creeks and estuaries, disconnected from the ocean by roads along the shoreline but remain connected via the water table. These wetland areas are visible on older maps but not easily detected in the very high resolution images in GEE. Habitat classifications within large contiguous areas such as pine forests, sand banks, and “grass” were more accurate. Narrow bands of beach or grassy areas within urban limits and along coasts had higher levels of confusion. Grass berms along roads within pine forests or near beach roads were not visible in the habitat map. Many beaches along the southern coast did not show up on the map. Roads less than one pixel width were generally not visible on the map except at a few intersections that coincided with the pixel at the point.

The most feasible improvements to the current satellite-based habitat model include adding data points and revising the habitat classification to represent the biological or environmental diversity expressed by stakeholders. Higher resolution data (< 30 m) would improve visual classification and model accuracy. Accounting for seasonal variation between wet and dry seasons or across years may help to determine if habitat change is important to the reported species.

Literature vs. eBird-based habitat use

Each focal species had records within eBird that agreed with their published habitat use preferences (except for Common Terns [n = 5] and Sharp-Shinned Hawks [n = 3 surveys]). All species also had habitat records outside their published habitat preferences. These results indicate an important gap in our understanding of habitat use for native and migrant species on Bahamian islands and need for landscape conservation beyond the habitats reported in the literature.

Bahamian endemic species such as the Bahama Woodstar Hummingbird may require significant study to determine their relationship with sand and grass habitats which account for more than 75% of the records in eBird, though they are only mentioned in urban and pine habitat in the literature. Species like La Sagra’s Flycatchers, Northern Mockingbirds, and Red-legged Thrushes are more commonly reported outside of their published habitat. The Great Blue Heron, which is described as being a wetland species, but is more often reported in sand and grass habitats. Not only does this create a need for different habitat conservation goals to protect these species, but also implies a need to protect other associated prey species. Great Blue Herons feed on land crabs, lizards, and insects on sand and grassy habitat as opposed to fish and amphibians in freshwater wetlands in the continental US.

Precision, accuracy, and resolution in citizen science data

This research relies on accurate and precise species and location descriptions. When an observer consistently identifies a species inaccurately, this inflates reports of wrong species and underrepresents the true species. Precision is the ability to consistently identify a species. Some observers classify multiple species as a single species or a single species as different species depending on context or behavior. Bahama Mockingbirds and Northern Mockingbirds, for example, may be mistaken for one another by novice birders and it is impossible to detect these errors in the eBird data. Some species names change over time by American Ornithological Union standards. eBird does well in this regard as species classifications are automatically updated with changes that apply to the entire database.

Location accuracy and specificity in eBird, when lacking, creates a challenge for habitat classification of ebird observations. Many observations are recorded as being “On Grand Bahama” or “In Freeport”. These data points are useful for the production of species lists and potentially for island wide conservation or monitoring activities. They are less useful than geographically specific observations with geographic coordinates. Better integration with open street-mapping data sets and engagement with local people to identify described locations on maps could significantly improve location accuracy and conservation potential. Accurate location data improves our knowledge of species communities within habitats and site or location popularity for conducting surveys. Managers and educators can use this information to improve survey representativeness. Observers can also be encouraged to identify habitat into broad categories to improve our understanding of habitat distribution and bird distributions in those habitats.

Potential drawbacks from and improvements to citizen science efforts

eBird includes locations not reported in scientific literature or other databases and are unlikely to have been formally surveyed. The primary drawback with eBird is the lack of standardization across observers. Observers may have very different levels of experience in bird identification, enthusiasm for birding or data entry, and do not report all observations. Periodic quizzes on birds recently entered by the observer or regional identification quizzes or certification can establish a birders credibility in identification and be used to filter novice and expert observations within eBird.

The eBird data shows a consistent underrepresentation of pine forests, wetlands, and

hwtc in survey data for the island of Grand Bahama and differences in monthly effort between habitats as well as for the island in general. Frequent analysis of citizen science bias and targeted engagement of the birding community to reduce bias and fill knowledge gaps can improve data quality and usefulness for conservation and species protection or management. In particular, the engagement of local birders to visit sites that tourists are less likely to visit or to conduct surveys at times of the year when tourists are not around could greatly improve our knowledge of the habitat use of birds and affect conservation and management activities.

Impact on logging/forest management

Several species recorded on Grand Bahama are of significant conservation importance. Woodpecker species were significantly impacted by the removal of pine forests and have since been rarely recorded on the island. The endemic Grand Bahama Nuthatch also relies on pine forests for breeding according to the literature and is confirmed by data presented here. As The Bahamas develops forestry management practices, using citizen science to advise formal study or determine conservation and monitoring targets is an important first step.

Future Work

Satellite-derived habitat mapping has great promise for improving the knowledge of Caribbean biodiversity, when informed by local stakeholder needs or values. Network maps should be generated to indicate how species use land under different management regimes as well as habitat type. This methodology can be used to rapidly generate habitat maps for all the islands of The Bahamas and Caribbean. Most of these islands have eBird surveys, and in combination with the habitat data, these surveys can advise conservation and management decisions both at national and regional levels. These novel habitat associations are fertile ground for future research on the individual species or habitats as well as the bird communities as a whole and how they partition these habitats.

This low-cost habitat mapping method is feasible for small nations, local NGOs, and Community Based Organizations. The cost savings come primarily from using freely available software and data. Using open-source multi-platform and web-based applications removes the need for specific hardware, software subscriptions, or significant formal or proprietary education or training. The maps presented here were generated quickly (approximately two weeks). This offers saving on personnel time. Rapid map generation for conservation and management provides opportunity loss prevention by having needed maps generated more quickly to serve reporting requirements for funding agencies. Most importantly, the code created to generate the map is also freely available and can be adapted for other locations or habitat classification needs.

Open source software and web-based applications also make products and processes readily available and transparent to collaborators.

Acknowledgments

We would like to thank the staff of the Government of the Bahamas' Department of Forestry, especially Danielle Hanek, and Christopher Russell for supplying map data for Grand Bahama Island. In addition, I thank Miami University's Graduate School, Department of Biology, and Department of Geography for funding and computing resources, Miami University graduate students Lauren Adkins and Tony Francis, Miami University undergraduate student

Justin Fain for R and python coding support, and the eBird group and all the observers that contributed data. Special thanks to Alma, Leo, and Oliver Davis for their support during the writing process.

Tables

Table 1. Habitat classes and basic descriptions for the satellite-derived habitat map of Grand Bahama.

Habitat Class	Land cover/ Pixel Value	Description
Water	0	Permanently submerged marine areas
Pine	1	Predominantly Caribbean Pine (*Pinus caribbaea*), includes coppice forests
Wetland	2	Waterlogged soils/ephemeral wetlands not connected to open ocean
Sand	3	Limestone sand
Urban	4	Manmade structures and roadways
Grass	5	Manicured lawns and nonnative grasses, includes small agricultural areas
hwtc	6	Plant communities dominated by palms and moisture tolerant species

Table 2. Focal species common names, four letter codes and scientific names

Common Name	Code	Scientific name
American Kestrel	AMKE	*Falco sparverius*
Barn Swallow	BARS	*Hirundo rustica*
Belted Kingfisher	BEKI	*Megaceryle alcyon*
Blue-winged Teal	BWTE	*Spatula discors*
Canvasback*	CANV	*Aythya valisineria*
Common Tern	COTE	*Sterna hirundo*
Gray Catbird	GRCA	*Dumetella carolinensis*
Great Blue Heron	GBHE	*Ardea Herodias*
Indigo Bunting	INBU	*Passerina cyanea*
Killdeer	KILL	*Charadrius vociferus*
Kirtland's Warbler*	KIWA	*Setophaga kirtlandii*
Lesser Scaup	LESC	*Aythya affinis*
Osprey	OSPR	*Pandion haliaetus*
Palm Warbler	PAWA	*Dendroica palmarum*
Piping Plover	PIPL	*Charadrius melodus*
Ring-necked Duck	RNDU	*Aythya collaris*
Ruddy Turnstone	RUTU	*Arenaria interpres*
Sanderling	SAND	*Calidris alba*
Sharp-shinned hawk	SSHA	*Accipiter striatus*
Turkey Vulture	TUVU	*Cathartes aura*
Bahama Mockingbird	BAMO	*Mimus gundlachii*
Bahama Swallow	BAHS	*Tachycineta cyaneoviridis*
Bahama Warbler	BAHW	*Setophaga flavescens*
Bahama Woodstar	BAWO	*Calliphlox evelynae*
Bahama Yellowthroat	BAYE	*Geothlypis rotrata*
Brown-headed Nuthatch[c]	BHNU	*Sitta pusilla*
Common Ground-Dove	COGD	*Columbina passerina*
Cuban Pewee	CUPE	*Contopus caribaeus*
Hairy Woodpecker	HAWO	*Picoides villosus*
La Sagra's Flycatcher	LSFL	*Myiarchus sagrae*
Loggerhead Kingbird	LOKI	*Tyrannus caudifasciatus*
Northern Mockingbird	NOMO	*Mimus polyglottos*
Olive-capped Warbler	OCAW	*Setophaga pityophila*
Pine Warbler	PIWA	*Dendroica pinus*
Red-legged Thrush	RLTH	*Turdus plumbeus*
Western Spindalis	WESP	*Spindalis zena*

Table 3. Habitat preference for Detroit Audubon Society species of international concern; X's indicate published habitat preference in commercially available field guides.

Common Name	Code	Urban	Pine	Coppice	Wetland	Beach	Agriculture
American Kestrel	AMKE	X					X
Barn Swallow	BARS	X	X		X	X	X
Belted Kingfisher	BEKI				X		
Blue-winged Teal	BWTE				X		
Canvasback*	CANV				X		
Common Tern	COTE				X	X	
Gray Catbird	GRCA			X			X
Great Blue Heron	GBHE				X		
Indigo Bunting	INBU			X			X
Killdeer	KILL	X		X	X		X
Kirtland's Warbler*	KIWA		X	X			
Lesser Scaup	LESC				X	X	
Osprey	OSPR				X	X	
Palm Warbler	PAWA		X	X	X	X	X
Piping Plover	PIPL				X	X	
Ring-necked Duck	RNDU				X	X	
Ruddy Turnstone	RUTU					X	
Sanderling	SAND					X	
Sharp-shinned hawk	SSHA		X				
Turkey Vulture	TUVU	X	X	X			X

Table 4. Habitat preference for Detroit Audubon Society species of international concern; X's indicate published habitat preference in commercially available field guides.

Common Name	Code	Urban	Pine	Coppice	Wetland	Beach	Agriculture
Bahama Mockingbird	BAMO	X		X			
Bahama Swallow	BAHS	X	X	X		X	X
Bahama Warbler	BAHW		X				
Bahama Woodstar	BAWO	X	X	X			
Bahama Yellowthroat	BAYE		X	X			
Brown-headed Nuthatch[c]	BHNU		X				
Common Ground-Dove	COGD	X	X	X	X	X	X
Cuban Pewee	CUPE		X	X			X
Hairy Woodpecker	HAWO	X	X	X			
La Sagra's Flycatcher	LSFL		X	X	X		
Loggerhead Kingbird	LOKI		X	X	X		X
Northern Mockingbird	NOMO	X			X		X
Olive-capped Warbler	OCAW		X				
Pine Warbler	PIWA		X				
Red-legged Thrush	RLTH	X	X	X			
Western Spindalis	WESP		X	X			

Table 5. Confusion Matrix and accuracy assessment for Random Forest Habitat Classification via Google Earth Engine.

		Reference Data								
		Water	Pine	Wetland	Sand	Urban	Grass	hwtc	Total	User's Accuracy (%)
Classified Data	Water	33	0	2	1	0	0	0	36	100%
	Pine	0	29	0	0	0	0	0	29	76%
	Wetland	0	3	22	0	0	0	5	30	49%
	Sand	0	0	2	38	1	0	0	41	84%
	Urban	0	0	0	6	21	1	1	29	88%
	Grass	0	2	7	0	2	57	1	69	95%
	hwtc	0	4	12	0	0	2	28	46	80%
	Total	33	38	45	45	24	60	35	280	
	Producer's Accuracy	92%	100%	73%	93%	72%	83%	61%		Total Accuracy 81%
	Kappa = 0.78 at 95% confidence interval									

Table 6. Summary of eBird localites, observers, surveys, and BAH and DAS species by satellite-derived habitat type in eBird records collected between 1 January 1988 and 31 December 2016 in which all detected species were recorded. NB: Some observers conducted surveys at multiple locations (and habitats) over time. Some species were reported at multiple locations and habitats

Habitat type	Water	Pine	Wetland	Sand	Urban	Grass	hwtc	Total
Total km2	44.63	321.96	247.92	82.53	32.71	225.98	173.59	1129.33
Localities	31	78	49	120	92	111	18	499
Observers	21	95	92	142	115	199	37	280
Surveys	41	325	358	1303	472	1671	72	4242
BAH Species	16	16	16	15	15	16	16	16
DAS Species	12	16	16	16	14	18	16	20
Total Species	108	155	176	217	163	247	113	278
BAH Observers	14	106	81	124	118	221	35	313
DAS Observers	18	100	94	136	107	194	35	280
Total Observers	26	117	101	160	143	233	41	339

Table 7. Jaccard dissimilarity for Species in habitats on Grand Bahama Island recorded in eBird, 1988-2016.

	Water	Pine	Wetland	Sand	Urban	Grass	hwtc
Water	0	0.406	0.505	0.556	0.424	0.574	0.433
Pine		0	0.320	0.347	0.233	0.392	0.366
Wetland			0	0.299	0.313	0.328	0.438
Sand				0	0.326	0.269	0.520
Urban					0	0.367	0.414
Grass						0	0.548
hwtc							0

Table 8. Percentage of eBird surveys for each Bahamian focal species in the satellite-derived habitat numbers in bold indicate published habitat preference; n = total number of surveys with said species; note coppice forest and agriculture classifications are not used.

Common Name	Water	Pine	Wetland	Sand	Urban	Grass	hwtc	n
Bahama Mockingbird	4%	35%	12%	8%	0%	38%	4%	26
Bahama Swallow	4%	31%	5%	14%	9%	34%	4%	140
Bahama Warbler	1%	6%	4%	2%	4%	24%	5%	103
Bahama Woodstar	1%	1%	6%	24%	1%	55%	1%	157
Bahama Yellowthroat	3%	47%	17%	9%	1%	20%	4%	101
Brown-headed Nuthatch^c	3%	70%	3%	0%	0%	18%	6%	33
Common Ground-Dove	1%	7%	7%	29%	12%	42%	1%	1436
Cuban Pewee	2%	36%	8%	7%	7%	36%	3%	96
Hairy Woodpecker	0%	13%	11%	17%	8%	49%	1%	663
La Sagra's Flycatcher	1%	7%	8%	12%	10%	61%	1%	687
Loggerhead Kingbird	1%	6%	7%	20%	9%	56%	0%	524
Northern Mockingbird	0%	6%	9%	30%	11%	42%	2%	2650
Olive-capped Warbler	2%	58%	6%	1%	>1%	26%	7%	145
Pine Warbler	2%	56%	5%	3%	2%	27%	5%	274
Red-legged Thrush	0%	7%	7%	22%	10%	53%	1%	1268
Western Spindalis	1%	27%	5%	3%	6%	56%	2%	448

Table 9. Percentage of eBird surveys for each DAS focal species in the satellite-derived habitat. Numbers in bold indicate published habitat preference; n = total number of surveys with said species; note coppice forest and agriculture classifications are not used

Common Name	Water	Pine	Wetland	Sand	Urban	Grass	hwtc	n
American Kestrel	0%	4%	8%	33%	60%	48%	1%	1233
Barn Swallow	0%	5%	1%	23%	12%	57%	1%	350
Belted Kingfisher	1%	3%	3%	42%	4%	40%	3%	676
Blue-winged Teal	0%	2%	27%	24%	1%	42%	3%	769
Common Tern	0%	20%	0%	0%	0%	20%	60%	5
Gray Catbird	0%	6%	14%	15%	7%	56%	1%	1107
Great Blue Heron	1%	2%	6%	39%	6%	41%	2%	1025
Indigo Bunting	0%	5%	10%	22%	9%	53%	2%	125
Killdeer	0%	2%	35%	35%	9%	52%	0%	726
Lesser Scaup	0%	0%	26%	33%	0%	40%	1%	143
Osprey	1%	1%	3%	42%	4%	42%	4%	517
Palm Warbler	1%	6%	10%	3%	10%	42%	2%	1324
Piping Plover	1%	1%	2%	73%	6%	17%	0%	307
Ring-necked Duck	0%	0%	25%	38%	0%	36%	0%	441
Ruddy Turnstone	1%	5%	3%	36%	13%	40%	1%	867
Sanderling	0%	8%	2%	70%	7%	12%	0%	275
Sharp-shinned Hawk	0%	0%	0%	0%	0%	100%	0%	3
Turkey Vulture	11%	11%	5%	26%	8%	46%	3%	1482

Chapter 2: Bird Species Diversity and Dissimilarity and their Relationships to Habitat and Observer Variation on Grand Bahama, The Bahamas

Abstract

The use of survey data from citizen science can be a powerful tool in understanding change in species diversity and composition across space and time. A challenge in applying citizen science data to conservation monitoring is to account for differences in observer behavior, sampling effort, and expertise in resulting patterns of species diversity. In this study, long-term data from eBird were used to analyze the bird species diversity and composition in relation to differences in observers across multiple habitat types on Grand Bahama, The Bahamas. Bird species richness was examined using species-accumulation curves as function of observer effort among habitat types. Variation in bird species composition was examined using multivariate ordination by habitat and survey effort, as well as the spatial covariation between bird species composition and observer composition. A total of 278 bird species were recorded on Grand Bahama by 280 different observers, but species-richness estimators suggests there were 20-60 species present that are not represented in the eBird data. Bird species richness varied significantly across seven habitat types, with high bird richness in grass and lowest in water habitats, but richness was strongly influenced by observer effort across habitats. Bird species composition varied significantly among habitat types forming three distinct clusters of habitat types: pine and high-water table communities, grass and wetland habitats, and urban, sand and water habitats. These patterns also reflect the differences in bird observers among habitat locations. Observer characteristics such as their physical ability to navigate a habitat or detect birds, their expertise, and their interest in recording particular birds are difficult to obtain for individual observers. Finally, Mantel tests showed that both bird species composition and observer composition were spatially structured at scales up to 20 km, and that bird and observer composition were positively associated across spatial scales. I share recommendations to reduce differences among observer ability, reduce variation in effort and improve data quality and usefulness.

