BAT POPULATION MONITORING IN NATIONAL PARKS OF THE GREAT LAKES REGION AND EVALUATION OF BAT ACOUSTIC ANALYSIS SOFTWARE

A Thesis
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Katy Rebecca Goodwin

In Partial Fulfillment of the Requirements
for the Degree of
MASTER OF SCIENCE

Major Program:
Environmental and Conservation Sciences

April 2019

Fargo, North Dakota

North Dakota State University
Graduate School

Title

Bat Population Monitoring in National Parks of the Great Lakes Region and Evaluation of Bat Acoustic Analysis Software

By

Katy R. Goodwin

The Supervisory Committee certifies that this ***disquisition*** complies with North Dakota State University's regulations and meets the accepted standards for the degree of

MASTER OF SCIENCE

SUPERVISORY COMMITTEE:

Dr. Erin Gillam
Chair

Dr. Mark Clark

Dr. Devan McGranahan

Approved:

4/10/2019
Date

Dr. Kendra Greenlee
Department Chair

ABSTRACT

North American bats face multiple threats, prompting an increase in bat research and conservation efforts in recent decades. Researchers often use acoustic monitoring, which entails recording bats' echolocation calls and subsequently identifying them to species, typically using automated software. Chapter 1 describes an acoustic monitoring program at eight U.S. national parks that aims to assess changes in bat populations over time. Data collected in 2016-2017 showed that activity levels of the little brown bat (*Myotis lucifigus*) decreased significantly while other species remained stable. Little brown bats have undergone similar population declines elsewhere due to the disease white-nose syndrome. Chapter 2 investigates whether different versions of bat call identification software are comparable to each other and how accurate they are. For the two software programs tested, agreement among versions was variable and species-dependent. Furthermore, newer versions were more conservative in assigning identifications, though not, on average, more accurate.

ACKNOWLEDGMENTS

I would like to thank my colleagues at the National Park Service Great Lakes Inventory and Monitoring Network, particularly Bill Route, Al Kirschbaum, and Rebecca Key, for their many important contributions to this project and their unfailing support. I would also like to acknowledge the dedicated efforts of numerous resource managers, biologists, and technicians who completed field data collection at each park. Several Northland College undergraduate students assisted with data management after the field season, notably Allison Boettcher, Taylor Pichler, and John Hermus. Additional collaborators include Karen Gilland of Normandeau Associates who provided data processing; Erik Olson of Northland College who contributed logistical and administrative support; and Lauren Hooton and Amanda Adams who generously shared their dataset of bat echolocation calls.

I appreciate the helpful feedback and guidance offered throughout this process by my advisor Erin Gillam, as well as committee members Devan McGranahan and Mark Clark. Particular thanks to Devan McGranahan and Raphaël Royauté for their assistance with statistical analyses. And finally, thank you to all my friends and family (near and far) for supporting and encouraging me these last two years.

Primary funding for this work was provided by the National Park Service. Additional support was received from the Biological Sciences Department and Environmental and Conservation Sciences Program at North Dakota State University.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

APIS .... Apostle Islands National Lakeshore

EPFU .... Big brown bat, *Eptesicus fuscus*

GLKN .... Great Lakes Inventory and Monitoring Network

GRPO .... Grand Portage National Monument

INDU .... Indiana Dunes National Park

ISRO .... Isle Royale National Park

LABO .... Eastern red bat, *Lasiurus borealis*

LACI .... Hoary bat, *Lasiurus cinereus*

LANO .... Silver-haired bat, *Lasionycteris noctivagans*

MISS .... Mississippi National River and Recreation Area

MYLU .... Little brown bat, *Myotis lucifugus*

MYSE .... Northern long-eared bat, *Myotis septentrionalis*

MYSO .... Indiana bat, *Myotis sodalis*

NPS .... National Park Service

NYHU .... Evening bat, *Nycticeius humeralis*

PESU .... Tri-colored bat, *Perimyotis subflavus*

SACN .... Saint Croix National Scenic Riverway

SLBE .... Sleeping Bear Dunes National Lakeshore

USFWS .... U.S. Fish and Wildlife Service

VOYA .... Voyageurs National Park

LIST OF APPENDIX TABLES

CHAPTER 1: BAT POPULATION MONITORING IN NATIONAL PARKS OF THE GREAT LAKES REGION

Abstract

Although widespread, common, and relatively diverse, North American bats are cryptic and difficult to study compared to other taxa. Baseline data regarding distributions, abundance, and use of the landscape are incomplete for many species. Over the past decade, a combination of threats including wind energy development, changing climatic conditions, and the fungal disease white-nose syndrome have prompted an increase in bat research and conservation efforts, including on public lands. At national parks in the Great Lakes region of the U.S., surveys targeting bats were relatively limited until a coordinated, comprehensive acoustic monitoring program was implemented in 2015. Acoustic monitoring is a widely-used method that entails recording echolocation calls of bats in their natural environment and subsequently identifying the calls to species, typically using automated software.

This chapter describes the bat acoustic monitoring effort I led at eight National Park Service sites in the Great Lakes region during 2016 and 2017, with the goal of assessing temporal changes in species-specific activity patterns. Six bat species were documented at all eight parks, while an additional three species were documented at one to a few parks each. I found that activity levels for eight of the nine study species remained stable across the two years. For the ninth species, the little brown bat (*Myotis lucifigus*), I observed a significant decrease in activity. This provides evidence that the little brown bat, which has already demonstrated severe population declines in the northeastern U.S. related to white-nose syndrome, is also declining in the Great Lakes region.

Introduction

Over the last few decades, North American bat populations have been increasingly threatened by a number of environmental pressures, prompting greater research, management, and conservation efforts. The most substantive threats facing bat communities include the disease white-nose syndrome, wind energy-related mortality, and changes in land use and climatic conditions.

White-nose syndrome is a disease caused by the fungal pathogen *Pseudogymnoascus destructans*. Both behavioral and physiological effects have been observed in infected bats, including fungal growth on the muzzle, ears, and wing membranes; increased frequency of arousals during hibernation; depletion of fat reserves and emaciation; and high rates of mortality (Blehert et al. 2009, Warnecke et al. 2012). In the 13 years since the pathogen was first documented in North America, it has spread to 33 U.S. states and 7 Canadian provinces (United States Geological Survey 2019), with an estimated total mortality of at least 6 million bats (United States Fish and Wildlife Service 2018*a*). Previous work has documented steep declines in local bat populations after the arrival of white-nose syndrome to an area. A variety of methods including winter hibernacula counts (Turner et al. 2011, Frick et al. 2015, Powers et al. 2015), summer capture surveys (Pettit and O'Keefe 2017), and summer acoustic surveys (Brooks 2011, Dzal et al. 2011) have all corroborated the declines and the disease continues to threaten new areas.

Wind energy development has grown substantially in the U.S. over the last two decades, reaching a current installed capacity of over 95,000 megawatts (American Wind Energy Association 2019). Estimates based on data prior to 2012 suggest 1-11 bats are killed per megawatt per (Arnett et al. 2015), or a total of 651,000-888,000 bat fatalities per year (Smallwood 2013), however current impacts are likely higher due to the continued increase in installed capacity. Although not all species are equally impacted (Kunz et al. 2007, Arnett et al. 2015), projections for one of the most affected species (hoary bat, *Lasiurus cinereus*) suggest that mortality due to wind energy could cause significant population declines and increased risk of extinction over the next 50-100 years (Frick et al. 2017).

Strong relationships have been observed between changing climatic conditions and some aspects of bat behavior (Frick et al. 2012) and physiology (Adams 2010). Modeling suggests climate change will lead to shifts in the geographic range of suitable conditions for both hibernacula (Humphries et al. 2002) and maternity colonies (Loeb and Winters 2013). Highly urbanized areas may have reduced species richness compared to nearby natural areas, though not necessarily a reduction in overall bat activity (Avila-Flores and Fenton 2005, Krauel and LeBuhn 2016). Specific factors associated with anthropogenic development, including impervious surfaces (Dixon 2012), roads (Kitzes and Merenlender 2014, Pourshoushtari et al. 2018) and artificial lighting (Cravens and Boyles 2019) have been shown to negatively impact bat activity, though varying responses are observed depending on species.