Introduction

Resource and wildlife managers increasingly rely on data from citizen science datasets. However, the ability or preferences of citizen scientists are not always measured or monitored and they may affect the data (Thogmartin et al., 2007). The ability and experience of the participant observer determines which species they are able to observe and identify (Johnston, Fink, Hochachka, & Kelling, 2018). Furthermore, observers may prefer to visit some habitat types over others, leading to differences in sampling effort as well as observer expertise. Observer knowledge of and access to biodiversity databases, and enthusiasm for recording the data also determine which species they enter into the database (Gregory S Butcher & Niven, 2007; Dickinson, Zuckerberg, & Bonter, 2010). Observers may also choose not to share data publicly even when it is entered into the data base (Groom, Weatherdon, & Geijzendorffer, 2017). I therefore hypothesize that observer effort will vary among habitats and that variation in observer effort among habitats will affect recorded bird diversity on Grand Bahama Island.

Large-scale species and habitat conservation may miss the mark when focused primarily on geographic boundaries, especially with limited and potentially biased data (Jenkins, Van Houtan, Pimm, & Sexton, 2015). Focusing on conservation action and human behavior has been shown to be more effective and economically efficient (Brown et al., 2015). Monitoring programs can use estimates of observer ability to evaluate the efficacy of citizen science programs (Johnston et al., 2018; Kelling et al., 2015). Citizen science training efforts and local conservation would benefit from knowledge of local expertise (Ormsby & Mannle, 2006; Ortega-Álvarez et al., 2012). Trainers can manage differences in observers, through capacity building, when those differences are acknowledged (Barlow et al., 2015). Managers can also use knowledge of effort gaps within the community to adjust efforts to promote citizen science.

Various studies use geospatial statistics to detect variation in observer effort across a physical landscape, but the observer composition is seldom considered and may have even greater variation and effect on the data. Environmental characteristics such as moisture, day length or elevation vary over physical distance between two locations and these variables, in turn, affect species communities (Ross, Brien, & Flynn, 2010). However, geographic bias in sampling efforts can significantly impact the evaluation of species distributions or conservation needs and outcomes (Leitão, Moreiraa, & Osborned, 2011).

Several studies have examined the spatial covariance in species diversity and composition with environmental variables (Chocron, Flather, & Kadmon, 2015; Hargrove, 2010; Merriam, 1898). However there are few studies that examine the covariance in the number and diversity of observers with species diversity and composition (G S Butcher & McCulloch, 1990; Dickinson et al., 2010; Lees & Martin, 2015). Habitat differences have been shown to affect the distribution and composition of species in the Bahamas (Lloyd & Slater, 2010; Wunderle et al., 2010). The effect of observer effort on recorded diversity has been discussed in various papers (Stier, Bolker, & Osenberg, 2016), and some focus on citizen science data (G S Butcher & McCulloch, 1990; Klemann-Junior, Vallejos, Scherer-Neto, & Vitule, 2017), however observer community composition is not considered in previous work. I propose the use of observer community dissimilarity as a distance measure in Mantel tests and Moran's I calculations to determine how the observer characteristics affect effort and recorded species diversity.

The physical distance between locations is typically represented in a two dimensional plane and the distances between locations can be used in Mantel tests and Moran's I calculations (Mantel & Valand, 2018; Moran, 1948). Spatial dependence in aggregate measures of community structure, such as species richness and abundance, are analyzed using Moran's I (F. Dormann et al., 2007), whereas pairwise dissimilarities in species composition, such as Jaccard's index, requires multivariate analyses Mantel tests or ordination methods (Mantel & Valand, 2018). These univariate and multivariate tests analyze spatial patterns of species diversity and composition and can also be applied to the community of observers. This leads to three kinds of analyses: (1) the relationships between spatial location and species diversity or composition; (2) the relationships between spatial location and observer diversity or composition; and (3) the spatial covariation between species and observers in their diversity or composition. The Jaccard or Bray-Curtis index of dissimilarity between species or observer communities can be represented as a distance using multivariate ordination (Kosub, 2016).

In this study, I first considered the variation in bird species diversity across habitat types, and examined how it relates to observer diversity and total effort on the island of Grand Bahama. Second, I used Mantel tests and Moran's I to describe the changes in species and observer diversity and survey effort over geographic distance (Mantel & Valand, 2018; Moran, 1948). Within geographic space, I hypothesized that observer diversity and observer effort would be negatively autocorrelated and that species diversity would be positively autocorrelated. Second, I hypothesized that there would be a positive correlation between geographic distance and both the dissimilarity in bird species composition and the dissimilarity in observer composition. Third, I hypothesized that bird species dissimilarity would be positive associated with observer dissimilarity over geographic space.

Methods

Data from Grand Bahama Island, Bahamas was downloaded from eBird in January 2017. The data were filtered to those surveys during which the observer attempted to identify all birds in the survey area and which occurred during the years 1988 to 2016. Habitat information was appended to each Locality ID (hereafter locality) using the methodology outlined in Chapter 1 and available at https://github.com/Ancilleno/habitat-effect-on-eBird-observers. The number of surveys during which each species or observer occurred at the locality was entered into species and observer detection matrices respectively.

Species and Observer Diversity

The total species detected at each locality by all observers within the study timeframe was calculated for each locality and summarized by habitat. The total unique observers that conducted surveys at each locality was also calculated. Summary statistics and species accumulation curves were calculated for the complete data set and individual habitats via the 'vegan' package, after (Oksanen et al., 2015). Total species richness was estimated using Bootstrap, Chao, and first-order Jackknife methods for the island wide dataset. I calculated species richness for all locations via the vegan package in R (Oksanen et al., 2015). Observer richness was calculated as the number of observers who conducted surveys at the location, during the time analyzed. The total number of surveys at each locality was also calculated from the data. I used a general linear model with a Poisson error distribution to analyze the effects of habitat type (as a categorical variable) and survey effort (expressed as the log number surveys) on bird species richness by locality. The statistical significance of the parameter estimates for each habitat type and log number of surveys were used to interpret differences in species richness by habitat and survey effort. To test the significance of the overall model, a likelihood ratio test was conducted on the model with and without the effects of habitat type and survey effort. A similar model was constructed for the effects of habitat type on observer richness, but without the effects of survey effort.

Bird Species Composition

The overall variation in bird species composition by habitat type and observer effort was analyzed with a constrained ordination using distance-based redundancy analysis (dbRDA; McArdle & Anderson, 2001). The analysis was conducted with the dbrda function in the 'vegan' package using Bray-Curtis dissimilarity among sample locations. Habitat type was used as a categorical predictor, and number of surveys was used as covariate in the ordination. The sample localities were plotted on the first and second multivariate dbRDA axes, together with the centroids of the localities by habitat and a biplot arrow for the numbers of surveys.

Mantel tests

Three distance matrices were created for the data set and for each habitat within the dataset. The Euclidean distance between each of the 501 survey locations was used to create a geographic distance matrix. The Bray-Curtis dissimilarity matrices were calculated separately for the observer identities and the bird species via the 'vegan' package in R (Oksanen et al., 2015).

Mantel tests were used to determine if the matrices of geographic distance and community dissimilarities were correlated using the Pearson correlation. If there was a significant overall Mantel test, I constructed Mantel correlograms to determine the spatial scales where bird species dissimilarity or observer dissimilarity was most important.

Moran's I

I used Moran's I (Gittleman & Kot, 1990) to determine if there was significant spatial autocorrelation among species diversity and effort measures across the locations in the data set via the "Moran.I" function in the 'ncf' package in R (Bjornstad & Cai, 2018). The Moran's I test in R returns an expected Moran's I value, the observed Moran's I, standard deviation (SD), and p-values based on randomization tests.

I used Moran's I to examine autocorrelation of reported species diversity and observer effort with geographic distance. I hypothesized that the species diversity and observer effort would be positively autocorrelated. I used Moran's I to test for autocorrelation of observer effort over species community dissimilarity. I hypothesized that observer effort would be positively autocorrelated with geographic distance.

Results

Bird species richness was significantly affected by both habitat type and the total numbers of surveys (Deviance = 6012, df=7, p <0.0001). Grass, pine, hwtc, water, and wetland had similar numbers of observed bird species, and these habitats had significantly higher numbers of bird species than sand and urban habitats (z-scores= -8.79 and -5.24, respectively; p <0.0001; Figure 11). The number of sampling events at a location showed a strong positive relationship to the number of bird species reported at the location (z score =84.8, p <0.001; Figure 12). Observer richness was also significantly different across habitat types (Deviance = 167, df=6, p <0.001), with grass significantly higher than all other habitat types (z-scores all p <0.0001) and pine, water and wetland having the lowest numbers of observers (Figure 13).

A total of 278 species were detected in the analyzed data and Bootstrap, Chao and Jackknife estimates of total species were 21 to 80 species higher (*Figure 14*). Habitat specific species accumulation curves were distinct except for hwtc and sand habitats which overlapped until about 18 localities which is the total localities in hwtc habitat. Species accumulation curves in pine and urban habitat overlapped until approximately 80 localities. Grass had the highest curve and water the lowest with wetlands, sand, hwtc, pine and urban habitats in between and in that order (highest to lowest).

Observer accumulation curves for the full dataset and within different habitats had a lower curve than bird species curves. Grass was the highest and water the lowest. Wetland habitat was lower than sand until they overlapped at 40 locations (*Figure 17*). Pine and urban were not as closely related as in species rarefaction curves and pine remained lower than urban habitat in the data set.

The distance-based redundancy analysis plot of variation in bird species composition using habitat and number of surveys as predictor variables showed high overlap in three groups of habitats (Figure 18). First, sand, urban and water habitats form a fairly small group with sand and urban 95% confidence ellipses overlapping with water but not each other. Grass and wetland habitats overlap almost completely with one another and more than 50% with hwtc. The hwtc group stretches along the first axis and completely overlaps pine habitat, which does not overlap with any other group. Sand, urban and water habitats are negatively constrained on the first two axes of the ordination while hwtc, Pine and wetland habitats are positive on both axes. The number of surveys conducted scores negative on axis 1 and positively on axis 2. A permutation test on species composition using 1000 permutations of the data against a null hypothesis showed that habitat and numbers of surveys were significant predictors of bird species composition (habitat – F-test=6.82, df=6, P=0.001; surveys – F-test=27.9, P=0.001).

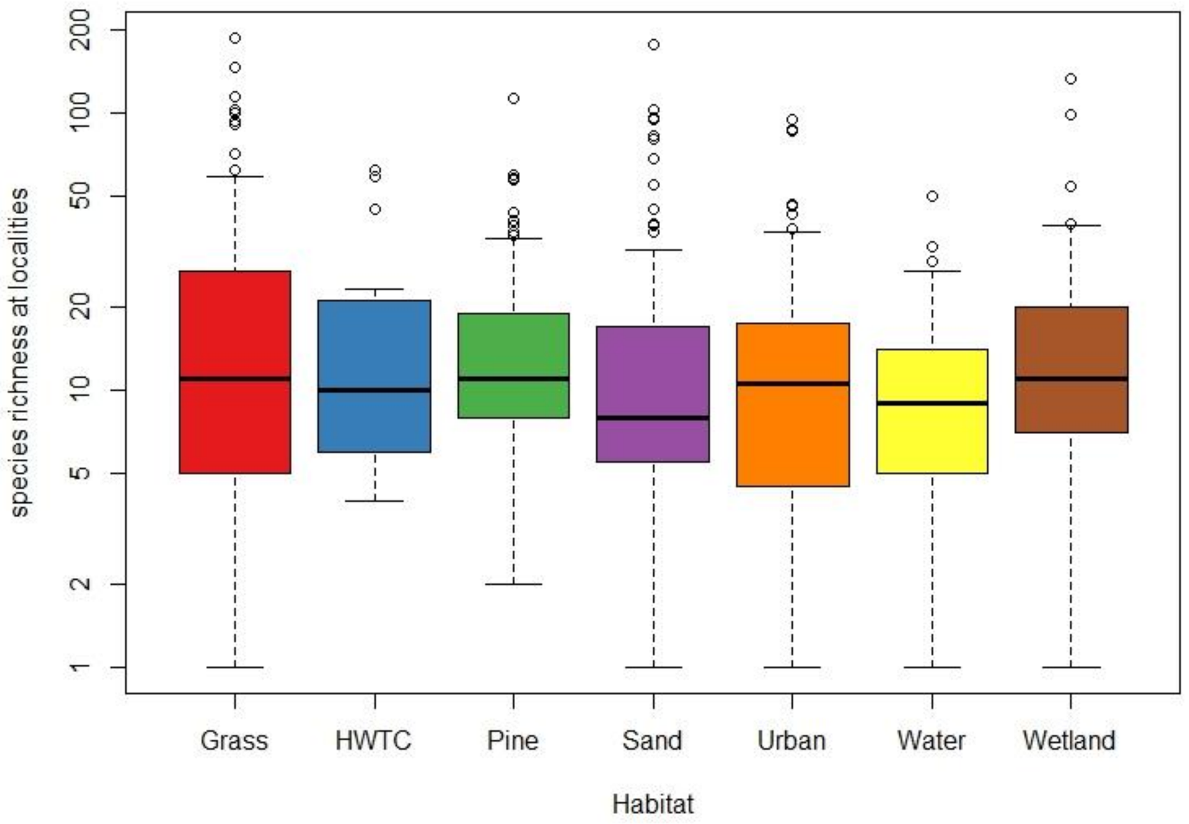


Figure 11. Boxplot of species richness at localities, by habitat on Grand Bahama, The Bahamas based on eBird data from 1 January 1988 to 31 December 2016: note: log scale on y-axis.

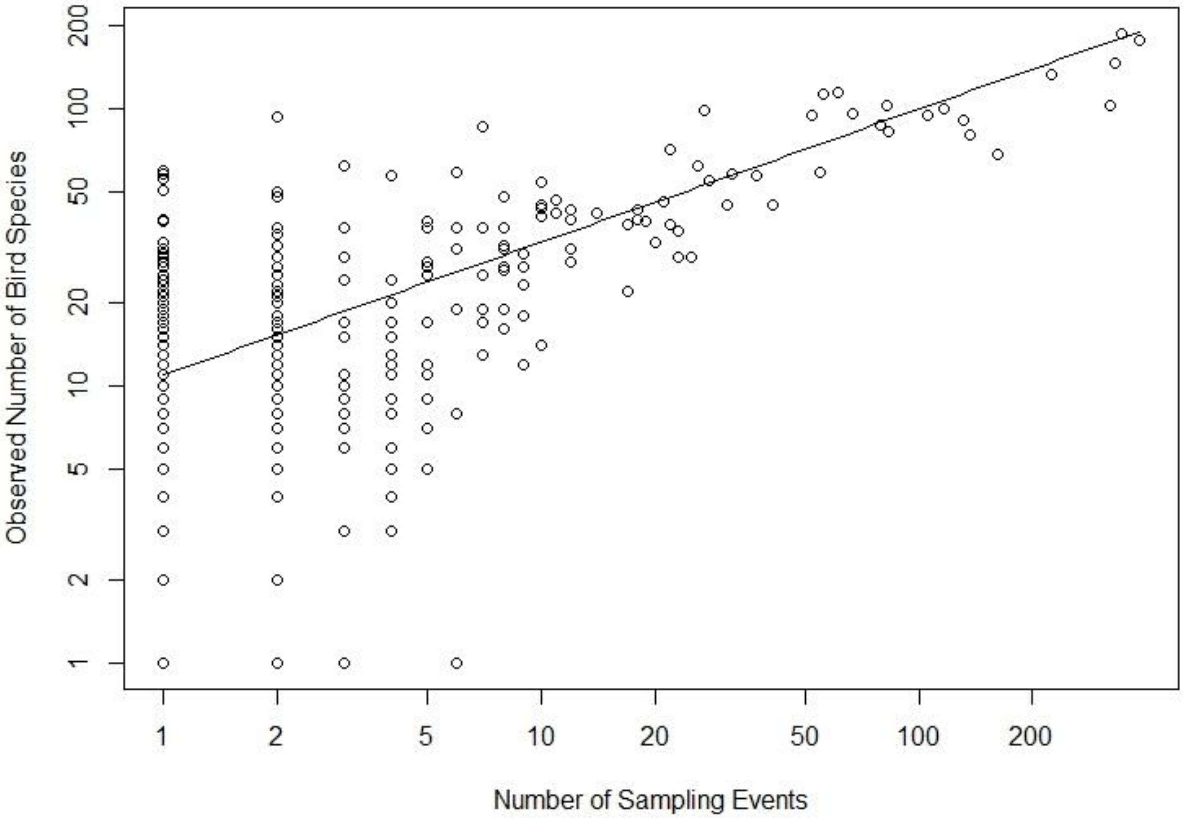


Figure 12. Scatterplot of the number of bird species detected at locations on Grand Bahama, The Bahamas and the number of sampling events conducted at the location. Trend line shows positive correlation. Based on data collected between 1 January 1988 and 31 December 2016.

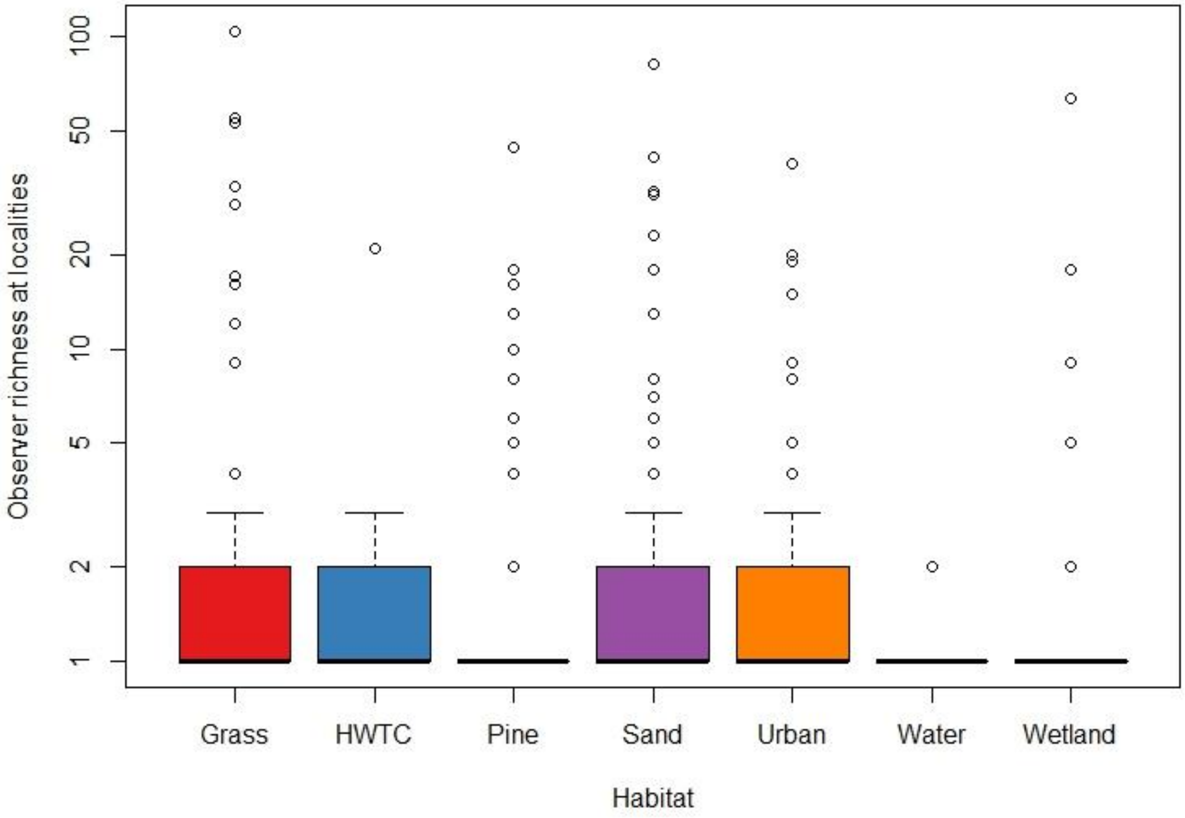


Figure 13. Boxplot of Observer richness at localities, by habitat on Grand Bahama, The Bahamas based on eBird data from 1 January 1988 to 31 December 2016: note: log scale on y-axis.

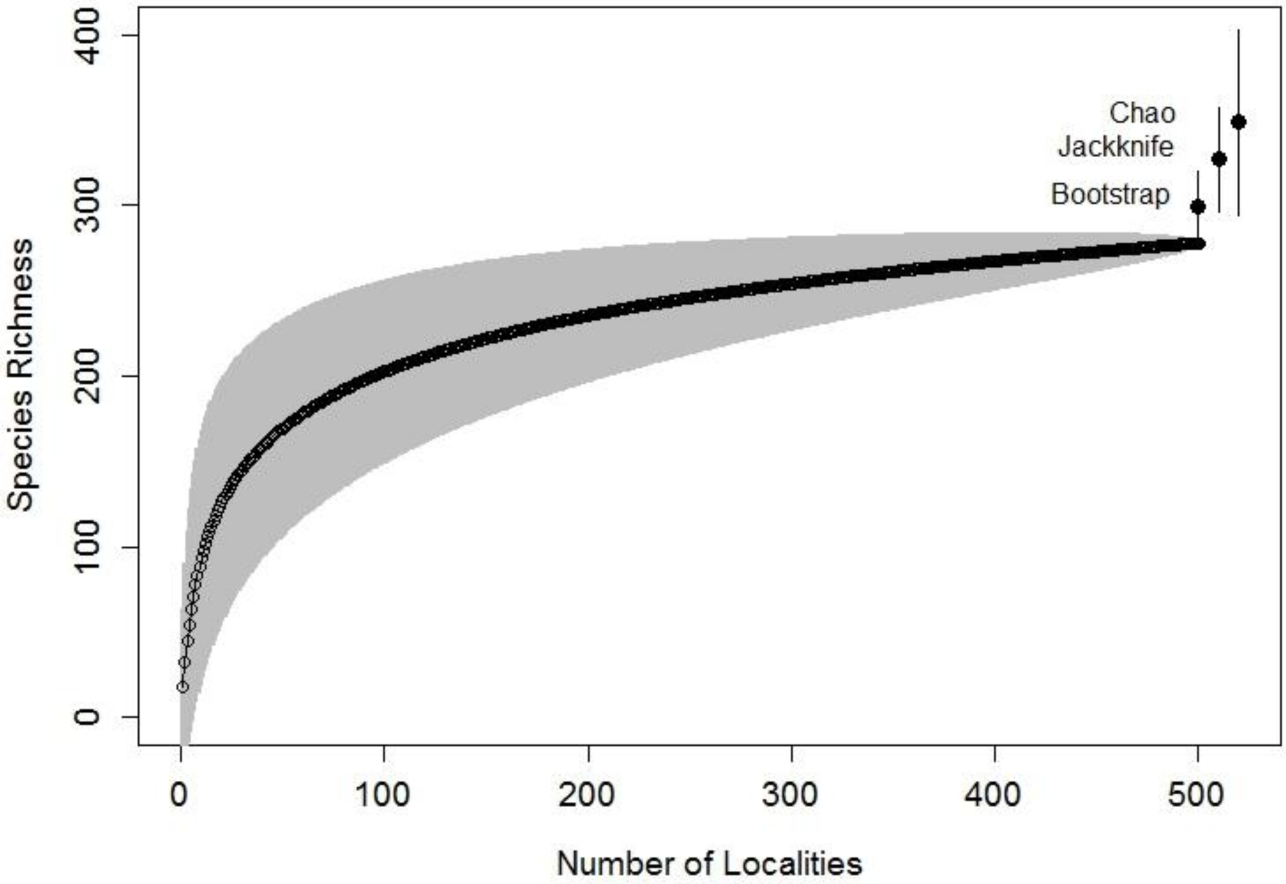


Figure 14 Species accumulation curves for localities on Grand Bahama, The Bahamas based on eBird data collected between 1 January 1988 and 31 December 2016 with Bootstrap, first order Jackknife and Chao estimates of true species richness.

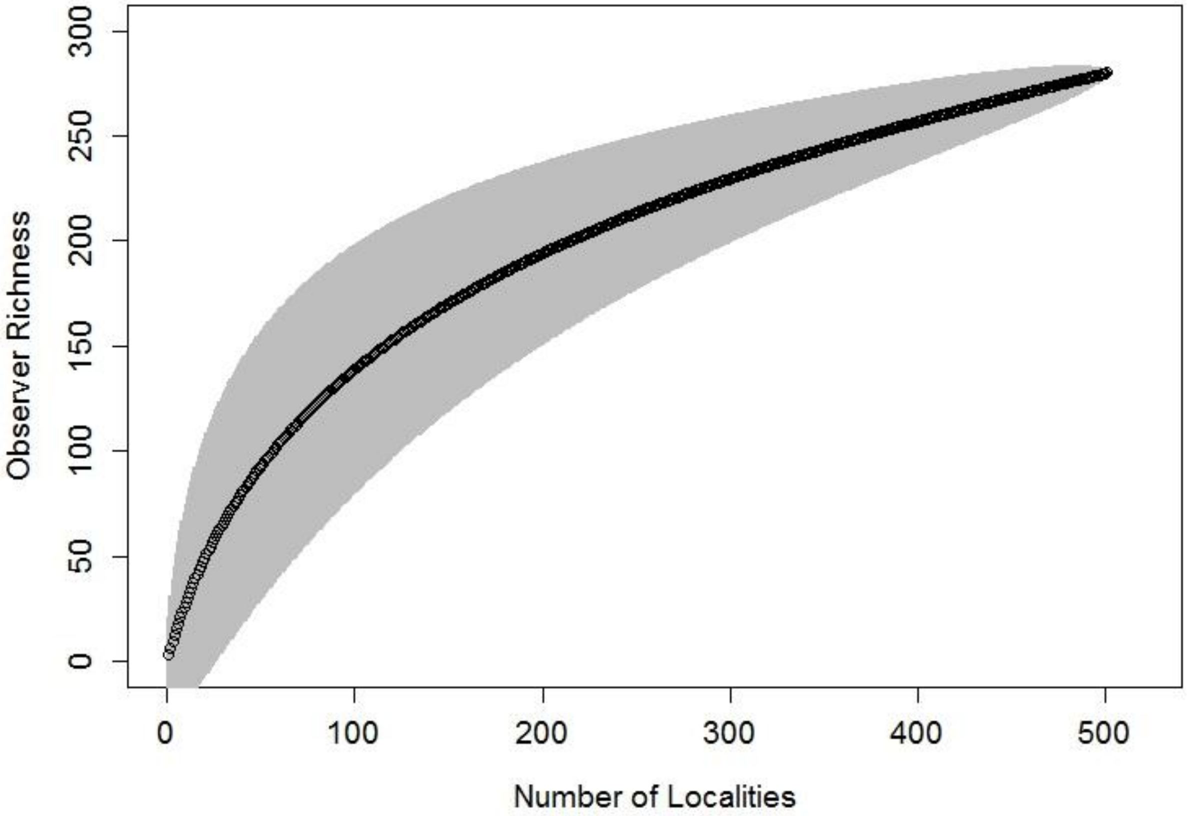


Figure 15. eBird observer accumulation curve for Grand Bahama, based on surveys conducted between 1 January 1988 and 31 December 2016

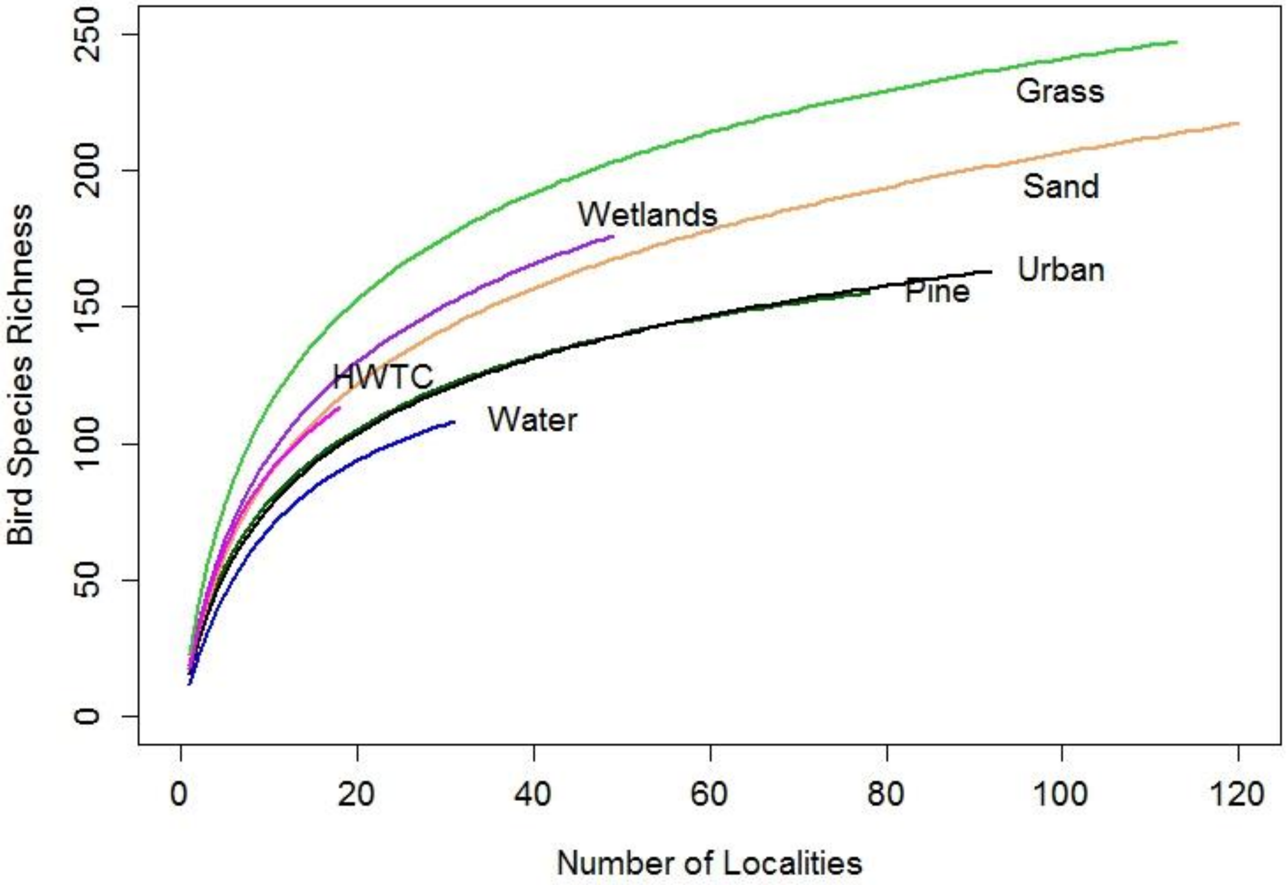


Figure 16. Habitat specific species accumulation curves by localities on Grand Bahama, The Bahamas, based on eBird data collected between 1 January 1988 and 31 December 2016.

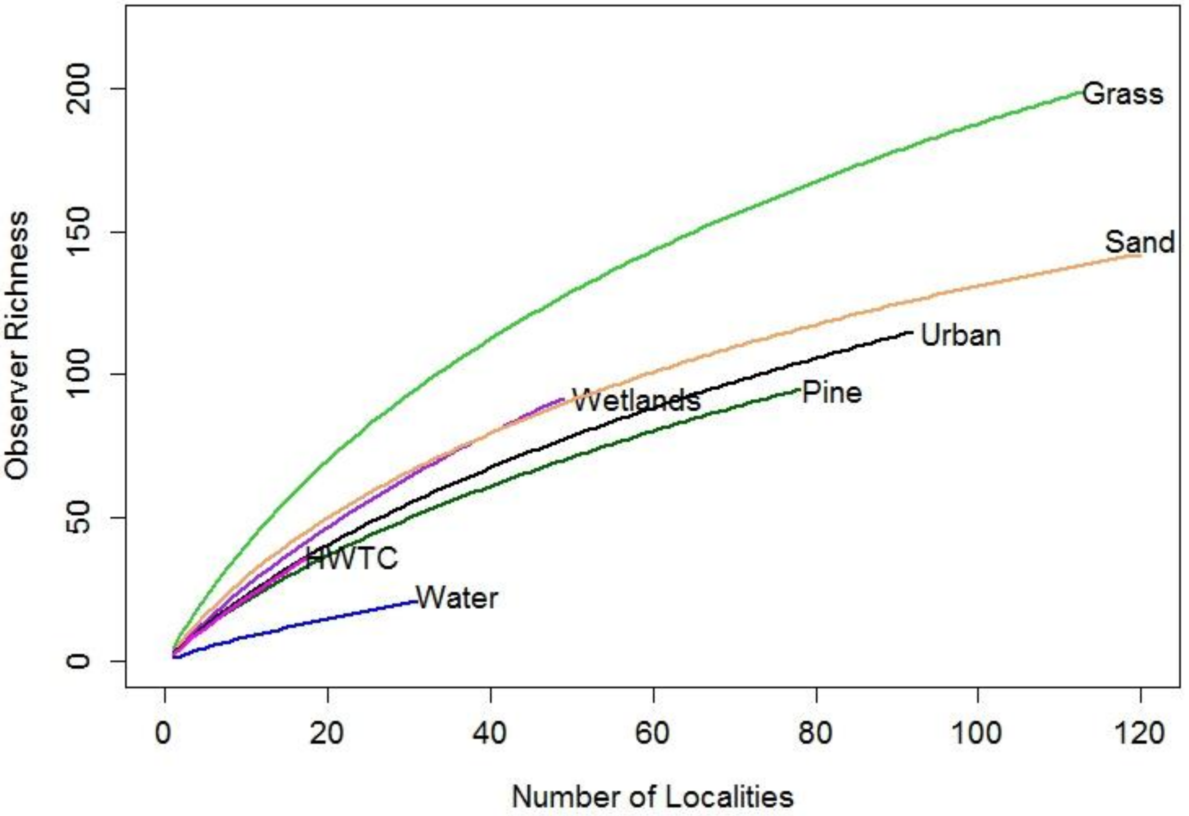


Figure 17. Observer richness rarefaction curves for habitats on Grand Bahama, THe Bahamas based on eBird survey data collected between 1 January 1988 and 31 December 2016

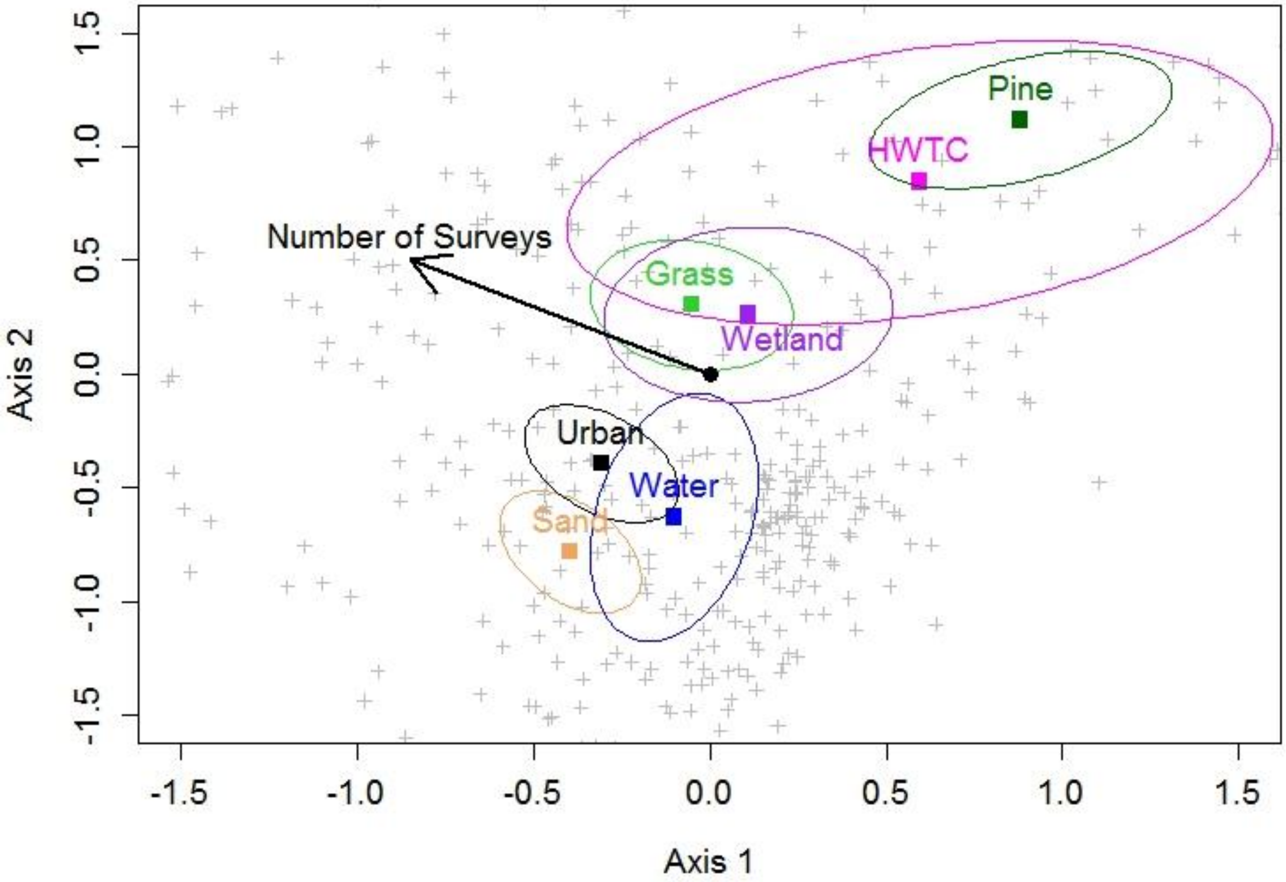


Figure 18. Ordination plot of Bray-Curtis species dissimilarity at localities in eBird data collected on Grand Bahama, The Bahamas between 1 January 1988 and 31 December 2016 in six habitat types (hwtc = High water table community).