Given these many threats, bat researchers have expanded efforts to obtain baseline data and conduct population monitoring, in order to better understand bat communities and assess changes over time and space. Passive acoustic detection, in which ultrasonic recording devices are systematically deployed throughout a landscape to record the echolocation calls of nearby bats, is a commonly used method to monitor bats and gain information on species richness (Skalak et al. 2012), occupancy and detection probability (Gorresen et al. 2008), and relative activity levels (Ford et al. 2011). This method is more efficient and less invasive than capturing and handling bats, considerations that may be particularly important when studying bat populations that are already rare or in decline. Following data collection, echolocation call sequences can be identified to species using advanced acoustic software, such as SonoBat (J. Szewczak, www.sonobat.com) or Kaleidoscope Pro (Wildlife Acoustics, Inc., www.wildlifeacoustics.com). Despite the fact that call variation within species, overlap among species, recording quality, and other factors place limitations on the accuracy of identification (Barclay 1999, Frick 2013, Russo et al. 2018), this is still a widely accepted and useful method.

Over the past decade, due to the increase in conservation concerns, many U.S. federal agencies have begun implementing or expanding bat research and monitoring programs so that up-to-date scientific data is available to inform their management decisions (Loeb et al. 2015, Rodhouse et al. 2016). The National Park Service (NPS) has been very active in this area, funding over 150 bat-focused research, conservation, and education projects at 78 parks since 2013 (National Park Service 2016). In the Great Lakes region, the NPS conducted several bat surveys prior to 2013, but there was no coordinated, consistent, or comprehensive region-wide bat survey effort. These older surveys, though limited in scope, provided important occurrence data and documented three to six species per park through a combination of acoustic and capture methods (Kruger and Peterson 2008, Miller 2010, Goodwin 2012, Route and Schaberl 2013).

In 2015, a long-term bat monitoring project was established by the NPS Great Lakes Inventory and Monitoring Network (GLKN) and national parks around the Great Lakes region, with a particular focus on documenting the status of bat populations before and after the arrival of white-nose syndrome (Gruver and Rabie 2015). When the project was initiated, the Great Lakes region was at the leading edge of the white-nose syndrome spread, with occurrences of the disease within 50 miles of most parks (United

States Geological Survey 2019), making the fungal pathogen a very real and imminent threat. The monitoring project entailed passive acoustic surveys at multiple sampling locations in each participating park, resampled yearly. This allowed parks to document baseline data on their bat populations and assess their status and trends over time as white-nose syndrome continued to move west, potentially causing severe population declines. Passive acoustic surveys were conducted at five parks in 2015, nine parks in 2016, and ten parks in 2017 and 2018. The current analysis assesses bat activity during the 2016-2017 summer seasons at eight of the ten parks. Specifically, I address which species are present and where, whether bat activity levels are changing, and whether similar trends are observed for all species. My hypothesis is that declines in bat activity will be observed for the species that are most susceptible to white-nose syndrome, while other species activity levels will remain stable.

Methods

Acoustic surveys were conducted according to the protocols developed by GLKN in coordination with an outside consultant (Gruver and Rabie 2015, Gruver et al. 2016). Although the methods were revised after the initial season of monitoring, the majority of the sampling protocol was consistent across all years.

Study Area

Eight units managed by the NPS were included in this survey: Apostle Islands National Lakeshore (APIS), Grand Portage National Monument (GRPO), Indiana Dunes National Park (INDU), Isle Royale National Park (ISRO), Mississippi National River and Recreation Area (MISS), Saint Croix National Scenic Riverway (SACN), Sleeping Bear Dunes National Lakeshore (SLBE), and Voyageurs National Park (VOYA) (Figure 1). They are located in Minnesota, Wisconsin, Indiana, and Michigan; together, they cover a total land area of approximately one million acres (Great Lakes Inventory and Monitoring Network 2008). The parks encompass diverse habitat types including floodplain forest along major rivers, mixed coniferous/deciduous northern forest, dry pine forest, oak savanna, wetlands, lakes, ponds, and sand dunes (National Park Service 2018).

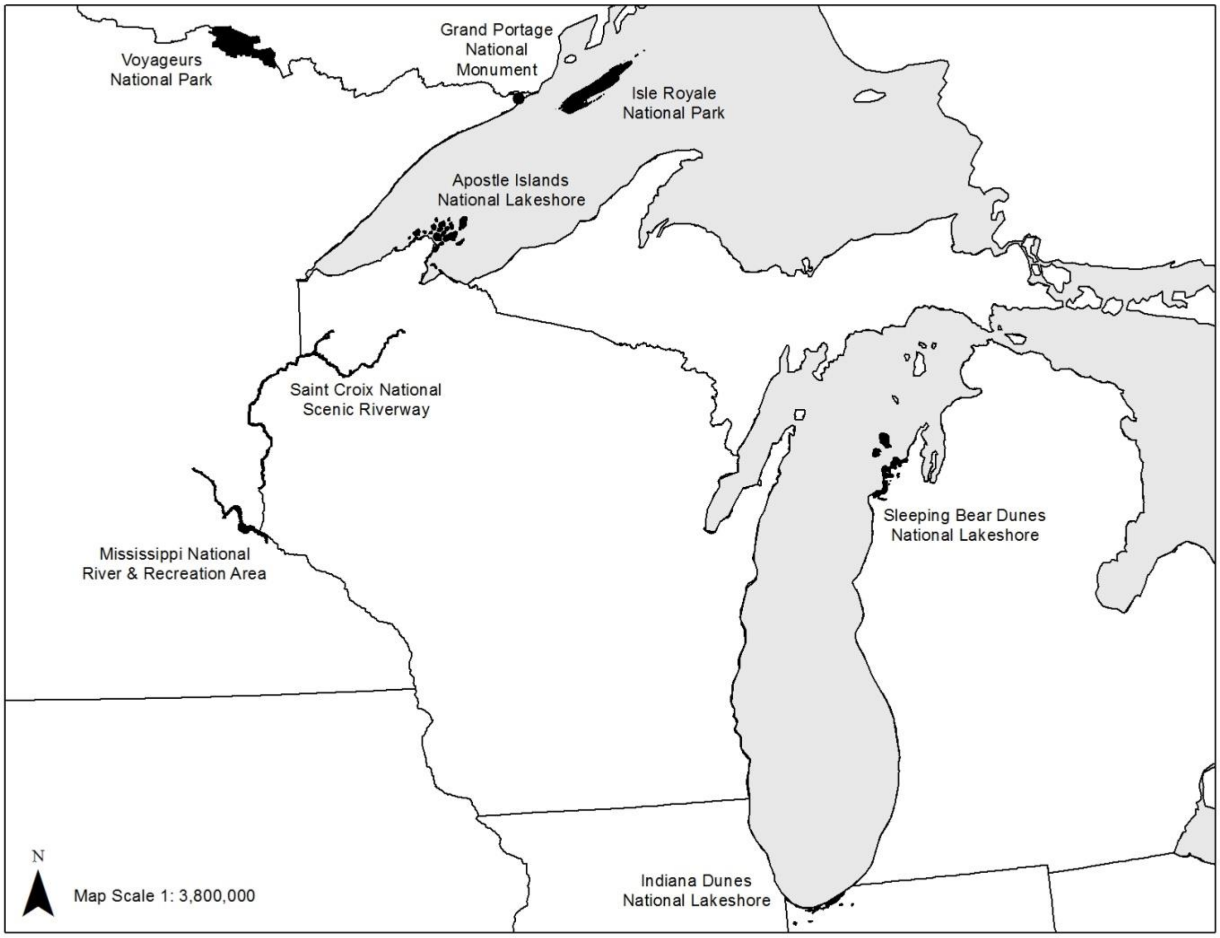

Figure 1. Location of eight national park units included in the study, within the upper Great Lakes region of the United States.