Mantel tests

Mantel tests showed a statistically significant but weak positive correlation between geographic distance and Bray Curtis dissimilarities in observer communities on Grand Bahama (Mantel r = 0.081, P≤0.001; *Figure 20*). Bird species dissimilarities exhibited a stronger positive association with geographic distance (Mantel r = 0.203, P≤0.001; *Figure 19*). Bird species dissimilarities and observer dissimilarities showed a similarly strong positive covariance across geographic distances (Mantel r = 0.209, P≤0.001). All relationships were significant at distances less than 20 km.

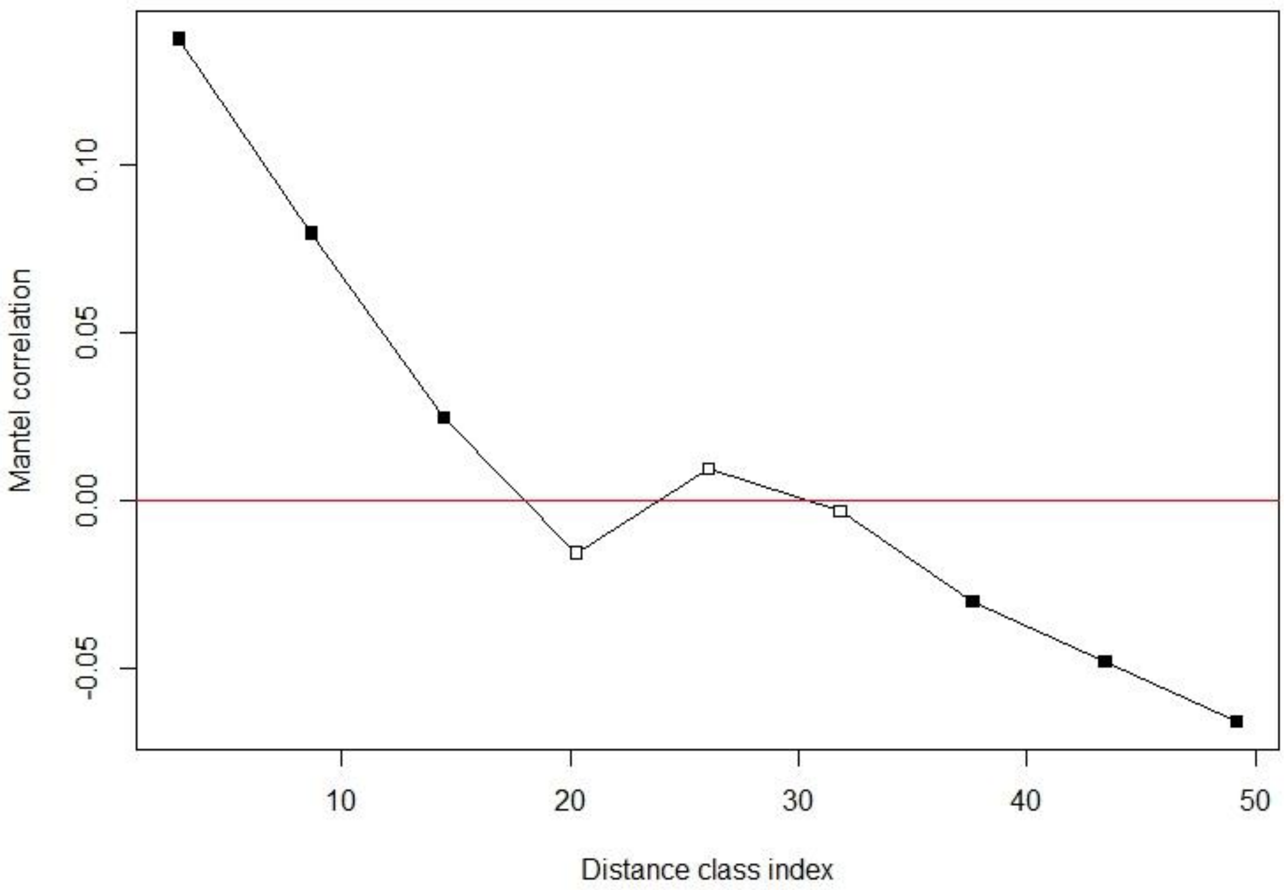


Figure 19. Mantel correlogram for species community dissimilarity over geographic distance. Significant positive correlations were found for distances less than 20 km.

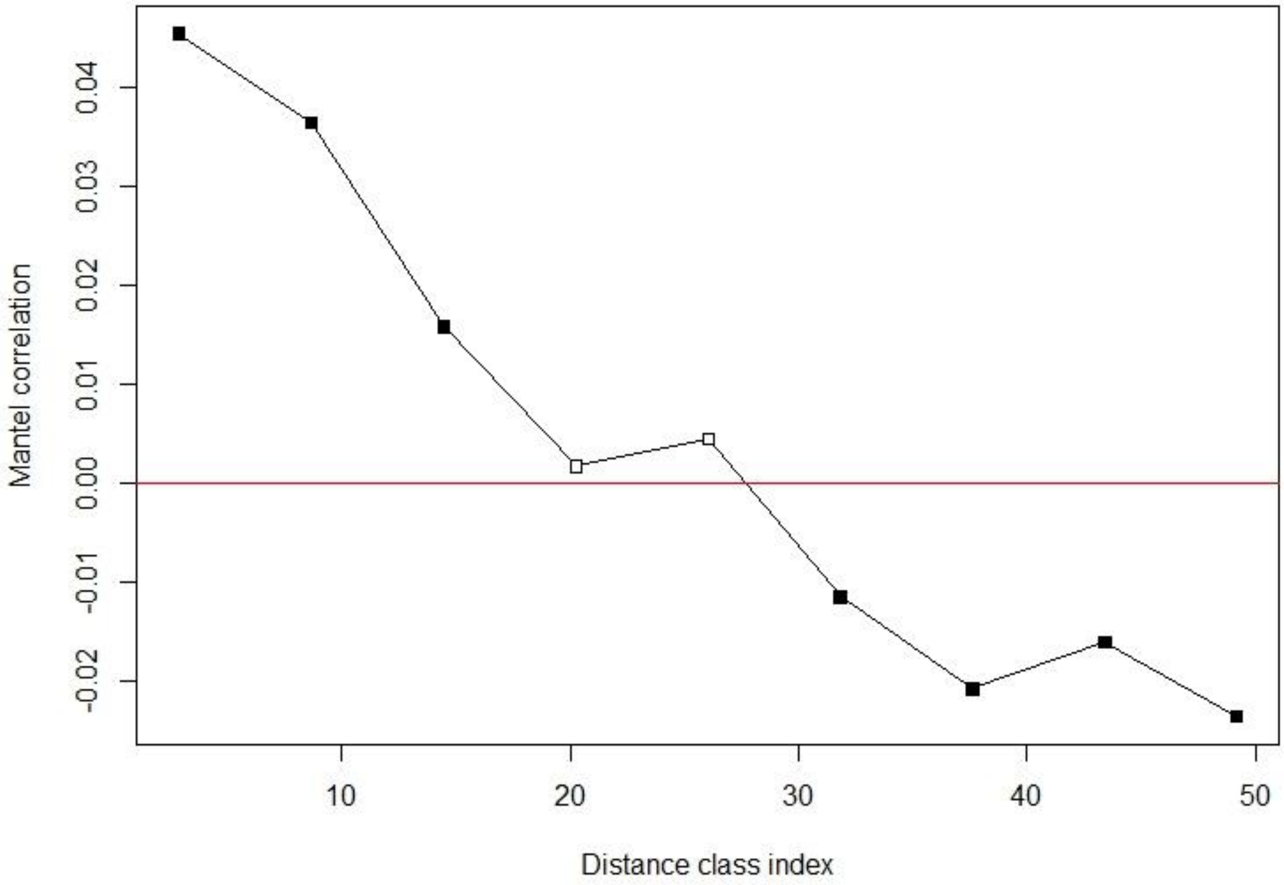


Figure 20. Mantel correlogram for observer community dissimilarity over geographic distance. Significant positive correlations were found for distances less than 20 km.

Moran's I with geographic distance

Moran's I tests for spatial autocorrelation were generally not statistically significant ($p>0.05$). I therefore cannot reject the null hypothesis that these values may be the result of spatial randomness. There were a few exceptions. Observer richness and the total surveys at the locations in the complete data set showed negative spatial autocorrelation. Locations that were near to one another were less likely to have similar values than could be expected at random. Shannon's species diversity was positively spatially autocorrelated in in the overall dataset, but not in any of the individual habitats.

Only water and sand habitats had significant spatial autocorrelation for the tested values. Water habitats exhibited positive spatial autocorrelation for the total surveys conducted at locations. Sand habitats exhibited negative spatial autocorrelation for the total minutes spent surveying at the locations.

Discussion

My results support other research that observer effort at a location is directly correlated to the number of species detected (G S Butcher & McCulloch, 1990) and that the species distributions are affected by the habitat in which surveys are conducted. Habitat type strongly influenced bird species richness and observer richness within the data set. Both habitat type and the number of surveys conducted at the location strongly influence species richness. This is expected as different habitats not only have more resources or naturally occurring species diversity, but different habitats also have different levels of visibility for birding. The higher species accumulation curve in grass habitat is a good example as that habitat is the most open and therefore birds using that habitat or forest edges would be more easily visible. Sand habitat is similarly open and birders are more likely to see active species. Similarly, pine and urban habitats are moderately open, but lines of sight are interrupted by pine trees and houses, respectively.

Interestingly, observer richness at a location does not strongly affect bird species richness. There are several reasons. Firstly, most locations have only one observer. Those observers sometimes record a single species but can also record dozens of species. On the other hand, very popular locations may have their species pool completely or nearly completely reported so as more observers visit the location this does not change the number of species already reported.

The differences and similarities in observer and species accumulation curves are primarily explained by observer behavior. Observers accrue faster in grass and sand habitats because those locations are more readily accessible to tourists and guided tour groups. Birders can readily find localities, in these habitats once they are listed in eBird. Tourism locations that are hotspots of tourist activity may also receive significant numbers of visiting birders by chance alone. These locations are primarily in grass and sand habitats such as golf and beach resorts or local parks and beaches along main roads. When tour guides bring a group to a location they also quickly increase the number of observers by reporting several observers on one survey or at the location, though they were all birding as a unit.

This difference indicates that some locations are visited much more heavily by birders than others in touristic areas. In particular, grass, wetlands and pine habitats showed a marked difference in observer localities. This can be expected as major tourist attractions related to birding such as the Garden of the Groves and the Garden of the Gates are mainstays of birding tours. The properties under protection by the Bahamas National Trust on the island are also heavily visited and may bias the data significantly. Other properties such as the beachfront properties of major hotels on the island help to lower the effective locality values in sand habitat. The convenience of the birding location immediately outside the hotel accommodation allows for convenient birding at multiple times during the day and can result in higher numbers of surveys for the same observer even if they are only on the island for a short period of time.

Mantel tests

Mantel tests showed that geographic distance correlates weakly to observer community dissimilarity and more strongly to species community dissimilarity. There was also a significant covariance between observer community dissimilarity and species community dissimilarity over geographic distance.

Species community differences across geographic distance has been well supported in other work (Fortel et al., 2014; Tryjanowski & Morelli, 2015) and is also supported here. This is expected on Grand Bahama, as the habitats are generally contiguous at the sample scale. Therefore, locations that are nearby are normally in the same habitat and have the same species.

The weak correlation between observer community dissimilarity and geographic distance can be explained by several characteristics of observer behavior in the data set. The first is the frequency of singletons. Observers that visit only one location, increase the number of unique sites, which have zero similarity with any other location. These observers form a large part of the population and their visitation necessarily shows no pattern. Another important group of observers are avid birders attempting to record as many different species as possible. These birders visit multiple locations, but are less likely to record a neighboring site in the same habitat after conducting one survey in that habitat. Here there may be a difference in reality and recording. A birder that starts a survey at a neighborhood intersection for example may walk several blocks in any direction but record the survey as a point location so the geographic location may not represent the actual effort and in some cases even overlap with other locations. Researchers with specific preferences may bias data within a habitat type when they enter their data in a citizen science data base. In our data, a single researcher conducted surveys in more than 20 pine forest localities. That single observer therefore created multiple locations with identical communities (he was the only observer at the locations he visited). At the same time, that researcher never visited any locations that had been visited by other observers. Therefore, all his locations were unique to him. Finally, tour guides are also less likely to take tour groups to multiple similar locations or multiple locations that are immediately adjacent to one another. This creates similar groups of observers that are recorded at spatially distant locations and often in different habitats.

The hwtc communities do show significant correlation between geographic distance and observer community differences. This correlation is primarily because of a small amount of observers that have conducted surveys in that habitat type and those observers conducted several different surveys in relative proximity to one another but distant from the other surveys in that habitat type. Because of the logistic cost of traveling to hwtc communities from housing and other amenities, and the difficulty in moving around within them, it is unlikely that most observers will visit hwtc habitat to conduct surveys, so these specialty observers drive the dynamic here.

The significant covariance between observer community dissimilarity and species community dissimilarity supports my hypothesis. Observer identity does affect the species recorded at a location in this dataset. Though the significant relationships were restricted to less than 20 km it is possible that this relationship will hold at multiple scales on different islands or in cities in the continental US.

Moran's I

The negative spatial autocorrelation in the eBird dataset for observer richness and total surveys conducted seems to primarily reflect individual and group behavior among birders. As individuals, birders tend to find a "good" birding spot and spend their time birding there. By the time they are done, their ideal birding time has expired and they do not go to another area nearby to look for birds again. Similarly, birders that are introduced to a high diversity spot will remain there birding with their group, and very few individuals will leave to go to other areas nearby for birding. This leads to high intensity spots surrounded by a few nearby locations with the occasional observer. Tour guides again are unlikely to take an entire group from a single location to a distinct location nearby to continue birding, instead they may walk a considerable distance in the vicinity but identify it all as one location.

When the habitats were considered separately, there was generally no significant spatial autocorrelation, either negative or positive for the variables assessed. Except for negative spatial autocorrelation in total minutes surveyed at sand locations and positive spatial autocorrelation at water locations. Clumped distribution of minutes surveyed at sand locations was found primarily around coastal hotels and resorts. Again, this is related to the observer convenience. The birders in this case seem to be more leisurely birders that spend a significant amount of time at a beach or pool near their accommodations and passively record birds they see. Some birders reported up to 15 hours of birding at a location though their trip comments, showed birding to be a secondary activity to golfing, banding piping plovers and a photo shoot. Other birders resorted to birding when they could not take part in another preferred activity, comments such as "course closed today" indicated that a golfer decided to go birding when golf was not an option.

Summary

Observers may have differing abilities to identify certain species especially among the woods warblers. Detection of cryptic species and audio detection of species by sound without visual confirmation is also a special skill that comes in useful in dense foliage, but not all observers have (Miller, Hallo, Sharp, Powell, & Lanham, 2014). Several species are limited to a few observers in the database. These observers, when present at a site may significantly change the reported species composition. Unfortunately, due to lack of temporal overlap between most observers, it is difficult to determine if the detection difference was simply due to that observer's ability or because the bird in fact was not there.

The correlation between observer and species community differences indicates that certain groups of observers are more likely to report certain groups of species within a habitat. This was expected for all habitats but was not seen in urban habitats. One potential reason for the lack of correlation within urban habitat is the prevalence of ubiquitous urban species such as invasive or generalist species (Northern Mockingbirds, House Sparrows, American Kestrels etc.). These species may increase the similarity of communities throughout the urban habitat making it more difficult to detect correlations. Urban habitats also have lower reports of easily disturbed or specialist species. It is more difficult to detect rare species that may require audio detection or that actively avoid humans.

Overall, it is reassuring that the observers on the island are so diverse in their behavior and ability as to exhibit differences in their species records visible in the data. However, this exposes a major weakness in citizen science data. The citizens are influencing the data in a tangible way. I propose several steps toward improving the quality of the data collected in eBird.

The first is being implemented via eBird and its partners already. They are involving more people in the use of eBird and in particular the mobile applications that improve the accessibility of the eBird program to people of various backgrounds. However, in the Bahamas, this can be improved by including some form of bird identification coursework in the standardized high school biology curriculum or job training programs for tourism businesses. In addition to incorporating the eBird program in student work, making it accessible to and collaborating with government agencies in official capacities could facilitate adoption at large scales.

Secondly, monitoring and/or standardizing observer ability is an essential component of quality assurance for citizen science data. Several marine ecosystem survey programs have integrated observer training to ensure the quality of reported data. The REEF fish ID project allows novice observers to volunteer data without any training, but then also offer expert training to certify observers at multiple levels and for different regions (Pattengill-Semmens & Semmens, 2003). This tiered format certification then allows researchers to evaluate the observer data more accurately by separating verified expert observers from novice observers with unknown identification skills. ReefCheck (Hodgson et al., 2004) and AGRRA (Atlantic and Gulf Rapid Reef Assessment; Pattengill-Semmens & Semmens, 1999; Wilkinson, Green, Almany, & Dionne, 2003) methods also include observer training. However, there is currently no standardized format or certification within the eBird dataset or the overall program. The nearest analog is the Bird Guide Training Program being implemented through BirdsCaribbean and the Caribbean Birding Trail, throughout the region to establish bird guides as experts in bird identification. If eBird could add a badge system that would certify an observer as having completed some formal certification, this would allow their identifications to carry more weight in analyses, when compared to unverified observers. It could also serve to increase the drive among the community to improve their skills in recognizable ways, though care must be taken not to make such certification required or prohibitively costly.

Lack of consistency or surveys of significant length or coverage by common observers also reduces the usefulness of the data. Further training for observers to standardize survey methodologies could significantly improve the data. By having multiple well-trained observers deliberately and representatively survey the island, it would be possible to better understand species distributions and also put the data collected by other untrained observers into perspective. Standardized minimum survey length, times and methods can be used to improve our understanding of species detectability and distribution. This could also be an opportunity to develop community in an area by recruiting willing observers through eBird to meet at specified dates and times to conduct surveys that fill in knowledge gaps.

Finally, integrating optional habitat or climatic variables would improve our understanding of the resource and environmental needs of the species detected. AGRRA (Lang, 2016) and ReefCheck (Hodgson et al., 2004) programs have additional requirements beyond species identification certification that also include understanding of the underwater habitat for coral reef communities and include fields for entering climatic variables such as temperature and depth. Options for including habitat, climatic variables and standardized survey parameters make it easier to calculate species distributions and other variables from the data in those programs. Some data such as the temperature at the location, wind speed and direction, precipitation, fruiting phenology of common plant species and human activity may be significant factors affecting bird activity and their detection (Wiley & Wunderle, 1993). In Chapter One I showed how important habitat type is in the data set, but these variables cannot be collected via satellite at such fine resolution. When observers are certified in Bahamian bird identification, further training in climatic variable reporting or inclusion of these variables on the digital form could significantly improve the usefulness of the data in terms of trip planning for observers or conservation planning for experts.

Chapter 3: Here together, there together: species and observer co-occurrence in eBird records on Grand Bahama, The Bahamas

Abstract

Large-scale Surveys for conservation and monitoring are time-consuming and expensive. Citizen science provides an inexpensive alternative, but variation in observer expertise and effort makes it difficult to interpret these data. In this study, I used a combination of indicator species, and co-occurrence analysis to determine associations among bird species, observers, and habitat types on Grand Bahama, The Bahamas. If year-round resident and migratory species share habitat, monitoring and conserving habitat for year-round residents will provide more sustainable job opportunities for local scientists and conservation agents and improve conservation outcomes. Bird species co-occurrence in observer records indicates similarities in species detection and may provide understanding of resource partitioning. Understanding co-occurrence in observer records lets us target knowledge gaps and determine species that have similar detectability. This knowledge can advise the development of training protocols for species identification and placing species into meaningful groups for identification guides. Most co-occurrence analysis focuses wildlife. In citizen science data, co-occurrence of observers, can tell us about the observers' behavior and help us discriminate among observers whose collective ability may skew the data or observers who influence on one another. Co-occurrence at localities shows us when site-specific factors such as marketing or shared interests such as golfing influence observers, and bring them to the same locations. Observers also co-occur among species; Observers with similar abilities may detect similar birds. Understanding this may allow trainers to fill knowledge gaps. Observer training can focus on species that behave differently from those being detected. With knowledge of the observer demographics we may also be able to determine when bird records are being artificially limited by some factor that affects the birder and not the birds. For example, golf course access is restricted in the Bahamas to affluent visitors normally and therefore those affluent visitors may co-occur with respect to common golf course species such as waterfowl. Finally, locality co-occurrence in respect to observers or species implies shared site characteristics or relationships such as management or policies that promote bird species or invite or restrict observer participation. I used probabilistic co-occurrence analyses to examine the co-occurrence of species pairs in the eBird dataset. I calculated probabilistic co-occurrence between bird species observers and localities on Grand Bahama and I determined indicator species in the data to detect proxy species that may be used to support habitat monitoring for less common species. I show how these co-occurrence analyses may support conservation management by determining essential site characteristics and monitoring proxies. I discuss the effect of management and marketing on observer communities and species communities. I also determine suites of indicator species that may be used as monitoring targets to represent important bird groups.

Introduction

Species monitoring and population assessments typically focus on a single species or group of species (Devictor, Whittaker, & Beltrame, 2010; Purcell, 2011; Scholes, Gill, Costello, Sarantakos, & Walters, 2017). Research often involves biological surveys that focus on a small area over a relatively short period (Joseph et al., 2009; Rapai, 2012). Researchers sometimes invest hundreds of hours to detect the species of interest, and then work to collect significant amounts of data from a few individuals and environmental data from the detection locations (Hayes, Barry, McKenzie, & Barry, 2004; Lloyd & Slater, 2011), which may represent a small part of the species' actual distribution (Sorenson & Wunderle, 2017). The researchers often focus on historic detection locations that may have changed significantly over the course of a year. These detection locations are often restricted to the detections of that single species and do not take into account the suite of species that may share the same habitat (Brudenell-Bruce, 1975; Rapai, 2012) but see Currie et al. (2005). The current model involves researchers waiting until the birds have migrated and then following the birds to the wintering grounds and conducting in situ searches at historic locations or similar sites (Radabaugh, 1974; Wunderle & Arendt, 2017). Focused systematic study is important in research but has limited usefulness for large-scale monitoring, management and protection (Loss, Loss, Will, & Marra, 2015; McKinley, Briggs, & Bartuska, 2013).

One way to improve efficiency and sustainability of biodiversity monitoring and research is reducing the time it takes to find the species of interest or improving the likelihood of species detections. Monitoring proxy species that indicate an increased likelihood of conservation target species may be an efficient method of reducing search time for rare species. Just as corals are used as climate proxies (Gagan et al., 2000) and birds are used as proxies for species diversity or ecosystem health (Tryjanowski & Morelli, 2015), perhaps one species can be used as a proxy for another target species in particular. One way of increasing the size and scope of species detections is to use citizen science data on rare migratory birds that are associated with year-round resident species or more common migratory birds in the Bahamas. In this study, I use a combination of single-species analysis, habitat indicators, and patterns of species co-occurrence to determine when species are more or less likely to co-occur at the same locations. Species co-occurrence models allow us to determine when two or more species share sites (positive co-occurrence) or when they are found in different sites (negative co-occurrence) (Sfenthourakis, Tzanatos, & Giokas, 2005; Veech, 2013). I also use the same method with observers to determine if they co-occur spatially and discuss how that co-occurrence may be the result of social, geographic, or ecological factors. Finally, I determine locality co-occurrence across species and observers. In this way, I consider localities that co-occur as types or groups, which share characteristics that allow them to support the species or observers they co-occur across. I use the same dataset to determine species that co-occur and use common species to predict the occurrence of rare species; to detect differences between groups of observers that are brought to the island by different circumstances; and to detect groups of localities that meet species conservation targets by attracting species or species groups of concern.

I hypothesize that species of conservation concern will have significant positive or negative co-occurrences with more common native and migrant species. I also hypothesize that visiting observers will have significant positive co-occurrences with local bird guides and that local bird guides will positively co-occur with one another. Finally, I hypothesize that localities within the same habitat will positively co-occur with other localities in that habitat and localities with similar management types will co-occur with other similarly managed localities with respect to species and observers.

Methods

Species of local concern include native endemics and birds that breed in the island's pine forests according to (Lloyd & Slater, 2011). Detroit Audubon Society species are primarily migrants recognized as occurring in both the Bahamas and Detroit River Valley (The Detroit Audubon Society, 2003). These species visit both the Bahamas and Michigan either as migrants or having a broad range that includes both locations. They are included here, because they provide a unique opportunity for transboundary conservation and interagency or international management. Species of local concern and Detroit Audubon Society species are listed in the tables in chapter 1. Rare species are the 50 least recorded species in the data set. The data were also filtered to focus on the 50 most popular localities and the 50 most active observers in the data set. Removal of less common observers and localities improved processing performance but did not change the dimensions of the co-occurrence matrices.

Indicator species analysis

Indicator species analysis was conducted using the indval function in the labdsv package in R (Roberts, 2016) to determine species that represented significant clusters of species with regard to habitats within the dataset. The method identifies those species from the larger data set that show significantly high levels of habitat specificity by comparing their occurrence and abundance to null distributions across habitats. Species were selected if the probability of their level of representation was $p < 0.05$. These species were then used to narrow analyses and discussions, especially where the detected indicator species also fell into one of the focal species categories.

Bird species co-occurrence at localities

Species co-occurrence matrices were calculated to determine relationships between species at the locality level using the 'co-occur' package in R (Griffith et al., 2016). When species were detected more often than would be expected at random, they were considered positively correlated. When they occurred less often than would be expected at random, their co-occurrence was considered negative. Co-occurrence matrices were generated for species in the full data set and after filtering to each terrestrial habitat type. I hypothesized that Bahamian native species would partition the habitat and therefore have negative co-occurrence at locations on the island. Several native focal species such as the Northern Mockingbird have native congeners like Bahama Mockingbirds and Bahama Yellowthroats have migrant congeners Common Yellowthroats. I hypothesized that Bahamian species and DAS species will have positive co-occurrences within individual habitats.

Species-observer co-occurrences

Co-occurrence matrices were calculated between pairs of bird species and the observers who recorded them. When species showed random co-occurrence across observers, that indicates that there is no significant difference in the ability or interest of observers to detect the species; however, if species show negative co-occurrence, it suggests that an observer who records one species is less likely to record the other species. It may indicate a difference in the observer qualities required to record a species or groups of specialty birders such as birders that focus on shore birds. Similarly, when species positively co-occur significantly among observers, it indicates that the observer qualities that allow for the detection and recording of the two species are similar. These matrices were calculated for all observations and separately for each of the six terrestrial habitats. Results are presented for the four habitats indicated in the indicator species analysis. The results from the two matrix groups were compared to determine if there were specific species pairs with relationships that could be used to predict one or the other.

Observer co-occurrence

Observer co-occurrence was calculated at localities and among species. Co-occurrence matrices were generated for observers in the full data set and after filtering to each terrestrial habitat type. Observers that co-occurred positively at localities were more likely than random to conduct surveys at the same location. Positive co-occurrence for observers at localities may indicate some similar mode of introduction or accessing the sites or some process that filtered the observers into the sites. Cruise ship visitors all land at the same port in the Bahamas for example and a tour group traveling with the same bird guide would visit the same locations. Observers that co-occurred positively among species were more likely than random to record the same species. This would indicate that the observers had similar species preferences or interests or perhaps they coincidentally recorded the same very rare species.

Locality co-occurrence

Locality co-occurrence was calculated among species and observers. Co-occurrence matrices were generated for localities in the full data set and after filtering to each terrestrial habitat type. Localities that co-occurred positively among species were more likely than random to have the same species reported. Positive co-occurrence for localities among species may indicate some similar habitat or characteristic of a site that makes it useful to species groups. Features like ponds may cause localities to co-occur among waterfowl or shallow sand flat localities may co-occur across wading species. Localities that co-occurred positively among observers were more likely than random to have the same observers conduct surveys at the locality. This would indicate that the localities had similar amenities, infrastructure or marketing avenues that targeted similar observers. Beach and sandy shore localities for example would co-occur among researchers interested in wading shorebirds.

Results

Indicator species analysis

Indicator species analysis detected 19 significant indicator species that were specific to one of four habitats. The first group of species included Cuban Emerald, Gray Catbird, Black-throated Blue Warbler, Ovenbird, Black-throated Green Warbler, and Worm-eating Warbler. These species are primarily associated with grass habitats in the database. The second group included the Common Tern and White-throated Sparrow; these species were very rare but occurred in hwtc habitat. The third group included Pine Warbler, Olive-capped Warbler, Bahama Warbler, Black-faced Grassquit, Bahama Yellowthroat, Zenaida Dove, and Brown-headed Nuthatch. These species were associated primarily with pine habitats then grass. The fourth and final group included Piping Plover, Sanderling, Black-bellied Plover, and Wilson's Plover, which were primarily associated with sand habitats.

Species co-occurrence at Grand Bahama Localities

When all observer data was used to create co-occurrence matrices at the 50 most popular localities in the dataset there are few significant relationships. Bahama Mockingbirds, Bahama Yellowthroats, and Bahama Warblers all show negative co-occurrences with one another; Common Terns co-occur negatively with Bahama Yellowthroats and Bahama Warblers but do not have a significant co-occurrence with Bahama Mockingbirds (*Figure 21*).

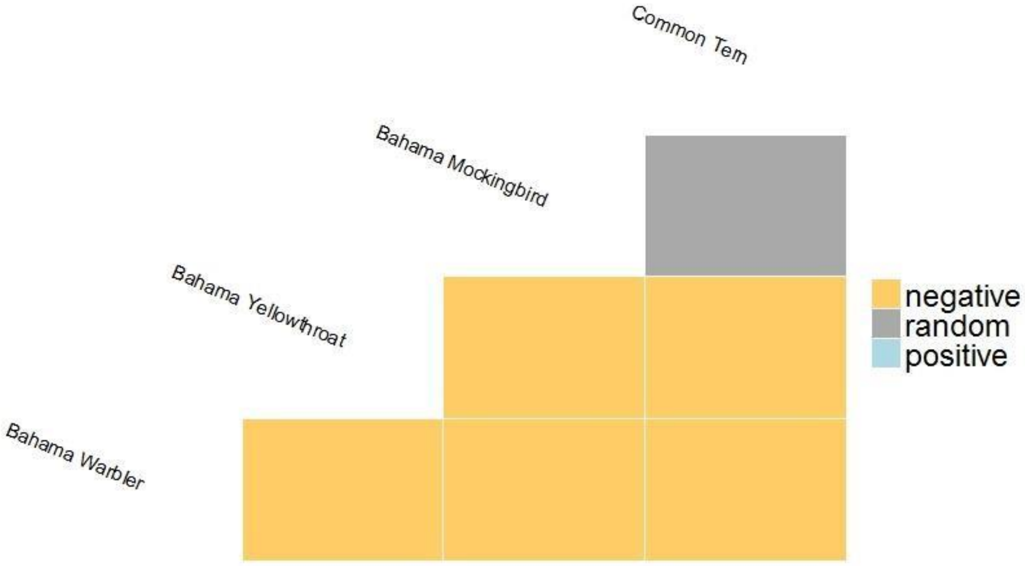


Figure 21 Species co-occurrence at the 50 most popular eBird localities on Grand Bahama. Using data from active observers between 1988 and 2016. Only significant relationships are shown.

Within the grass habitat, most of the Bahamian endemic species exhibited significant co-occurrences. Bahama Mockingbirds, Bahama Swallows, Bahama Warblers, Bahama Woodstars and Bahama Yellowthroats are present in the co-occurrence matrix (*Figure 22*); however, the Bahama Mockingbird no longer has a significant negative co-occurrence with the Bahama Warbler or Yellowthroat. This may be due to low numbers of occurrences or truly random distribution. Indigo Buntings are also present in the matrix. Along with the Bahama Woodstar and Bahama Swallow, the Indigo Bunting exhibits negative co-occurrence with all the Endemic species in this habitat.

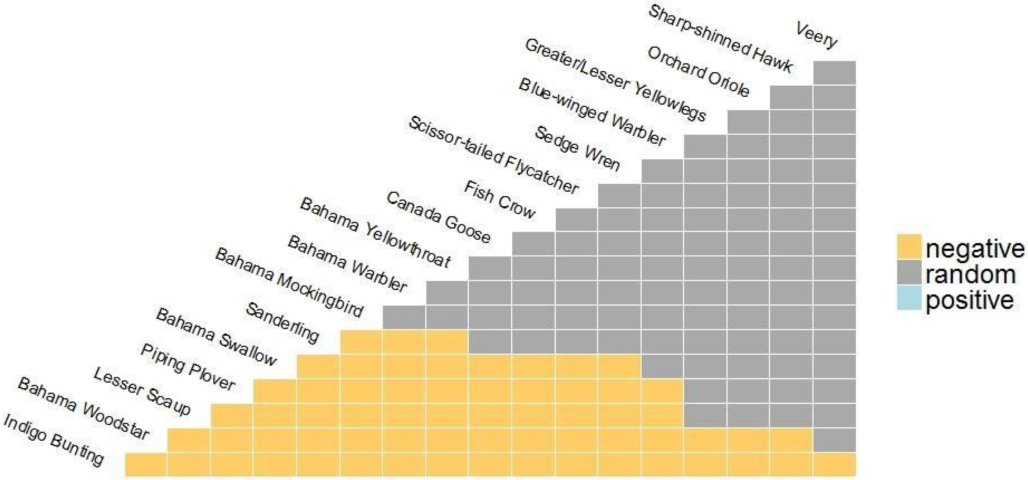


Figure 22 Species co-occurrence in grass localities on Grand Bahama only significant relationships are shown.