Nine species of bats are found in the Upper Midwest/Great Lakes region (Kurta 2017). All are insectivores belonging to the Family Vespertilionidae, the largest and most common group of bats in North America. Great Lakes bat species can be divided into two groups. The tree-roosting, migratory bats include the eastern red bat (*Lasiurus borealis*, LABO), hoary bat (*Lasiurus cinereus*, LACI), silver-haired bat (*Lasionycteris noctivagans*, LANO), and evening bat (*Nycticeius humeralis*, NYHU). The cavity-roosting, hibernating bats include the big brown bat (*Eptesicus fuscus*, EPFU), little brown bat (*Myotis lucifugus*, MYLU), northern long-eared bat (*Myotis septentrionalis*, MYSE), Indiana bat (*Myotis sodalis*, MYSO), and tri-colored bat (*Perimyotis subflavus*, PESU). All five hibernating species are confirmed to be susceptible to white-nose syndrome; additionally, red and silver-haired bats have been documented with *P. destructans*, though not showing signs of the disease itself (United States Fish and

Wildlife Service 2018*a*). The U.S. Fish and Wildlife Service (USFWS) lists the Indiana bat as federally endangered and the northern long-eared bat as federally threatened under the Endangered Species Act (United States Fish and Wildlife Service 1967, 2016). Furthermore, all nine species are listed by one or more states as endangered, threatened, or of special concern (Michigan Natural Features Inventory 2009, Minnesota Department of Natural Ressources 2013, Wisconsin Department of Natural Resources 2015, Indiana Division of Fish & Wildlife 2018).

Field Methods

Sampling locations were selected using generalized random-tessellation stratified sampling design, which is both randomized and spatially balanced (Stevens and Olsen 2004, Rodhouse et al. 2011, Loeb et al. 2015). A 1 km grid was overlaid on each park, grid cells were randomly numbered, and cells to be sampled were then selected by starting with Cell 001 and working through the list in consecutive order. Some cells were eliminated based on poor access, safety concerns, or lack of suitable habitat (e.g. cell fell in a lake). Within each sampled 1 km^2 cell, the specific sampling location was selected by identifying an area that was reasonably accessible and had appropriate habitat with relatively low clutter to optimize recording quality. This was typically done through a combination of examining aerial imagery and scouting potential sites on the ground.

Acoustic surveys were conducted during the summer between approximately June 1 and August 15 in 2016 and 2017. This is similar to the time period recommended by the USFWS for surveying for the two federally listed species (Indiana bat and northern long-eared bat) (United States Fish and Wildlife Service 2018*b*). Each site was sampled one time each year with a deployment period of 7-14 nights.

At each sample site, one Wildlife Acoustics Song Meter 3 or Song Meter 4 device was deployed to passively record bat echolocation calls. An ultrasonic microphone was attached either directly to a port on the device or via a connecting cable. Acoustic recordings were stored as WAV files on SD cards mounted in the Song Meter device. Photographs, GPS coordinates, and basic habitat information were collected for each deployment. The Song Meter device automatically generates a summary text file providing information on the device’s status throughout the deployment (battery voltage, number of files recorded, etc.). These files were retained alongside the associated WAV files.

Data Management and Call File Classification

Field data were reviewed for quality control and organized by park. All location and deployment information was stored in a Microsoft Access database developed by the NPS. An outside consultant hired by the NPS was provided with the database and the raw acoustic data in WAV file format. The consultant classified echolocation call files to species using the automated bat call identification software program Kaleidoscope Pro (version 4.0.0, classifier Bats of North America 3.1.0, "-1 More Sensitive" setting). Call files were excluded from species classification and all further analyses in the following situations: Deployments were excluded if they did not have at least four valid nights of recording, with a valid night defined as having a minimum of six hours of recording. The duration of recording was determined by examining the summary file automatically generated by the Song Meter, which indicates the times at which the device turned on at night and off in the morning. If the summary file was not available, duration was determined from the timestamps of the first and last audio files. Deployment length was cut off at a maximum of sixteen nights, with any call files recorded after the sixteenth night excluded from analysis. Deployments were also excluded if less than four nights fell within the protocol sampling period (June 1 – August 15). If at least four nights fell within these dates, and additional nights extended beyond, all nights in that deployment were included.

Kaleidoscope Pro was programmed to only allow certain candidate species during call file classification. The list of candidate species varied by park (ranging from 6-9 species) and was determined in advance by NPS biologists based on known occurrences and ranges (Table 1). Following automated classification, manual vetting was performed on 1% of call files (minimum of 10) for each park-species combination to verify identifications made by the software. Complete manual vetting of all call files was not feasible due to the large volume of data.

Statistical Analyses

Acoustic data do not provide a valid estimate of abundance, as there is no way to distinguish how many individual bats are producing the recorded echolocation call files (e.g. one bat calling ten times vs. ten bats calling one time each). Population trends were therefore analyzed using measures other than abundance, including species richness, occupancy, and activity levels. When conducting statistical analyses, a site was excluded from the year-to-year comparisons when the sampling location changed or when a deployment failed in one year or the other.

Table 1. Candidate bat species allowed for each park location during call file classification in Kaleidoscope Pro. An “X” indicates the species was allowed. Parks are located in Minnesota, Michigan, Wisconsin, and Indiana. Park and species abbreviations are provided on page vii.

					Species				
Park	EPFU	LABO	LACI	LANO	MYLU	MYSE	MYSO	NYHU	PESU
APIS	X	X	X	X	X	X			
GRPO	X	X	X	X	X	X			
INDU	X	X	X	X	X	X	X	X	X
ISRO	X	X	X	X	X	X			
MISS	X	X	X	X	X	X			X
SACN	X	X	X	X	X	X			X
SLBE	X	X	X	X	X	X			
VOYA	X	X	X	X	X	X			

Species richness was defined as the total number of species observed (Gotelli and Chao 2013), calculated per park. Occupancy was assessed by calculating at what proportion of sample sites a given species was recorded, for each park. Changes in occupancy were evaluated using the z-test for two proportions, with Yates’ continuity correction applied when the expected number of sites with detections and/or sites without detections was less than five.

Activity levels were measured using the metric of call files per deployment night. A “call file” here refers to an audio file that has been determined by the Kaleidoscope Pro software to contain bat echolocation sounds identifiable to species. A “call sequence” is defined as a series of vocalizations made by one individual bat as it passes the microphone (Loeb et al. 2015). Although it is possible for an audio file to contain multiple call sequences, for this study I am equating one call sequence to one call file. A deployment night is defined as the period from the evening of one day to the morning of the following day, when one acoustic detector was deployed. Bat activity levels were calculated for each species/site combination in each year by taking the total number of call files (over the entire deployment) identified to that species divided by the number of deployment nights at that site.

Statistical analyses were done in R (version 3.5.2) (R Core Team 2018). I used the *lmer* function in the *lme4* package (Bates et al. 2015) to perform linear mixed effects modeling in order to examine the relationship between bat activity levels and species. Preliminary model selection evaluated five models using *Park*, *Species*, and their interaction as possible fixed effects (Table A1). The response variable for all models was *Difference in Call Files per Deployment Night* (calculated as the call files per deployment night in 2017 minus the call files per deployment night in 2016). Models were compared using Akaike's Information Criterion corrected for small sample size (AICc), using the *aictab* function in the *AICcmodavg* package (Mazerolle 2017). I performed a Tukey pairwise comparison to contrast each pair of parks, using the *glht* function in the *multcomp* package (Hothorn et al. 2008). This resulted in only one significantly different pair (park locations INDU and GRPO, $z = 3.4$, $p < 0.05$). Based on the small difference between the top two models (*Park* + *Species* and *Species Only*, ΔAICc = 3.39) and the lack of explanatory power of the *Park* variable, the *Species Only* model was selected. The final version of the model utilized *Species* as the only fixed effect and *Site* nested within *Park* as the random effect, allowing for random intercepts but not random slopes (Table A1). After running the final model, I calculated 95% confidence intervals around the model coefficients for each species using the *confint* function in package *lme4* (Bates et al. 2015). In addition, two-sided multiple comparisons hypothesis testing was performed using the *glht* function in the *multcomp* package (Hothorn et al. 2008) in order to evaluate the null hypothesis that there was no difference in bat call files per deployment night between the two years, for each species.

Results

A total of 202 sites were sampled in 2016, while 206 sites were sampled in 2017, with a range of 17 to 35 sites per park (Table A2). For year-to-year comparisons, a total of 185 sites were retained that were surveyed in both years and met all protocol specifications. Mean number of recording nights per site was 9.4 (range 5 to 16) in 2016 and 8.8 (range 4 to 16) in 2017. After filtering out noise and bat recordings that were not identifiable to species, the total number of call files with species classifications was 384,032 in 2016 and 312,496 in 2017. Manual vetting was performed on 4,060 call files from 2016 and 3,354 call files from 2017 (Table A3). Automated call classification results showed that all expected species (Table 1) were documented for every park in both years. Species richness per park ranged from six to nine species, and did not differ from 2016 to 2017.

Multiple comparisons testing showed little brown bat activity had a significant decline from 2016 to 2017 ($p < 0.0001$). For the remaining species, p-values were not significant and confidence intervals overlapped zero (except eastern red bat), indicating a lack of difference in activity level between the two years (Figure 2, Table 2).

Table 2. Model predictions for final model relating *Difference in Call Files per Deployment Night* to *Species* with *Site*/*Park* as nested random effects. Multiple comparisons results for each of the nine observed bat species and 95% confidence intervals are also shown. Significant p-value indicated in bold.

Species	Coefficient	Lower CI	Upper CI	z-value	p-value
Big brown bat	8.29959	-1.26802	17.69105	1.746	0.5131
Eastern red bat	-12.3017	-21.8693	-2.91022	-2.588	0.0816
Hoary bat	0.05076	-9.51685	9.442223	0.011	1.0000
Silver-haired bat	3.32668	-6.24093	12.71814	0.7	0.9968
Little brown bat	-42.1184	-51.686	-32.7269	-8.86	**<0.0001**
Northern long-eared bat	-5.07552	-14.6431	4.315937	-1.068	0.9441
Indiana bat	-10.1236	-34.4017	13.90337	-0.847	0.9873
Evening bat	-13.8098	-38.088	10.21714	-1.156	0.9124
Tri-colored bat	-4.7863	-19.9159	9.997615	-0.638	0.9984

Considering only the group of parks where a given species was expected (Table 1), the overall percentage of sample sites at which that species was detected ranged from 84 to 100% in 2016 and from 69 to 100% in 2017 (Tables A4, A5). The proportion of sites with detections in 2016 vs. 2017 was compared for each park/species combination. Significant changes were observed for red bats at APIS (Chi-squared = 4.9, df = 1, $p < 0.05$) and northern long-eared bats at GRPO (Chi-squared = 6.5, df = 1, p

< 0.05). In both cases, the proportion of sites with detections was smaller in 2017. For the remaining park/species combinations there were no significant differences in the proportion of sites with detections across years. In combination with the linear mixed effects model results above, this indicates that little brown bat activity was reduced overall, but the species was detected at a similar number of sample sites across the two study years.

Discussion

North American bat populations have been on the decline for over a decade, largely due to the rapid spread and high mortality rates of white-nose syndrome. In the Great Lakes region, the first occurrences of the disease were in eastern Ontario in the winter of 2010-2011. The following year, *P. destructans* was detected for the first time in Minnesota. By the spring of 2015, when this project was initiated, white-nose syndrome was confirmed or suspected in over 30 counties throughout the project area (Michigan, Indiana, Wisconsin, and Minnesota). Currently in 2019, that number has jumped to over 60 counties and the disease has moved on to several western states (United States Geological Survey 2019).

The main goals of this project were to provide baseline data describing the bat populations of national parks around the Great Lakes region, and to track trends in these populations as white-nose syndrome began impacting the area. Prior NPS studies have been very limited in scope and basic knowledge about species presence/absence, distributions, and relative abundance was lacking or incomplete. Silver-haired, little brown, and northern long-eared bats were all previously documented at four of the parks (APIS, GRPO, ISRO, and VOYA) through mist-netting and acoustic survey efforts. Similarly, big brown, red, and hoary bats had all been previously documented at three of the parks (APIS, GRPO, and ISRO). However, bat data for the other parks in this project (MISS, SACN, SLBE, and INDU) has not been reported on.

Results from two years of the monitoring program (2016-2017) suggest that all expected species (Table 1) are present and most are widely distributed throughout each park (Tables A2, A3). Manual vetting verified that over 60% of the call files were correctly identified for four species: red bat, hoary bat,

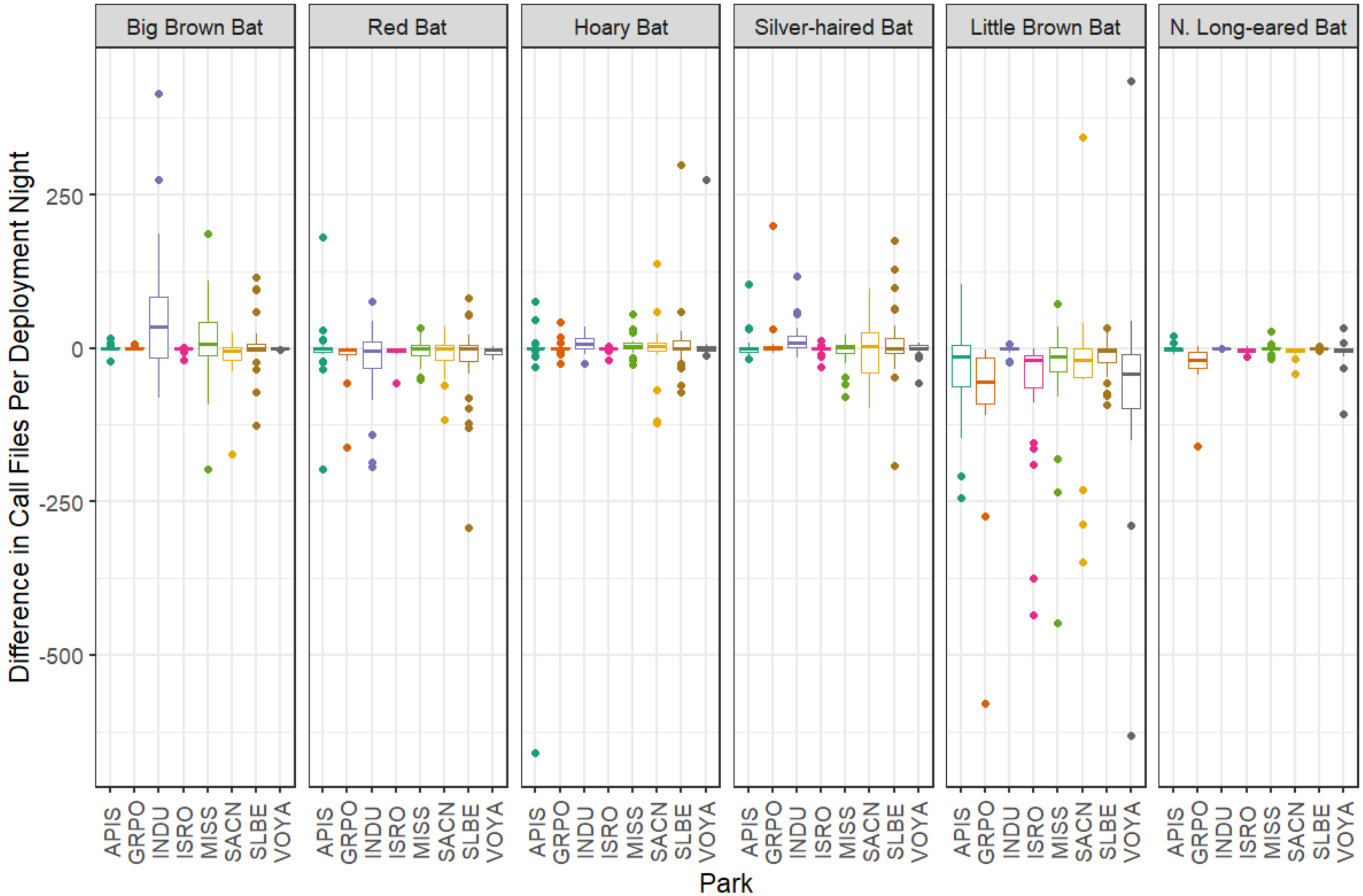

Figure 2. Difference in call files per deployment night by species and park. Difference was calculated by subtracting 2016 values from 2017 values, such that a positive difference indicates increased activity in 2017 and negative difference indicates decreased activity in 2017. Data only shown for the species common to all parks. Park abbreviations are provided on page vii.

little brown bat, and tri-colored bat (Table A4). This high level of agreement increases the confidence that these species are in fact present. However, for the remaining species there is more uncertainty.

Big brown and silver-haired bats can produce calls with many overlapping characteristics and are thus hard to distinguish (Betts 1998). During the manual vetting process, the reviewer was permitted to assign call files to species groups when a species-specific identification could not be made with confidence. Although there are lower percentages of exact agreement between the manual reviewer and the software (only 35% for big brown bat and 53% for silver-haired bat), this is a result of over 1,200 call files being assigned to the "EPFU/LANO" group (Table A4). It is not known which species produced these calls; however it can at least be narrowed down to these two species that have similar echolocation call structures.