High Water Table Communities include American Kestrels in negative co-occurrence with Bahama Swallows, Bahama Warblers, Bahama Woodstars and Bahama Yellowthroats (*Figure 23*). The Gray Catbird is a North American winter migrant that shows negative co-occurrence with the American Kestrel and Bahama Swallows and Yellowthroats, but do not show a significant relationship with the Bahama Woodstar or Warbler in this habitat.

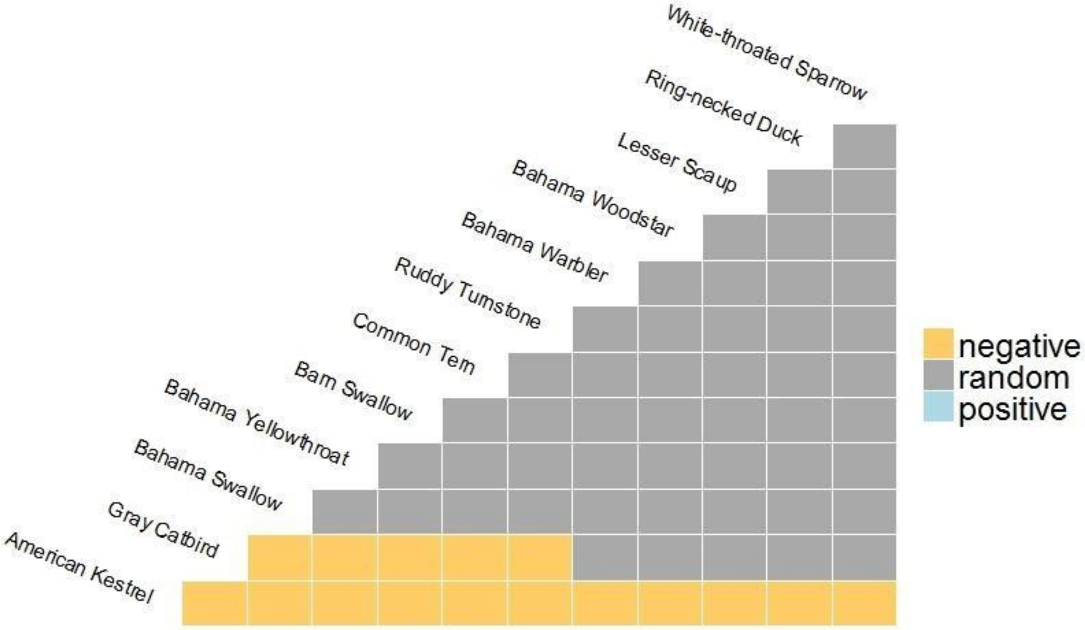


Figure 23. Species co-occurrence at High Water Table Community (hwtc) localities on Grand Bahama. Only significant relationships are shown.

The American Kestrel negatively co-occurs with several Bahamian endemic species in pine forest localities (*Figure 24*). These Bahamian Species include the Bahama Warbler, Bahama Swallow, and Bahama Yellowthroat. However, the relationship with Bahama Mockingbird is again random. The relationship with the indigo Bunting is also random, but the Bunting exhibits negative co-occurrence with Gray Catbirds and Bahama Warblers.

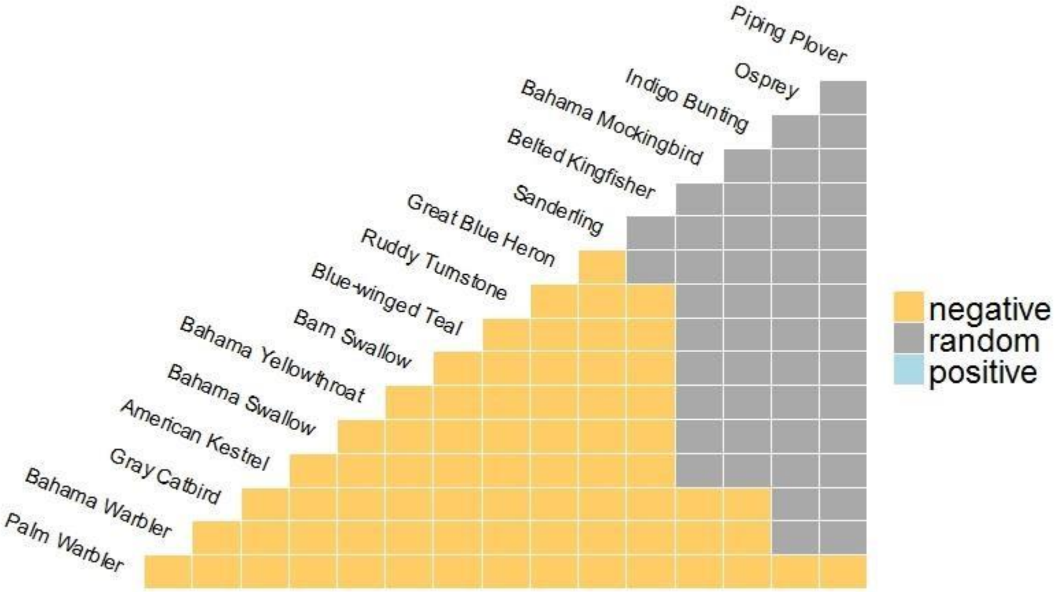


Figure 24. Species co-occurrence at pine localities on Grand Bahama. Only significant relationships are shown.

Species co-occurrence at sand localities showed very few significant interactions. However, the Bahama Yellowthroat, Bahama Swallow and Bahama Woodstar show significant relationships. Bahama Woodstars co-occurred negatively with the other two endemic species, but Yellowthroats and swallows did not have significant relationships. Indigo Buntings again showed a significant negative co-occurrence with all the endemic species.

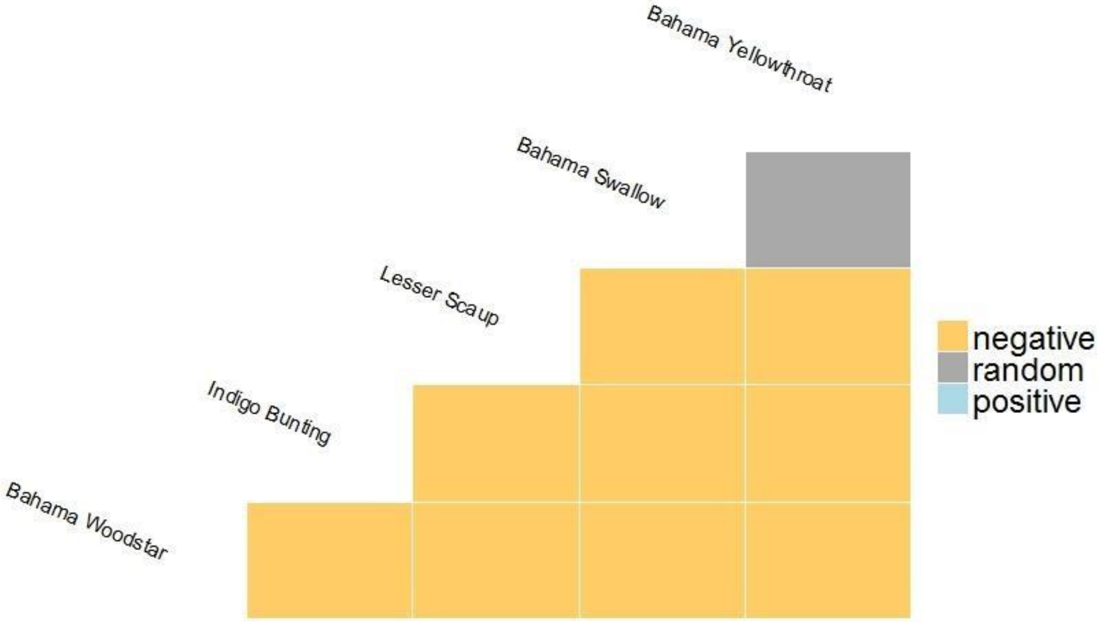


Figure 25. Species co-occurrence at sand localities on Grand Bahama. Only significant relationships are shown.

In urban localities, Bahama Swallows, Bahama Warblers, and Indigo Buntings show no significant negative correlations. However, the Bahama Woodstar negatively co-occur with Bahama Swallows and Indigo Buntings. Indigo Buntings are present with the endemic Bahama Mockingbirds, Swallows, Warblers, Woodstars and Yellowthroats in the wetland co-occurrence matrix. However, only the Indigo Bunting, Bahama Yellowthroat and Bahama Woodstar negatively co-occur significantly with one another. The other relationships between these species are not significant. When the data are filtered to include only the data from the top 50 observers and only the 50 most popular locations, The Bahama Yellowthroat, Bahama Warbler and Brown-headed Nuthatch negatively co-occur. The Bahama Mockingbird Negatively co-occurred with the Bahama Yellowthroat and Warbler, but they do not show significant co-occurrence with the Brown-headed Nuthatch.

Species and observer co-occurrences

When data collected by the top 50 observers are used to create species and observer co-occurrence matrices, there are many significant negative co-occurrences in each habitat. However, Pine habitat is the only habitat with a significant positive co-occurrence between any species and their observers: Ospreys and Belted Kingfishers (*Figure 26*). This means that in the 50 top observers, Ospreys and Belted Kingfishers were seen by the same observer, more often than would be expected at random. In the remaining habitat types, the Ospreys and Belted Kingfishers actually share a negative co-occurrence in wetland, urban and hwtc communities.

Some of the negative relationships seen in the locality co-occurrences were also present in the observer co-occurrence matrices however, some of the relationships changed. Bahama Yellowthroats and American Kestrels for example have negative co-occurrences in both the observer and locality matrices in pine habitat and the locality matrix for hwtc habitat, however there is no significant co-occurrence in any other matrix.

The Indigo Bunting co-occurs negatively with Bahama Yellowthroats in wetland and grass localities, but among observers negative co-occurrences only occur in grass, wetland and pine habitats. In addition, while the Indigo Bunting co-occurred negatively with the American Kestrel and Bahama Yellowthroat among observers, there was no significant relationship in the locality matrix. In the wetland habitat the Osprey and Belted Kingfisher relationship was reversed and they shared a negative co-occurrence in among observers. Indigo Buntings continue to share a negative co-occurrence with Bahama Yellowthroats in wetland among observers as in Pine, despite their lack of a significant co-occurrence among pine localities.

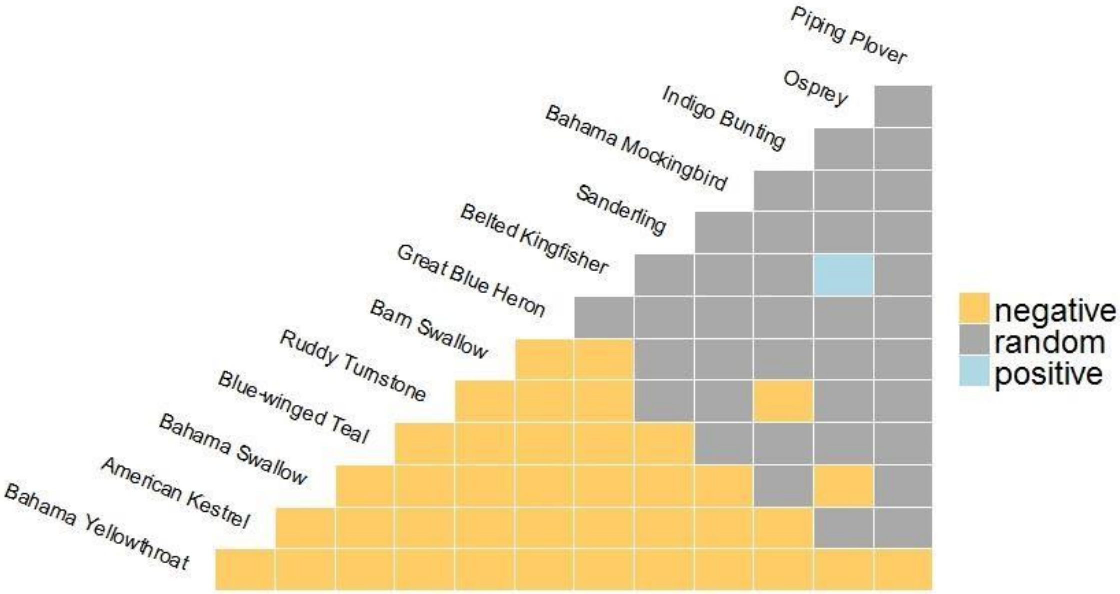


Figure 26. Species co-occurrence among observers within pine habitats on Grand Bahama, The Bahamas 1988-2016.

Locality Co-occurrences

When co-occurrence data are considered as localities in the data set where there is co-occurrence among species or observers, we can determine when localities support similar groups of species or observers. Species characteristics in fact change much more slowly than localities, especially when humans directly manage that locality in ways that impact species or observer access.

Within the wetland habitat, the Sharp Point Wetlands positively co-occur with Ensign Road and Sweeting Cay, however, Sweeting Cay and Ensign Road do not co-occur among species (*Figure 27*). Ensign Road is forested location near unpaved roads at a large intersection, with a few buildings in the vicinity, about 2 km from the nearest coastline. The sweeting Cay location is near an inland blue hole about 0.18 km away from the southern coast of Grand Bahama. The Sharp Point wetlands parallel the southern coast and are separated from a shallow bay by a narrow coastal road and sand bank less than 0.11 km from the coastline. Both the Sharp Point Wetlands and Sweeting cay have negative co-occurrences with Martha's Yard and the "Drive to West End" survey locations among species. Thirteen species have been reported on Ensign Road with 20 at the Sharp Point Wetlands and 40 on Sweeting Cay. Each of these three locations have only one observer. The same observer visited both Sharp Point Wetlands and Sweeting Cay and a different observer visited Ensign Road. Martha's yard had two observers that detected 68 total species while the Drive to West End had five observers who detected seven total species. The five observers on the Drive to west end were actually driving together as a group, which included a leading Bahamian bird guide and one of the observers from Martha's yard.

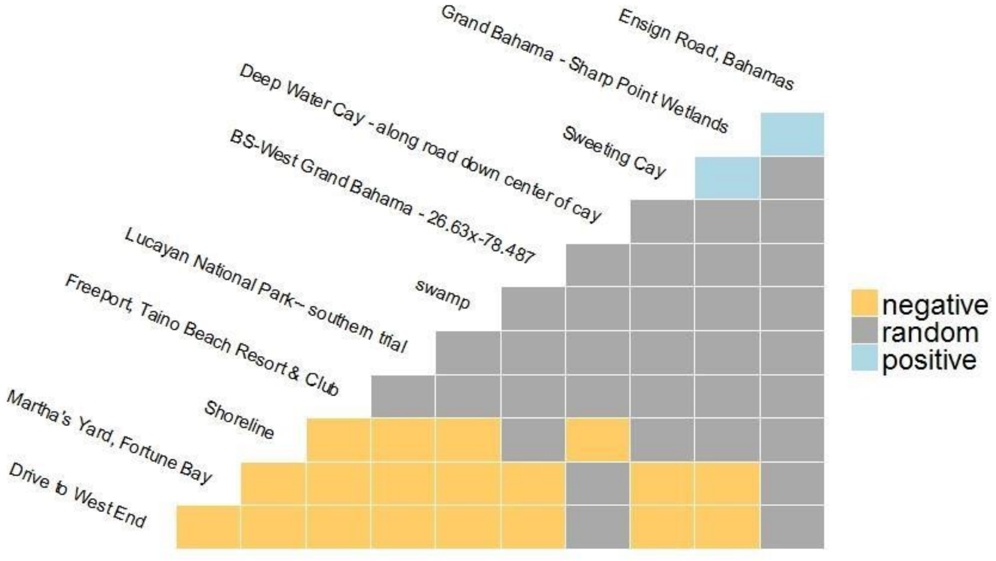


Figure 27. Locality co-occurrence among species for wetland habitat on Grand Bahama from eBird data 1988-2016

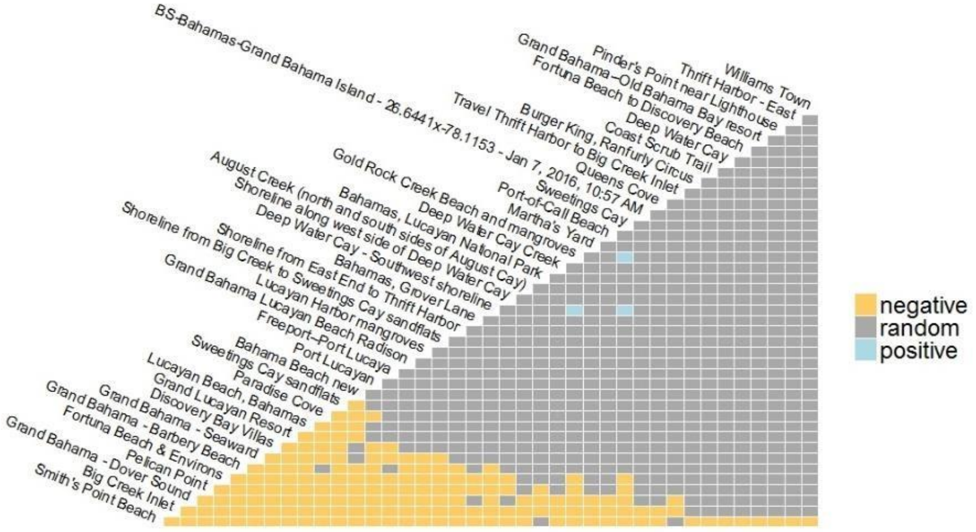


Figure 28. Locality co-occurrence among species for sand habitat on Grand Bahama from eBird data 1988-2016

As in the wetland habitat there are three sand localities with positive co-occurrences in species data (*Figure 28*). However, the three of these localities co-occur positively with one another and were each visited by a different observer. Sweetings Cay, Gold Rock Creek and

Beach and Wetlands and the Deep Water Cay South West Shoreline were reported as locations for 27, 19 and 13 species respectively. All three locations share four species (Great Blue Herons, Palm Warblers, Ruddy Turnstones and Turkey Vultures). Gold Rock Creek and Deep Water Cay respectively share five and three more species with Sweetings Cay.