Similarly, species in the genus *Myotis* (little brown bat, northern long-eared bat, and Indiana bat) produce very similar calls and some recordings may not be identifiable to species due to the high degree of overlap (Ratcliffe and Dawson 2003, Broders et al. 2004). Manual vetting results showed high agreement for the little brown bat (82%), low agreement for the northern long-eared bat (23%), and no agreement for the Indiana bat (Table A4). My dataset had a relatively small number of calls assigned to the latter two species. For example, across the two years at INDU, only 48 call files were classified as northern long-eared bats and 157 call files were classified as Indiana bats. The infrequency of detection in combination with the uncertainty due to overlapping call characteristics suggests caution should be taken when interpreting the results for these species. To increase certainty of their presence, a manual review of all calls assigned to these species could be conducted in the future.

I hypothesized that I would detect declines in bat activity for the five species that are most affected by white-nose syndrome: big brown bats, little brown bats, northern long-eared bats, Indiana bats, and tri-colored bats (United States Fish and Wildlife Service 2018*a*). Model results indicated a significant decline in little brown bat activity from 2016 to 2017, but no declines for the other species. For little brown bats, the raw number of call files (including only sites with valid deployments in both years at the same location) declined from 131,630 in 2016 to 59,621 in 2017, a drop of almost 55%. However, I did not observe any difference in the geographic distribution of the little brown bat, as it was detected at a similar number of sample sites in both years.

Other studies that used acoustic methods during the summer period have also documented significant, rapid declines following the arrival of white-nose syndrome. Surveys in New York found that little brown bats decreased by 78% over a one year period (Dzal et al. 2011) while in Massachusetts a decline of 72% was observed for all *Myotis* species combined (Brooks 2011). Researchers using non-acoustic methods have obtained similar results. A decline of 88% was observed over three years of emergence counts at a maternity colony in New York (Dobony and Johnson 2018). A long-term mist-netting effort in Indiana documented a decline in captures per unit effort of almost 80% for the little brown bat when comparing pre- and post-white-nose syndrome periods (Pettit and O'Keefe 2017). The observed decline in the current analysis is therefore not surprising in the context of this species' population trends in other white-nose affected areas, and is in fact less severe than comparable studies.

In addition to the little brown bat, four other species in the Great Lakes region are also susceptible to white-nose syndrome (the big brown bat, northern long-eared bat, Indiana bat, and tri-colored bat), yet I did not observe significant changes in their activity levels in this dataset. I did observe a reduction in the number of sample sites with detections for northern long-eared bats at one park (Grand Portage). With the exception of the big brown bat, surveys in other regions have often found major declines in these species due to white-nose syndrome, of a similar magnitude to declines in the little brown bat. For example, hibernacula surveys showed decreases in Indiana, tri-colored, and northern long-eared bats ranging from 16% - 99% (Langwig et al. 2015, Powers et al. 2015). Based on capture data, declines were documented in Indiana bats (59.6%) and tri-colored bats (12.5%), but not in northern long-eared bats (Pettit and O'Keefe 2017). Acoustic surveys also revealed declines in the Indiana bat and the northern long-eared bat, but no change for the tri-colored bat (Ford et al. 2011). The lack of observed declines in my own results for these species may be influenced by the relatively small sample size, as call files assigned to these species comprised less than five percent of the total call files each year. As additional years of data are included in the analysis, we may be better able to detect trends for these less common species.

For big brown bats, past studies aimed at examining the effects of white-nose syndrome have either not included the species (Dzal et al. 2011, Langwig et al. 2015, Powers et al. 2015, Dobony and Johnson 2018), not found any significant change (Brooks 2011, Ford et al. 2011), or found a slight increase (Pettit and O'Keefe 2017). A survey specifically focused on big brown bats suggests that this species, though susceptible, is less severely impacted by white-nose syndrome because it is able to maintain higher body fat content and normal torpor/arousal patterns through the hibernation period (Frank et al. 2014). My data did not show any significant decrease in big brown bat activity, supporting the idea that this species has greater resistance to the disease as compared to the three *Myotis* species and the tri-colored bat.

My findings did not detect any significant changes in activity level for the remaining four species present in the study area (red, hoary, silver-haired, and evening bats). These species have not been documented with symptoms of white-nose syndrome and are not thought to be impacted by the disease because they are migratory rather than hibernators (United States Fish and Wildlife Service 2018*a*). I did, however, observe a reduction in the number of sample sites with detections for red bats at one park (Apostle Islands). Red bats, along with silver-haired and hoary bats, are highly susceptible to mortality from wind turbines (Kunz et al. 2007, Frick et al. 2017) and could experience declining populations as wind energy development increases.

Although this project only assessed two years of data, it has already helped fill in information gaps by providing further evidence to verify species presence at the parks. My results also show a serious decline in the little brown bat, which is likely related to the arrival of white-nose syndrome to the region, but no significant changes for any other species. Although much work has been done in eastern states where white-nose syndrome was originally detected, this study is among the first to document changing bat populations on a regional scale in the Midwest. Future work will expand on this analysis to include acoustic data from all years 2015-2019, allowing greater power to detect population trends, even for the rarer species. This information will be invaluable to park natural resource staff as they consider management decisions and implement bat conservation efforts.

Literature Cited

Adams, R. A. 2010. Bat reproduction declines when conditions mimic climate change projections for western North America. Ecology 91:2437–2445.

American Wind Energy Association. 2019. Wind Facts at a Glance. <https://www.awea.org/wind-101/basics-of-wind-energy/wind-facts-at-a-glance>. Accessed 25 Feb 2019.

Arnett, E. B., E. F. Baerwald, F. Mathews, L. Rodrigues, A. Rodriguez-Duran, J. Rydell, R. Villegas-Patraca, and C. C. Voigt. 2015. Impacts of Wind Energy Development on Bats: A Global Perspective. Pages 295–323 in C. C. Voigt and T. Kingston, editors. Bats in the Anthropocene: Conservation of Bats in a Changing World.

Avila-Flores, R., and M. B. Fenton. 2005. Use of spatial features by foraging insectivorous bats in a large urban landscape. Journal of Mammalogy 86:1193–1204.

Barclay, R. M. R. 1999. Bats are not birds--A cautionary note on using echolocation calls to identify bats: A comment. Journal of Mammalogy 80:290–296.

Bates, D., M. Mächler, B. Bolker, and S. Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software 67.

Betts, B. J. 1998. Effects of interindividual variation in echolocation calls on identification of big brown and silver-haired bats. Journal of Wildlife Management 62:1003–1010.

Blehert, D. S., A. C. Hicks, M. Behr, C. U. Meteyer, B. M. Berlowski-Zier, E. L. Buckles, J. T. H. Coleman, S. R. Darling, A. Gargas, R. Niver, J. C. Okoniewski, R. J. Rudd, and W. B. Stone. 2009. Bat white-nose syndrome: An emerging fungal pathogen? Science 323:227.

Broders, H. G., C. S. Findlay, and L. Zheng. 2004. Effects of clutter on echolocation call structure of Myotis septentrionalis and M. lucifugus. Journal of Mammalogy 85:273–281.

Brooks, R. T. 2011. Declines in summer bat activity in central New England 4 years following the initial detection of white-nose syndrome. Biodiversity and Conservation 20:2537–2541.

Cravens, Z. M., and J. G. Boyles. 2019. Illuminating the physiological implications of artificial light on an insectivorous bat community. Oecologia 189:69–77. Springer Berlin Heidelberg.

Dixon, M. D. 2012. Relationship between land cover and insectivorous bat activity in an urban landscape. Urban Ecosystems 15:683–695.

Dobony, C. A., and J. B. Johnson. 2018. Observed resiliency of little brown myotis to long-term white-nose syndrome exposure. Journal of Fish and Wildlife Management 9:102017-JFWM-080.

Dzal, Y., L. P. McGuire, N. Veselka, and M. B. Fenton. 2011. Going, going, gone: The impact of white-nose syndrome on the summer activity of the little brown bat (Myotis lucifugus). Biology Letters 7:392–394.

Ford, W. M., E. R. Britzke, C. A. Dobony, J. L. Rodrigue, and J. B. Johnson. 2011. Patterns of acoustical activity of bats prior to and following white-nose syndrome occurrence. Journal of Fish and Wildlife Management 2:125–134.

Frank, C. L., A. Michalski, A. A. McDonough, M. Rahimian, R. J. Rudd, and C. Herzog. 2014. The resistance of a North American bat species (Eptesicus fuscus) to white-nose syndrome (WNS). PLoS ONE 9:1–14.

Frick, W. F. 2013. Acoustic monitoring of bats, considerations of options for long-term monitoring. THERYA 4:69–78.

Frick, W. F., E. F. Baerwald, J. F. Pollock, R. M. R. Barclay, J. A. Szymanski, T. J. Weller, A. L. Russell, S. C. Loeb, R. A. Medellin, and L. P. McGuire. 2017. Fatalities at wind turbines may threaten population viability of a migratory bat. Biological Conservation 209:172–177.

Frick, W. F., S. J. Puechmaille, J. R. Hoyt, B. A. Nickel, K. E. Langwig, J. T. Foster, K. E. Barlow, T. Bartonička, D. Feller, A. J. Haarsma, C. Herzog, I. Horáček, J. van der Kooij, B. Mulkens, B. Petrov, R. Reynolds, L. Rodrigues, C. W. Stihler, G. G. Turner, and A. M. Kilpatrick. 2015. Disease alters macroecological patterns of North American bats. Global Ecology and Biogeography 24:741–749.

Frick, W. F., P. M. Stepanian, J. F. Kelly, K. W. Howard, C. M. Kuster, T. H. Kunz, and P. B. Chilson. 2012. Climate and weather impact timing of emergence of bats. PLoS ONE 7:e42737.

Goodwin, K. R. 2012. Acoustic and Visual Surveys of Bats (Chiroptera) in Isle Royale National Park. Unpublished National Park Service report.

Gorresen, P. M., A. C. Miles, C. M. Todd, F. J. Bonaccorso, and T. J. Weller. 2008. Assessing bat detectability and occupancy with multiple automated echolocation detectors. Journal of Mammalogy 89:11–17.

Gotelli, N. J., and A. Chao. 2013. Measuring and Estimating Species Richness, Species Diversity, and Biotic Similarity from Sampling Data. Encyclopedia of Biodiversity. Volume 5. Elsevier Ltd.

Great Lakes Inventory and Monitoring Network. 2008. The Great Lakes Inventory and Monitoring Network.

Gruver, J., and P. Rabie. 2015. Proposed Acoustic Monitoring Protocol for Bats in Five National Parks in the Great Lakes Monitoring Network.

Gruver, J., P. Rabie, L. A. Starcevich, M. Romanski, and B. Route. 2016. Monitoring Protocol for Bats in the Great Lakes Inventory & Monitoring Network.

Hothorn, T., F. Bretz, and P. Westfall. 2008. Simultaneous Inference in General Parametric Models. Biometrical Journal 50:346–363.

Humphries, M. M., D. W. Thomas, and J. R. Speakman. 2002. Climate-mediated energetic constraints on the distribution of hibernating mammals. Nature 418:313–316.

Indiana Division of Fish & Wildlife. 2018. Endangered and Special Concern Species List.

Kitzes, J., and A. Merenlender. 2014. Large roads reduce bat activity across multiple species. PLoS ONE 9:e96341.

Krauel, J. J., and G. LeBuhn. 2016. Patterns of bat distribution and foraging activity in a highly urbanized temperate environment. PLoS ONE 11:e0168927.

Kruger, L., and R. Peterson. 2008. Occurrence of Temperate Bat Species at Three National Parks in the Great Lakes Region. Natural Resource Technical Report NPS/GLKN/NRTR–2008/128.

Kunz, T. H., E. B. Arnett, W. P. Erickson, A. R. Hoar, G. D. Johnson, R. P. Larkin, M. D. Strickland, R. W. Thresher, and M. D. Turtle. 2007. Ecological impacts of wind energy development on bats: Questions, research needs, and hypotheses. Frontiers in Ecology and the Environment 5:315–324.

Kurta, A. 2017. Mammals of the Great Lakes Region. Third Edit. University of Michigan Press, Ann Arbor.

Langwig, K. E., J. R. Hoyt, K. L. Parise, J. Kath, D. Kirk, W. F. Frick, J. T. Foster, and A. Marm Kilpatrick. 2015. Invasion dynamics of white-nose syndrome fungus, Midwestern United States, 2012–2014. Emerging Infectious Diseases 21:1023–1026.

Loeb, S. C., T. J. Rodhouse, L. E. Ellison, C. L. Lausen, J. D. Reichard, K. M. Irvine, T. E. Ingersoll, J. T. H. Coleman, W. E. Thogmartin, J. R. Sauer, C. M. Francis, M. L. Bayless, T. R. Stanley, and D. H. Johnson. 2015. A Plan for the North American Bat Monitoring Program (NABat). Gen. Tech. Rep. SRS-208. Asheville, NC: U.S. Department of Agriculture Forest Service, Southern Research Station. Volume 208.

Loeb, S. C., and E. A. Winters. 2013. Indiana bat summer maternity distribution: Effects of current and future climates. Ecology and Evolution 3:103–114.

Mazerolle, M. J. 2017. AICcmodavg: Model selection and multimodel inference based on (Q)AIC(c). R package version 2.1-1.

Michigan Natural Features Inventory. 2009. Michigan's Special Animals: Endangered, Threatened, Special Concern, and Probably Extirpated.

Miller, B. W. 2010. Revised Relative Abundance Estimates and Temporal Activity of Bats at Three Great Lakes National Parks Based on Acoustic Data. Natural Resource Report NPS/GLKN/NRR - 2010/178.

Minnesota Department of Natural Ressources. 2013. Minnesota's List of Endangered, Threatened, and Special Concern Species.

National Park Service. 2016. Protecting Bats in Parks. <https://www.nps.gov/subjects/bats/protecting-bats-in-parks.htm>. Accessed 3 Mar 2019.

National Park Service. 2018. Great Lakes Inventory & Monitoring Network: Parks & Partners. <https://www.nps.gov/im/glkn/parks-partners.htm>. Accessed 4 Mar 2019.

Pettit, J. L., and J. M. O'Keefe. 2017. Impacts of white-nose syndrome observed during long-term monitoring of a Midwestern bat community. Journal of Fish and Wildlife Management 8:69–78.

Pourshoushtari, R. D., B. P. Pauli, P. A. Zollner, and G. S. Haulton. 2018. Road and habitat interact to influence selection and avoidance behavior of bats in Indiana. Northeastern Naturalist 25:236–247.

Powers, K. E., R. J. Reynolds, W. Orndorff, W. M. Ford, and C. S. Hobson. 2015. Post-White-nose syndrome trends in Virginia's cave bats, 2008-2013. Journal of Ecology and The Natural Environment 7:113–123.

R Core Team. 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Ratcliffe, J. M., and J. W. Dawson. 2003. Behavioural flexibility: The little brown bat, Myotis lucifugus, and the northern long-eared bat, M. septentrionalis, both glean and hawk prey. Animal Behaviour 66:847–856.

Rodhouse, T. J., T. E. Phillipi, W. B. Monahan, and K. T. Castle. 2016. A macroecological perspective on strategic bat conservation in the U.S. National Park Service. Ecosphere 7:1–17.

Rodhouse, T. J., K. T. Vierling, and K. M. Irvine. 2011. A practical sampling design for acoustic surveys of bats. Journal of Wildlife Management 75:1094–1102.

Route, B., and J. Schaberl. 2013. A Cursory Survey of Bats in Voyageurs National Park, Minnesota. Natural Resource Technical Report NPS/GLKN/NRTR—2013/701.

Russo, D., L. Ancillotto, and G. Jones. 2018. Bats are still not birds in the digital era: Echolocation call variation and why it matters for bat species identification. Canadian Journal of Zoology 96:63–78.

Skalak, S. L., R. E. Sherwin, and R. M. Brigham. 2012. Sampling period, size and duration influence measures of bat species richness from acoustic surveys. Methods in Ecology and Evolution 3:490–502.

Smallwood, K. S. 2013. Comparing bird and bat fatality-rate estimates among North American wind-energy projects. Wildlife Society Bulletin 37:19–33.

Stevens, D. L., and A. R. Olsen. 2004. Spatially balanced sampling of natural resources. Journal of the American Statistical Association 99:262–278.

Turner, G. G., D. M. Reeder, and J. T. H. Coleman. 2011. A five-year assessment of mortality and geographic spread of white-nose syndrome in North American bats, with a look to the future. Bat Research News 52:13–27.