Observer Co-occurrence at localities

The majority of significant positive co-occurrences between observers occur between local bird guides and their guests and other bird guides. One observer, a scientific researcher, has negative co-occurrences with the bird guides. This observer created more than twenty unique sites and was the only person to visit those sites.

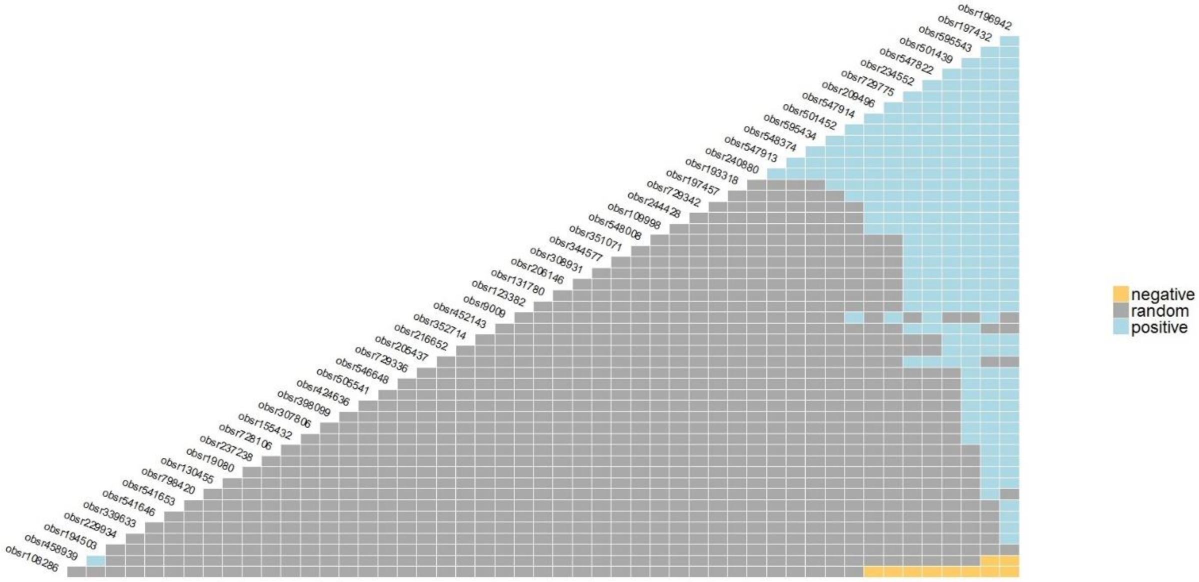


Figure 29. Observer co-occurrence matrix for locations on Grand Bahama, based on eBird data from 1988 to 2016.

Discussion

Indicator species analysis

The four indicator species groups include six species commonly found in thick brush and coppice or dense canopies these species may be more active in the early morning or cooler hours of the evening except in heavily shaded areas where their activity is more constant (pers. obs.). Viewing these species is most likely when the forests edge meets an open area so that observers can move for better views. These species were mostly associated with grassy areas such as parks landscaped properties and golf courses. The White-throated Sparrow and Common Tern are extremely rare in the data set. Their selection as indicator species is most likely a result of the combination of their mutual rarity and their coincidence at the locations at which they were detected they were found in hwtc communities but the number of occurrences does not make them likely candidates for monitoring. I recognize the third group of species as those that use an open understory forest or public area where large trees border open areas and walkways, though they use different parts of the trees and paths this group was mostly associated with the Pine habitat in the data base. The final group of plovers are distinctly a group of shorebirds associated with sand habitat. The piping plover in particular is an important species because many of the surveys included comments that referred to piping plover research activity being conducted in parallel with the survey.

Species co-occurrences

Several possibilities are available for birds in the dataset and their correlations. For example: birds regularly recorded during the same survey may have significant overlap in resource needs. Two species that are detected at the same location and during the same time of day, but never by the same observer may have a significant difference in their behavior, detectability, or may be easily confused causing observers to choose one or the other during identification. When two birds share the same locations, but are not detected at the same time of day or the same season of the year, the habitat is temporally partitioned by the species and this relationship may be useful in observing year-round resident bird distribution and using it to predict migrant species distribution during the winter months. This is especially useful when multiple species co-occur at locations with the target species.

This proposed methodology, can reduce the amount of time spent on *in situ* searches for the migrant species so that more time is available for collecting biometric and ecological data. The reduced cost and the utilization of citizen science data can then be used to support the development of local initiatives to educate or train observers in species monitoring.

Species co-occurrence at localities

Species co-occurrence relationships at localities vary among habitats. Few species are present in all habitats and those that coexist in multiple habitats may or may not maintain the same relationships across all habitats. Observers reported Indigo Buntings in all six terrestrial habitat types on Grand Bahama and it showed significant negative co-occurrence with other species in all but hwtc habitats. However, the Indigo Buntings did not show the same significant co-occurrences across all habitats. Significant negative co-occurrence with Indigo Buntings only occurred in one habitat for Gray Catbirds (pine) and Bahama Mockingbirds (Grass). Interestingly the Bahama Mockingbirds and Gray Catbirds only negatively co-occurred in pine habitat. The Buntings negatively co-occurred with the five endemic species (Mockingbirds, Swallows, Woodstars, Warblers and Yellowthroats) in the grass habitat localities. In the urban habitat, however, it only co-occurred with the Bahama Woodstar Hummingbird. The two species with which the Indigo Buntings shared the most negative co-occurrences were the Bahama Woodstar Hummingbird (wetland, sand, urban and grass) and the Bahama Yellowthroat (wetland, sand and Grass).

Species co-occurrence among observers

Species co-occurrence matrices were created among observers. Two species positively co-occurred when the likelihood of being recorded by the same observer was higher than would be expected if the co-occurrence was random. The Indigo Bunting showed no significant positive or negative co-occurrence with the Bahama Mockingbird or Bahama Warbler or the Gray Catbird in any of the terrestrial habitats. In urban habitats the Indigo Bunting co-occurred negatively with only the Bahama Woodstar Hummingbird. In sand habitats the Indigo Bunting negatively co-occurred with Bahama Yellowthroats, Woodstars and Swallows. In fact, only three species showed significant negative co-occurrence with indigo Buntings among observers. Those were the Bahama Yellowthroat (pine, wetland, and sand habitat); Bahama Woodstar (sand and urban Habitat) and Bahama Swallow (sand Habitat).

Positive co-occurrences

The only species pair to show positive co-occurrence among observers was the Osprey and Belted Kingfisher in the pine habitat. The Osprey and Belted Kingfisher only occurred in 3 and 9 localities respectively in the pine habitat and similarly were only seen by 4 and 14 observers respectively.

Species that share a habitat even if they use it at different times may have similar resource needs. Citizen Science data such as eBird bird detection data includes year round information on bird distribution. Year round data from native species and common migratory species can be used to identify high quality habitat for groups of birds. This information can then be used to select locations that are more likely to support populations of target migrant species. Here I look at species correlations at locations within the eBird dataset. Such co-occurrences indicate that two species use similar locations though they may partition them temporally. We then determine if those co-occurrences are the same across observers within the data set.

There was only one positive co-occurrence detected for any pair of species. The co-occurrence between Osprey and Belted Kingfishers among observers occurred over 4 observers who saw Ospreys at 3 localities on three dates in three separate years. Two of those observers happened to be surveying in the same location at the same time and were in fact, the only observers that conducted surveys in the pine habitat on that day. Those two observers also recorded the Belted Kingfisher, Turkey Vulture and Great Blue Heron, but never conducted another survey in the pine habitat on Grand Bahama Island. Likewise, only one observer conducted surveys on either of the other two dates. On January 1, 2001 a single observer recorded 12 species which included the Osprey, Bahama Warblers, Bahama Yellowthroats, Belted Kingfishers, Blue-winged Teals, Gray Catbirds, Great Blue Herons, Northern Mockingbirds, Palm Warblers, Ring-necked Ducks, Turkey Vultures, and Western/Eastern Wood Pewees. This observer also never returned to the pine habitat. On 25 December 2013 a single observer recorded an osprey. This observer returned to the pine forest twice that week on the 28th and 31st, but then never returned to the pine habitat and did not record another osprey though he recorded a Belted Kingfisher at the same location on a subsequent trip.

Several elements of observer behavior lead to this rare co-occurrence: timing, effort, group activity and habitat selection. These observations occurred during the winter season when Ospreys may be in migration (Wiley, Poole, & Clum, 2014), but also during times when it is unlikely for locals to conduct surveys. Three observers saw the ospreys on Wednesday during a workweek between 2:30 p.m. and 5:30 p.m. when local adults are at work. The other observer saw the osprey on Sunday, a holy day for most Bahamians when they either attend church or prefer to relax.

Negative co-occurrences

The negative co-occurrence across localities for the Bahama Yellowthroat and the two species American Kestrel and Indigo Bunting, may be simply explained. A single observer, who happened to be engaged in a study focused on the endemic birds of the pine forest (Lloyd, Slater, & Metcalf, 2007). This single observer entered 24 of the 41 surveys that included Bahama Yellowthroats at 24 unique locations. This observer also contributed to significant bias in the co-occurrences of the other endemic species. Unfortunately, not enough observers recorded endemic species to allow statistically significant correlations without this data. The endemic species were also not recorded in enough locations outside of this observer’s range for statistically significant correlations to be made. This observer conducted surveys for only one month in 2007 and contributed the majority of the records for the endemic species in our study between 1988 and 2016. The same observer contributed approximately 25% of the detection locations for Bahama Yellowthroats in the wetland and grass habitat where there were also negative co-occurrences across localities.

Species with no significant co-occurrences

Some species like the Brown-Headed Nuthatch (also Grand Bahama Nuthatch) did not show up in any of the co-occurrence matrices. The number of records were simply too low. This may be because of the bird’s behavior and the observers’ ability in general. Observers most commonly identify the Grand Bahama Nuthatch first by call then by sight. A rare endemic bird, the Nuthatch is often difficult to find and a significant specialty bird. The majority of the records for the species occurred during guided tours to “the Lucayan Estates Nuthatch Spot” and the Pine forests by Bahamian members of the BirdsCaribbean guide team during the 2011 regional meeting of BirdsCaribbean. The observers were professionals, experts and accomplished birders. Local guides found the birds prior to the surveys during which they were recorded. The expertise and enthusiasm of the birders reduced the likelihood that other species would find observer co-occurrence, within them, because these observers sought out, identified and recorded as many species as they could. The birds also occurred at so few locations that it is much less likely for a significant negative or positive co-occurrence to be detected.

Using species co-occurrences to detect proxy species may continue to be useful for some species, however, when using citizen science data it is essential that we know the observers' ability and motivation before we attempt to use the data. In the case of observers that record a single species during a survey, we need some way to ascertain that they did in fact search for all species or perhaps to engage them in a way that teaches them to avoid bias. Here, in the case of the endemic species, it was simple to determine the inherent bias in the records made by one observer. This observer contributed to the citizen science data, but also published research that aligned with these observations. Other observers may have different biases that are less apparent or not well documented.

Observer co-occurrence

The majority of observers visit few locations and often generate unique locations. Therefore, when the entire data set was used, the co-occurrence was seen as unlikely for virtually any co-occurrence. When the singletons were removed by narrowing the data set to only the top fifty locations, this greatly reduced the number of positive observer co-occurrences. However, this also made two important co-occurrence interactions much more evident. Bird guides positively co-occur with their guests and researchers are a distinct group that negatively co-occur with others.

The co-occurrence of bird guides with their visitors highlighted the very deliberate bias toward easy birding locations. Visitors often want to see a certain species or group of species. In unfamiliar surroundings many tourists may also be hesitant to venture out on their own to find a new birding location. Many of the co-occurring visitors were seen to only visit the locations the guide took them too. Several of the checklists also have group id's which indicates that not only did the guide take them there and choose the site, but the guide probably also completed the online entry of the species checklist for the trip.

The researcher that negatively co-occurred with the other active observers was particularly focused on pine habitat and targeting certain species within that habitat (Lloyd & Slater, 2011). This very particular behavior eliminated the possibility of significant co-occurrence with most observers who were visiting sand and grass habitat. He also created all of his locations based on scientific protocols, which do not take into account ease of access, driving distance or the availability of amenities for tourist guests. This makes the sampling locations he chose less biased geographically and perhaps more representative of the habitat, but less representative of the localities visited by typical observers.

Summary and General Discussion

The analyses completed here are meant to encourage the use of eBird data in new ways from new perspectives to not only support the conservation and study of the birds as wildlife but to also engage with the local communities on Grand Bahama and in other locations in the Bahamas, Caribbean and worldwide. The evidence provided here shows that the eBird data set is a useful tool not only for single bird species distribution but also for broader conservation of entire groups of species. Habitat manipulations to improve diversity at multiple locations and to encourage observer participation. Based on these analyses I present the following recommendations for monitoring groups, especially those that use eBird.

Engage locals to improve conservation monitoring and sustainability

Species monitoring and population assessments typically focus on a single species or group of species (Bernardo, Lloyd, Olmos, Cancian, & Galetti, 2011; J. Daniel Lambert, Thomas P. Hodgman, Edward J. Laurent, Gwenda L. Brewer, Marshall J. Iliff, 2009). Intense surveys and high cost research focus on a small area over a relatively short period (Burnett, 2013; Dreitz, Lukacs, & Knopf, 2006). Researchers detect the species of interest, and then work to collect significant amounts of data from a few individuals and the detection locations (Wunderle, Lebow, White, Currie, & Ewert, 2014). This framework is inefficient and unsustainable for migratory species protection and management on the wintering grounds such as in the Bahamas. Engaging local communities in monitoring is more efficient and can generate more if not better data.

The cost of bringing international researchers to the wintering grounds is prohibitive for most international funding mechanisms and even more so for local (wintering ground) resources. Travel and accommodation for visiting researchers can quickly amount to thousands of dollars for flights, hotel accommodations and the import of research equipment and work time lost during travel. Migratory species research is subject to artificial scheduling limits such as university or agency research calendars. Natural stochastic events such as tropical storms may also prevent researchers visiting the wintering grounds. Having local scientists in place reduces this risk to data collection and long term project sustainability and data quality.

The research presented here, above all else makes the case for improving local capacity and autonomy for conducting science on the wintering grounds. The use of local trainers to improve the habitat models will be invaluable to research on habitat use and occupancy by bird species, but local observers are essential in training such models quickly and accurately. Local observers can also respond quickly to update maps to represent stochastic changes in habitat after fires, hurricanes or human impacts. Local training of trainers to further develop local capacity is an important part of sustainable biodiversity monitoring and conservation efforts (J. Daniel Lambert, Thomas P. Hodgman, Edward J. Laurent, Gwenda L. Brewer, Marshall J. Iliff, 2009; Pattengill-Semmens & Semmens, 2003; Vanderstraete, Goossens, & Ghabour, 2005). The work presented here shows the important effect observer differences can have on the dataset. Standardization of observer training may balance out differences in ability and greatly reduce noise in the data (Johnston et al., 2018).

Citizen science data are a valuable tool for conservation planning, community engagement and education. eBird is a leader in the citizen science genre. Despite its broad availability and accessibility there are some improvements that can be made to increase its usefulness.

The addition of habitat data to the eBird dataset at a meaningful scale for small island developing states makes the data much more meaningful by presenting the species presence and absence data in context with the surrounding environment as needed by conservation practitioners (Brooks, Waylen, & Borgerhoff Mulder, 2012; Klingbeil & Willig, 2015). The methods presented here are efficient and adaptable, and though they are not perfect, they are easy to update. Appending the habitat data to the eBird data also showed the significant differences in the amount of activity between the habitats. Knowing these differences can help resource and wildlife managers to better plan for species and habitat conservation in the context of biodiversity and forestry resource management or even tourism development for ecotourism.

Acknowledge nuances in data

The comparison of the observer and species community distances with the geographical distances reinforces the unique format of the citizen science data. The data set does not necessarily follow the conventions of typical scientific research. Lack of standardization across observers introduces nuance and variation to the data set that we are not yet managing. Recognizing those differences in observers and their tangible effect on the data that is collected can be a weakness in the data set if not acknowledged, but when recognized, we can capitalize on the different abilities of the observers that are actively participating in the data set. The recommendations set forth here to assess, monitor and improve the ability of the individual observer can have significant effects in the long run. We can improve the data set, increase engagement with the platform and better support conservation and management of the species and natural resources. By confirming observer ability, we may improve the observers' skills on the individual level, but also create a more engaged community when we implement tiered skill levels and mentorship to allow users to engage with one another.

Change perspective on data

Repurposing the Mantel tests and Moran's I to view the data over observer and species dissimilarity is also intended to support the recognition of spatial dependence in species and observer data. While all the data collected may represent real world values, in the end, their meaning is dependent on our interpretations of what those data mean. I recognize each observer as a pseudolocation with its own qualities and characteristics, where each observer can be more or less similar to other observers just as research locations across the landscape vary in similarity. Similarly, I simultaneously recognize each bird species as 1) an organism seeking a location that holds the required resources and 2) a set of unique requirements, into which a location can fit. This allows us to re-envision the entire dataset. As humans, we actually have much more power over the configuration of our physical landscape than we do the needs of the species we try to conserve. Instead of chasing each species to find the ideal habitat for that species, when we interpret the species or species communities as goals for our localities, we can manipulate the habitat to achieve higher biodiversity. So now the data is not just a snapshot of the past, but a roadmap to the future where we can look at the locations that have our target species and improve other locations to approximate those characteristics.

Focus on the goals of monitoring

Finally, we best meet our ultimate goals, of predicting the location of a species and better conserving that species, when the species is considered as part of a broader community and that entire community is protected. The explorations of species relationships in the third chapter approach the goal of placing the species of interest in context with the other species in the eBird data set. Indigo Buntings and Gray Catbirds are not the highest priorities in migratory species protection. Their relationships set the foundation for explorations into other species and when conducted with greater computing resources we can expand those explorations to national levels in the Bahamas or other Caribbean nations. Most importantly, we can begin to integrate species managers do not normally consider in conservation plans, as monitoring proxies. This will improve comprehensive habitat and ecosystem protection. Holistic conservation plans that integrate multiple species will be more effective. When endemic native species are integrated into international conservation goals, it becomes more likely to engage local communities and succeed at the local level in smaller tight knit communities.

LITERATURE CITED

Barlow, K. E., Briggs, P. A., Haysom, K. A., Hutson, A. M., Lechiara, N. L., Racey, P. A., … Langton, S. D. (2015). Citizen science reveals trends in bat populations: The National Bat Monitoring Programme in Great Britain. *Biological Conservation*, *182*, 14–26. https://doi.org/10.1016/j.biocon.2014.11.022

Beals, E. W. (1984). Bray-curtis ordination: An effective strategy for analysis of multivariate ecological data. *Advances in Ecological Research*, *14*(C), 1–55. https://doi.org/10.1016/S0065-2504(08)60168-3

Bjornstad, O. N., & Cai, J. (2018). Package "ncf": Spatial Covariance Functions. Retrieved from https://cran.r-project.org/web/packages/ncf/ncf.pdf

Brooks, J. S., Waylen, K. A., & Borgerhoff Mulder, M. (2012). How national context, project design, and local community characteristics influence success in community-based conservation projects. *Proceedings of the National Academy of Sciences*, *109*(52), 21265–21270. https://doi.org/10.1073/pnas.1207141110

Brown, C. J., Bode, M., Venter, O., Barnes, M. D., McGowan, J., Runge, C. A., … Possingham, H. P. (2015). Effective conservation requires clear objectives and prioritizing actions, not places or species. *Proceedings of the National Academy of Sciences*, *112*(32), E4342–E4342. https://doi.org/10.1073/pnas.1509189112

Brudenell-Bruce, P. G. C. (1975). *The Birds of New Providence and the Bahama Islands* (1st ed.). New York, NY: Taplinger Publishing Co., Inc.

Butcher, G. S., & McCulloch, C. E. (1990). Influence of observer effort on the number of individual birds recorded on Christmas bird counts. In *Survey designs and statistical methods for the estimation of avian population trends* (pp. 120–129). https://doi.org/10.1016/B978-1-4377-1986-4.00034-2

Butcher, G. S., & Niven, D. K. (2007). Combining Data from the Christmas Bird Count and the Breeding Bird Survey to Determine the Continental Status and Trends of North America Birds. *National Audubon Society*, 34. Retrieved from https://www.researchgate.net/publication/237545005_Combining_Data_from_the_Christmas_Bird_Count_and_the_Breeding_Bird_Survey_to_Determine_the_Continental_Status_and_Trends_of_North_America_Birds

Chocron, R., Flather, C. H., & Kadmon, R. (2015). Bird diversity and environmental heterogeneity in North America: A test of the area-heterogeneity trade-off. *Global Ecology and Biogeography*, *24*(11), 1225–1235. https://doi.org/10.1111/geb.12353

Currie, D., Wunderle, J. M., Ewert, D. N., Davis, A., & McKenzie, Z. (2005). Winter avian distribution and relative abundance in six terrestrial habitats on southern Eleuthera, The Bahamas. *Caribbean Journal of Science*, *41*(1), 88–100.

Devictor, V., Whittaker, R. J., & Beltrame, C. (2010). Beyond scarcity: Citizen science programmes as useful tools for conservation biogeography. *Diversity and Distributions, 16*(3), 354–362. https://doi.org/10.1111/j.1472-4642.2009.00615.x

Dickinson, J. L., Zuckerberg, B., & Bonter, D. N. (2010). Citizen Science as an Ecological Research Tool: Challenges and Benefits. *Annual Review of Ecology, Evolution, and Systematics, 41*(1), 149–172. https://doi.org/10.1146/annurev-ecolsys-102209-144636

F. Dormann, C., M. McPherson, J., B. Araújo, M., Bivand, R., Bolliger, J., Carl, G., … Wilson, R. (2007). Methods to account for spatial autocorrelation in the analysis of species distributional data: A review. *Ecography, 30*(5), 609–628. https://doi.org/10.1111/j.2007.0906-7590.05171.x

Fortel, L., Henry, M., Guilbaud, L., Guirao, A. L., Kuhlmann, M., Mouret, H., … Vaissière, B. E. (2014). Decreasing abundance, increasing diversity and changing structure of the wild bee community (hymenoptera: anthophila) along an urbanization gradient. *PLoS ONE, 9*(8). https://doi.org/10.1371/journal.pone.0104679

Gagan, M. K., Ayliffe, L. K., Beck, J. W., Cole, J. E., Druffel, E. R. M., Dunbar, R. B., & Schrag, D. P. (2000). New views of tropical paleoclimates from corals. *Quarternary Science Reviews, 19,* 45–64. Retrieved from http://www.ingrid.ldgo.columbia.edu.

Gittleman, J. L., & Kot, M. (1990). ADAPTATION: STATISTICS AND A NULL MODEL FOR ESTIMATING PHYLOGENETIC EFFECTS. *Systematic Zoology*, *39*(3), 227–241. Retrieved from https://academic.oup.com/sysbio/article-abstract/39/3/227/1727758

Griffith, D. M., Veech, J. A., & Marsh, C. J. (2016). **cooccur** : Probabilistic Species Co-Occurrence Analysis in *R*. *Journal of Statistical Software*, *69*(Code Snippet 2), 1–17. https://doi.org/10.18637/jss.v069.c02

Groom, Q., Weatherdon, L., & Geijzendorffer, I. R. (2017). Is citizen science an open science in the case of biodiversity observations? *Journal of Applied Ecology*, *54*(2), 612–617. https://doi.org/10.1111/1365-2664.12767

Hargrove, L. J. (2010). Limits to Species ' Distributions : Spatial Structure and Dynamics of Breeding Bird Populations Along an Ecological Gradient A Dissertation submitted in partial satisfaction of the requirements for the degree of Doctor of Philosop, (March).

Hayes, W. K., Barry, R. X., McKenzie, Z., & Barry, P. (2004). Grand Bahama's Brown-headed Nuthatch: A Distinct and Endangered Species. *Bahamas Journal of Science, 12*(1), 21–28.

Hodgson, G., Kiene, W., Mihaly, J., Liebeler, J., Shuman, C., & Maun, L. (2004). *Reef Check Instruction Manual: A Guide to Reef Check Coral Reef Monitoring*. *Futures*. https://doi.org/10.3109/0142159x.2012.684914

Jenkins, C. N., Van Houtan, K. S., Pimm, S. L., & Sexton, J. O. (2015). US protected lands mismatch biodiversity priorities. *Proceedings of the National Academy of Sciences, 112*(16), 5081–5086. https://doi.org/10.1073/pnas.1418034112

Johnston, A., Fink, D., Hochachka, W. M., & Kelling, S. (2018). Estimates of observer expertise improve species distributions from citizen science data. *Methods in Ecology and Evolution, 9*(1), 88–97. https://doi.org/10.1111/2041-210X.12838

Joseph, E., Wunderle, J. M., Currie, D., Ewert, D. N., White, J. D., & Carey, E. (2009). A Brief History of the Kirtland's Warbler Research and Training Program. In *Proceedings of the 12th Symposium on the Natural History of the Bahamas* (pp. 59–65). San Salvador, Bahamas: Gerace Research Station. https://doi.org/ISBN 0-935909-89-3

Kelling, S., Johnston, A., Hochachka, W. M., Iliff, M., Fink, D., Gerbracht, J., … Yu, J. (2015). Can observation skills of citizen scientists be estimated using species accumulation curves? *PLoS ONE*, *10*(10). https://doi.org/10.1371/journal.pone.0139600

Klemann-Junior, L., Vallejos, M. A. V., Scherer-Neto, P., & Vitule, J. R. S. (2017). Traditional scientific data Vs. Uncoordinated citizen science effort: A review of the current status and comparison of data on avifauna in Southern Brazil. *PLoS ONE*, *12*(12). https://doi.org/10.1371/journal.pone.0188819

Klingbeil, B. T., & Willig, M. R. (2015). Bird biodiversity assessments in temperate forest: the value of point count versus acoustic monitoring protocols. *PeerJ*, *3*, e973. https://doi.org/10.7717/peerj.973

Kosub, S. (2016). *A note on the triangle inequality for the Jaccard distance*. Retrieved from https://arxiv.org/pdf/1612.02696.pdf

Lang, J. (2016). AGRRA : Introduction to Coral Transects Revision : 2016-09-22.

Lees, A. C., & Martin, R. W. (2015). Exposing hidden endemism in a Neotropical forest raptor using citizen science. *Ibis, 157*(1), 103–114. https://doi.org/10.1111/ibi.12207

Leitão, P. J., Moreiraa, F., & Osborned, P. E. (2011). Effects of geographical data sampling bias on habitat models of species distributions: A case study with steppe birds in southern Portugal. *International Journal of Geographical Information Science, 25*(3), 439–453. https://doi.org/10.1080/13658816.2010.531020

Lincoln, F. C. (1879). *Migration of Birds*. (P. A. Anatasi, Ed.), *Nature* (3rd, revis ed., Vol. 19). Manhattan, KS: Division of Biology, Kansas State University, Manhattan, KS. https://doi.org/10.1038/019481d0

Lloyd, J. D., & Slater, G. L. (2010). *Rapid Ecological Assessment of the avian community and their habitats on Andros , The Bahamas*. *Assessment*.

Lloyd, J. D., & Slater, G. L. (2011). Abundance And Distribution Of Breeding Birds In The Pine Forests Of Grand Bahama, Bahamas. *Journal of Caribbean Ornithology, 24*(1), 1–9. Retrieved from https://www.ecoinst.org/wp-content/uploads/2014/03/Lloyd_Slater_2011_Abundance_distribution_breeding_birds_grand_bahama.pdf

Loss, S. R., Loss, S. S., Will, T., & Marra, P. P. (2015). Linking place-based citizen science with large-scale conservation research: A case study of bird-building collisions and the role of professional scientists. *Biological Conservation*, *184*, 439–455. https://doi.org/10.1016/j.biocon.2015.02.023

Mantel, N., & Valand, R. S. (1970). A Technique of Nonparametric Multivariate Analysis. *Biometrics*, *26*(3), 547. https://doi.org/10.2307/2529108

Mantel, N., & Valand, R. S. (2018). A Technique of Nonparametric Multivariate Analysis Author (s): Nathan Mantel and Ranchhodbhai S . Valand Published by : International Biometric Society Stable URL : https://www.jstor.org/stable/2529108 International Biometric Society is collaborating wi. *Biometrics*, *26*(3), 547–558.

McArdle, B. H., & Anderson, M. J. (2001). Fitting Multivariate Models To Community Data : A comment on Distance-Based Redundancy Analysis. *Ecology*, *82*(1), 290–297.

McKinley, D. C., Briggs, R. D., & Bartuska, A. M. (2013). Reprint of: When peer-reviewed publications are not enough! Delivering science for natural resource management. *Forest Policy and Economics*, *37*, 9–19. https://doi.org/10.1016/j.forpol.2013.09.004

Merriam, C. H. (1898). *Life Zones and Crop Zones of the United States*. Washington. Retrieved from https://ia801900.us.archive.org/21/items/lifezonescropzon10merr/lifezonescropzon10merr.pdf

Miller, Z. D., Hallo, J. C., Sharp, J. L., Powell, R. B., & Lanham, J. D. (2014). Birding by Ear: A Study of Recreational Specialization and Soundscape Preference. *Human Dimensions of Wildlife*, *19*(6), 498–511. https://doi.org/10.1080/10871209.2014.921845

Moran, P. A. P. (1948). *The Interpretation of Statistical Maps*. *Source: Journal of the Royal Statistical Society. Series B (Methodological)* (Vol. 10). Retrieved from https://www.jstor.org/stable/pdf/2983777.pdf

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., Hara, R. B. O., … Wagner, H. (2015). Package “vegan” : Community Ecology Package. https://doi.org/ISBN 0-387-95457-0

Ormsby, A., & Mannle, K. (2006). Ecotourism benefits and the role of local guides at Masoala National Park, Madagascar. *Journal of Sustainable Tourism*, *14*(3), 271–287. https://doi.org/10.1080/09669580608669059

Ortega-Álvarez, R., Sánchez-González, L. A., Rodríguez-Contreras, V., Vargas-Canales, V. M., Puebla-Olivares, F., & Berlanga, H. (2012). Birding for and with people: Integrating local participation in avian monitoring programs within high biodiversity areas in Southern Mexico. *Sustainability*, *4*(9), 1984–1998. https://doi.org/10.3390/su4091984

Pattengill-Semmens, C. V., & Semmens, B. X. (2003). Conservation and management applications of the reef volunteer fish monitoring program. *Environmental Monitoring and*

Assessment, *81*(1–3), 43–50. https://doi.org/10.1023/A:1021300302208

Pattengill-Semmens, C. V, & Semmens, B. X. (1999). Assessment and Monitoring Applications of a Community-Based Monitoring Program: The Reef Environmental Education Foundation. *National Coral Reef Institute Meeting*, (April), 1–13.

Purcell, K. L. (2011). Long-term avian research at the San Joaquin Experimental Range: Recommendations for monitoring and managing oak woodlands. *Forest Ecology and Management*, *262*(1), 12–19. https://doi.org/10.1016/j.foreco.2010.07.039

Radabaugh, B. E. (1974). Kirtland's Warbler and its Bahama Wintering Grounds. *The Wilson Bulletin*, *86*(4), 374–383.

Raffaele, H., Wiley, J., Garrido, O., Keith, A., & Raffaele, J. (1998). *A Guide to the Birds of the West Indies*. Princeton, New Jersey: Princeton University Press.

Rapai, W. (2012). *The Kirtland's Warbler*. https://doi.org/10.3998/mpub.4072745

Roberts, D. W. (2016). Package "labdsv": Ordination and Multivariate Analysis for Ecology. Retrieved from https://cran.r-project.org/web/packages/labdsv/labdsv.pdf

Ross, M. S., Brien, J. J. O., & Flynn, L. J. (2010). Ecological Site Classification of Florida Keys Terrestrial Habitats Published by : The Association for Tropical Biology and Conservation Ecological Site Classification of Florida Keys Terrestrial Habitats1. *Audubon*, *24*(4), 488–502.

Scholes, R. J., Gill, M. J., Costello, M. J., Sarantakos, G., & Walters, M. (2017). *Working in Networks to Make Biodiversity Data More Available*. *The GEO Handbook on Biodiversity Observation Networks*. https://doi.org/10.1007/978-3-319-27288-7_1

Sfenthourakis, S., Tzanatos, E., & Giokas, S. (2005). Species co-occurrence: The case of congeneric species and a causal approach to patterns of species association. *Global Ecology and Biogeography*, *15*(1), 39–49. https://doi.org/10.1111/j.1466-822X.2005.00192.x

Sorenson, L., & Wunderle, J. M. (2017). First Photographic Record of Kirtland's warbler in Cuba!

Stier, A. C., Bolker, B. M., & Osenberg, C. W. (2016). Using rarefaction to isolate the effects of patch size and sampling effort on beta diversity. *Ecosphere*, *7*(12), 15. https://doi.org/10.1002/ecs2.1612

Sullivan, B. L., Wood, C. L., Iliff, M. J., Bonney, R. E., Fink, D., & Kelling, S. (2009). eBird: A citizen-based bird observation network in the biological sciences. *Biological Conservation*, *142*(10), 2282–2292. https://doi.org/10.1016/j.biocon.2009.05.006

Thogmartin, W. E., Gray, B. R., Gallagher, M., Young, N., Rohweder, J. J., & Knutson, M. G. (2007). Power to Detect Trend in Short-Term Time Series of Bird Abundance. *The Condor*, *109*(4), 943–948. Retrieved from http://www.jstor.org/stable/40072337

Tryjanowski, P., & Morelli, F. (2015). Presence of Cuckoo reliably indicates high bird diversity: A case study in a farmland area. *Ecological Indicators*, *55*(AUGUST), 52–58. https://doi.org/10.1016/j.ecolind.2015.03.012

Veech, J. A. (2013). A probabilistic model for analysing species co-occurrence. *Global Ecology and Biogeography*, *22*(2), 252–260. https://doi.org/10.1111/j.1466-8238.2012.00789.x

Wardle, C., Gape, L., & Moore, P. (2014). *Beautiful Bahama Birds: Comon Birds of the Bahama Islands* (first).

White, A. W. (1998). *A Birder's Guide to The Bahama Islands (Including Turks and Caicos)*. (V. Maynard, B. Berman, & H. Willoughby, Eds.) (First Edit). American Birding Association.

Wiley, J. W., & Wunderle, J. M. (1993). The effects of hurricanes on birds, with special reference to Caribbean islands. *Bird Conservation International*, *3*(4), 319–349. https://doi.org/10.1017/S0959270900002598

Wilkinson, C., Green, A., Almany, J., & Dionne, S. (2003). *Monitoring Coral Reef Marine Protected areas a Practical Guide on How Monitoring Can Support Effective Management of MPAs*. *Austral* (1st ed., Vol. 1). ian Institute of Marine Science.

Wunderle, J. M., & Arendt, W. J. (2017). The plight of migrant birds wintering in the caribbean: Rainfall effects in the annual cycle. *Forests*, *8*(4). https://doi.org/10.3390/f8040115

Wunderle, J. M., Currie, D., Helmer, E. H., Ewert, D. N., White, J. D., Ruzycki, T. S., … Kwit, C. (2010). Kirtland's Warblers in Anthropogenically Disturbed Early-Successional Habitats on Eleuthera, the Bahamas. *The Condor*, *112*(1), 123–137. https://doi.org/10.1525/cond.2010.090134