United States Fish and Wildlife Service. 1967. Endangered Species. Federal Register. Volume 32.

United States Fish and Wildlife Service. 2016. Endangered and Threatened Wildlife and Plants; 4(d) Rule for the Northern Long-Eared Bat. Federal Register. Volume 81.

United States Fish and Wildlife Service. 2018a. White-nose syndrome: The devestating disease of hibernating bats in North America.

United States Fish and Wildlife Service. 2018b. Range-wide Indiana Bat Survey Guidelines.

United States Geological Survey. 2019. White-nose Syndrome occurrence map - by year (2019). <https://www.whitenosesyndrome.org/resources/map>. Accessed 4 Mar 2019.

Warnecke, L., J. M. Turner, T. K. Bollinger, J. M. Lorch, V. Misra, P. M. Cryan, G. Wibbelt, D. S. Blehert, and C. K. R. Willis. 2012. Inoculation of bats with European Geomyces destructans supports the novel pathogen hypothesis for the origin of white-nose syndrome. Proceedings of the National Academy of Sciences 109:6999–7003.

Wisconsin Department of Natural Resources. 2015. Wisconsin Endangered and Threatened Species Laws and List.

CHAPTER 2: EVALUATION OF BAT ACOUSTIC ANALYSIS SOFTWARE

Abstract

Passive acoustic monitoring is a common method of studying bats which involves recording the echolocation calls of bats in their natural environment. Call sequences are then identified to species using specialized automated software. Particularly for long-term monitoring efforts, a limitation of these automated software programs is that newer versions use different algorithms and may therefore provide results that are not directly comparable to older versions. However, there is little available information regarding how much or in what ways the versions differ, or which versions are most accurate.

For this chapter, I evaluated two popular software programs used for automated bat call identification, Kaleidoscope Pro and SonoBat. I tested a common set of echolocation call files across multiple versions of each program and compared their outputs. Results demonstrated that the level of agreement among versions varied widely by species, with higher rates of agreement on species with more distinctive call characteristics. In addition, newer versions were more conservative, assigning fewer species-level identifications. However, newer versions were not found to be substantially more accurate than older versions. This information will be useful to bat researchers as they plan and perform their data analyses.

Introduction

Bat researchers are often interested in examining population trends to better understand the conservation and management needs of different bat species. However, because they are nocturnal, bats are difficult to observe directly, forcing researchers to use indirect monitoring methods. One common technique is to deploy specialized acoustic equipment to record the ultrasonic echolocation calls produced by bats as they travel and forage. Such passive acoustic monitoring is an efficient method, as ultrasonic bat detectors can be deployed across the landscape to independently record echolocation calls of passing bats, without the researcher needing to be present. Passive acoustic monitoring can provide information on species richness (Skalak et al. 2012), occupancy and detection probability (Gorresen et al. 2008), and relative activity levels (Ford et al. 2011). This method is more cost-effective and less invasive than capturing and handling bats, although it leaves greater uncertainty in species identification compared to having bats in the hand.

Limitations on the accuracy of call identification include call variation within species, overlapping call characteristics among species, recording quality, environmental conditions, and other factors (Barclay 1999, Frick 2013, Goerlitz 2018, Russo et al. 2018). Advanced acoustical analysis software programs, such as SonoBat (J. Szewczak, www.sonobat.com) or Kaleidoscope Pro (Wildlife Acoustics, Inc., www.wildlifeacoustics.com), have been developed to assist with this challenge of identifying echolocation call sequences to species. A typical workflow begins with running acoustic files through automated software that assigns each echolocation call sequence to the species that is most likely to have produced the sound. To achieve this, the software measures a suite of quantitative call parameters (such as frequency, duration, and amplitude) and compares these to a reference set of known-species calls (Russo and Voigt 2016). Automated analysis is often followed by manual vetting on a portion of the dataset, which involves visual review of the spectrogram by a trained individual to verify that the identification made by the software appears to be correct. This is an important step in validating results, however there is no standardized method and the reviewer's individual bias may impact results (Fritsch and Bruckner 2014). Because of the large size of many datasets (thousands or hundreds of thousands of files), and the time and expertise required, manual vetting is not always feasible, rendering a high degree of accuracy in the software even more important.

Currently, there are four commercially available automated software programs to aid in identifying North American bat calls to the species level: BCID (R. Allen, Bat Call Identification, Inc., http://www.batcallid.com), EchoClass (E. Britzke, U.S. Army Engineer Research and Development Center, www.fws.gov/midwest/Endangered/mammals/inba/surveys/inbaAcousticSoftware.html), Kaleidoscope Pro, and SonoBat. While auto-classification provides a method of objectively and efficiently processing calls, it does have several drawbacks. The software can be expensive to purchase (current prices up to approximately $1500) and data processing time can be significant (Froidevaux et al. 2014). Furthermore, each program is designed with different algorithms and is based on different underlying "call libraries". A call library is a collection of known-species bat calls, typically obtained by recording bats that have been captured and released with light-tags as markers to enable following them, or by recording bats at a known roost site (Clement et al. 2014). Previous studies have documented that the various automated software programs show only moderate agreement on species identifications when compared against a common test dataset (Janos 2013, Lemen et al. 2015, Rydell et al. 2017). Although a small number of cross-program comparisons have been done, there is a need for additional independent validation (Russo and Voigt 2016) and one unaddressed area is comparisons *within* programs.

When working with large datasets collected over an extended timespan, the frequent release of software upgrades presents two quandaries for users. First, it is unclear how much difference there is between software versions, in terms of whether results from different versions can be legitimately compared to each other. Second, although users may assume each upgrade is an advancement over the last in terms of auto-classification accuracy, there is little available data to demonstrate whether there is an actual improvement. Without any information to address these issues, it is difficult for a software user to know whether investing in and employing the upgraded version will be advantageous or not.

For researchers, it is not necessarily useful to change software versions within the time frame of a long-term study. If successive versions are used as upgrades become available, it will be difficult to determine if temporal trends are due to the change in software or due to an actual change in bat populations. On the other hand, using only a single software version would necessitate either reanalyzing previously-analyzed data (requiring significant time investment) or continuing to use an older version of software that was available at the outset of the project (possibly sacrificing accuracy). More information is clearly needed to help users make appropriate choices between software versions.

The objectives of this study are to compare multiple versions of two popular software programs to describe and quantify the differences between them, and to evaluate whether later versions provide improved classification accuracy, both overall and at a species-specific level.

Methods

Software Programs and Versions

Two software programs were selected for this analysis: Kaleidoscope Pro and SonoBat. These were selected because they are both widely used by bat researchers and are being actively maintained and updated by the developers. For Kaleidoscope Pro, three versions of the “Bats of North America

Classifier" were compared: Version 3.1.0, Version 4.3.0, and Version 5.1.0. For SonoBat, two versions of the program were compared: Version 3.2.2 and Version 4.3.1.

Test Datasets

The software programs were tested against three datasets. The first two contain data collected through passive acoustic monitoring at Apostle Islands National Lakeshore (Wisconsin) and Indiana Dunes National Park (Indiana). Sampling at both locations took place between June and August 2016 using Wildlife Acoustics Song Meter detectors. Full-spectrum data was recorded in WAV file format. For details of sampling methodology, refer to Chapter 1. Due to the large number of files collected, I took a random subset of 20,000 files from each park to make data processing more manageable. The true identity of the bat species creating each call sequence is unknown for these two datasets.

In contrast, the third dataset contains audio files recorded either at known roosts or at foraging sites where the species and individual bat were able to be identified unambiguously (Hooton and Adams, unpublished data; Adams 2013). This dataset contains 312 call files from free-flying bats of seven species. These recordings were made with Avisoft equipment and collected in Ontario, Saskatchewan, British Columbia, and New York. Because each call is identified to species, this dataset can be used to assess accuracy of the software programs. However, one limitation is that these files were included in the call library used to develop Kaleidoscope Pro and thus the program's "familiarity" with the files may inflate accuracy. Since I am primarily interested in comparisons within programs rather than across programs, and since it can be assumed that accuracy would be inflated equally for all versions of Kaleidoscope Pro, this is not necessarily problematic for answering the primary research questions.

Based on published range maps (Kurta 2017), there were up to nine Great Lakes region bat species that I considered to be possible within these test datasets. These include the big brown bat (*Eptesicus fuscus*, EPFU), eastern red bat (*Lasiurus borealis*, LABO), hoary bat (*Lasiurus cinereus*, LACI), silver-haired bat (*Lasionycteris noctivagans*, LANO), little brown bat (*Myotis lucifugus*, MYLU), northern long-eared bat (*Myotis septentrionalis*, MYSE), Indiana bat (*Myotis sodalis*, MYSO), evening bat (*Nycticeius humeralis*, NYHU), and tri-colored bat (*Perimyotis subflavus*, PESU). The list of species used for each test dataset is provided in Appendix B.

Data Analysis

For Kaleidoscope Pro, files were converted to zero-crossing format prior to species classification. This step was performed to account for the fact that the software programs approved by the U.S. Fish and Wildlife Service for endangered/threatened species surveys can only be used with zero-crossing formatted files (United States Fish and Wildlife Service 2018*b*). For SonoBat, files were kept in the full-spectrum format. For both Kaleidoscope Pro and SonoBat, software tests used the same or most equivalent settings for all versions of the program (Appendix B). Settings in Kaleidoscope Pro and SonoBat were not the same because the two programs do not provide all of the same options.

Both programs automatically generate output files containing the file-by-file classification information that was used for this analysis. This includes both species-specific classifications and additional categories that are not associated with one particular species (e.g. “No ID”). Classification categories for each software program, plus modifications I made during analysis, are explained in Appendix C. All comparisons were made across different versions of the same program; I did not attempt to directly compare Kaleidoscope Pro versus SonoBat.

To evaluate how sensitive or conservative each software version is, I first determined what proportion of files each version classified to the species level. The software programs assigned files to one of three groups: identified as a particular species, identified as an unknown bat species, or not identified. For each software program, I calculated the percent agreement between each possible pair of versions. Percent agreement was calculated as the percent of files given a certain classification by the first (earlier) version that had the same classification in the second (later) version. I used the same type of calculation to further examine the subset of files that were identified to the species level by both software versions in a comparison pair.

Using only the dataset of known-species calls, I evaluated the accuracy of the software programs. There are numerous ways to measure accuracy (Knight et al. 2017, Tharwat 2018) but I have chosen to use the two metrics “true positive rate” (true positive identifications/total number of files tested for that species) and “correct classification rate” (true positive identifications/number of files given a species-level identification) for ease of comparison with other published studies and with performance data provided by the software developers. Performance data for Kaleidoscope Pro was obtained from the website (https://www.wildlifeacoustics.com/products/kaleidoscope-pro/classifiers) and by emailing customer support. Performance data for SonoBat was obtained from help menus within the program itself.

Results

Software Version Comparisons

Kaleidoscope Pro results showed that the proportion of files classified to a specific bat species was highest in the earliest version of the program, Version 3.1.0 (Figure 3) and the two datasets (Apostle Islands and Indiana Dunes) had very similar patterns. The decrease in species-specific classifications in later versions was associated with an increase in unknown bat classifications. Meanwhile, the proportion of noise/not classified files remained constant across all three versions.

SonoBat similarly had a higher proportion of species-specific classifications in the earlier version of the program, Version 3.2.2 (Figure 3). For both datasets, there was a large increase in the proportion of noise/not classified files when using Version 4.3.1. For Indiana Dunes only, the proportion of files classified as unknown bats also decreased substantially for the later version.

For Kaleidoscope Pro, overall percent agreement was highest between the latter two versions of the program and slightly higher for the Apostle Islands dataset than the Indiana Dunes dataset (Table 3). All three versions shared a very high level of agreement for files classified as hoary bat or noise. For SonoBat, overall percent agreement was relatively low for both datasets (Table 4). At the species level, high rates of agreement were observed for big brown bat and hoary bat in the Apostle Islands dataset, and for big brown bat and noise/not classified files in the Indiana Dunes dataset.

I further examined species-specific agreements and disagreements between software versions by looking at only the subset of files that were assigned a species-level classification by both versions. This resulted in different subsets of files for each pair of versions compared. For the Apostle Islands dataset analyzed in Kaleidoscope Pro, I found high levels of agreement between Versions 3.1.0 and 4.3.0 (76-100%), and between Versions 3.1.0 and 5.1.0 (79-100%) for five of the six species (Tables D1 and D3). The exception was the northern long-eared bat with only 30-40% agreement. When comparing Versions 4.3.0 and 5.1.0 for the Apostle Islands data, I found even higher percent agreement for the same five species (97-100%), as well as much better agreement for northern long-eared bat (78%) (Table D5).

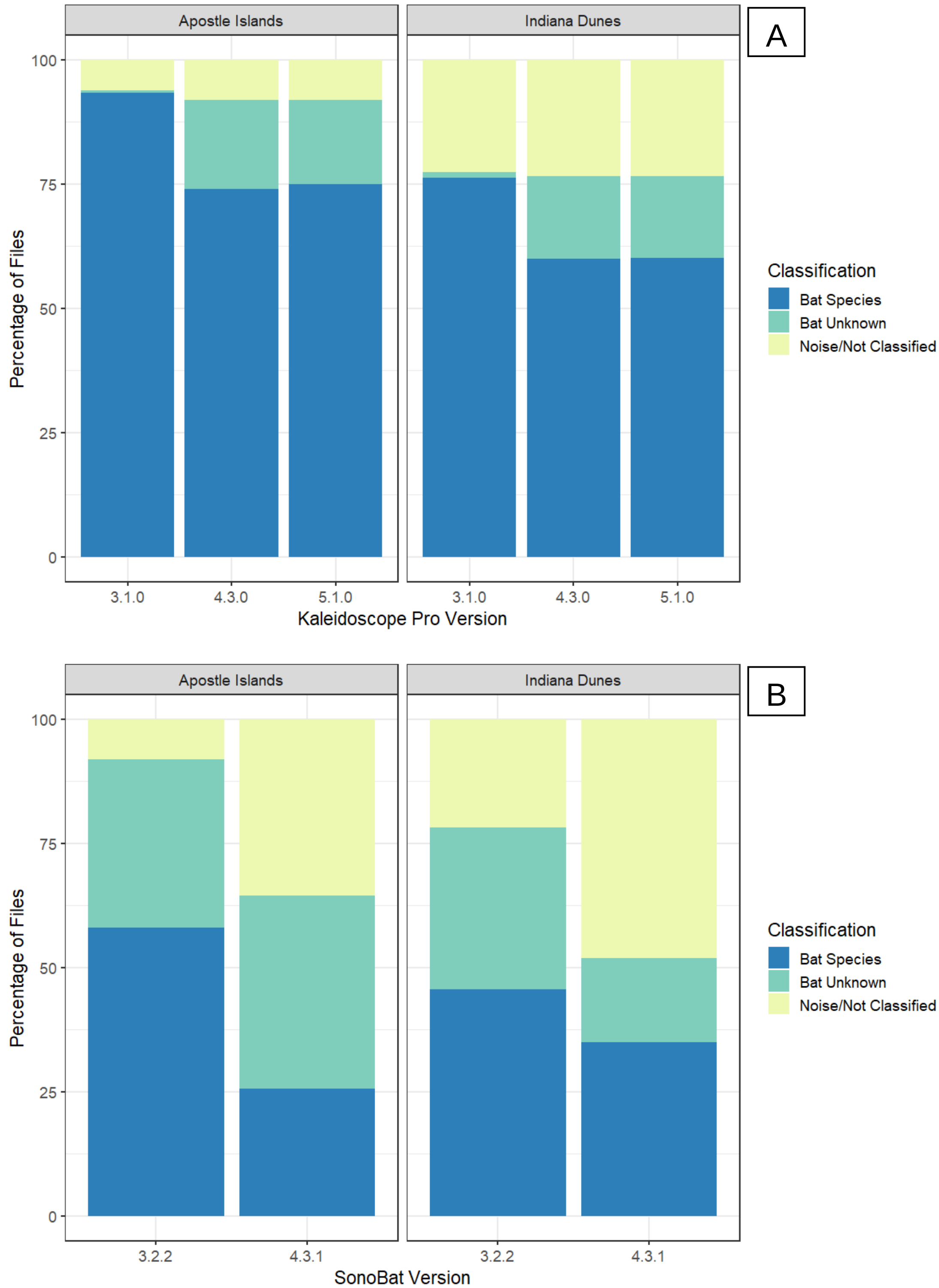

Figure 3. Percentage of files classified as a specific bat species (or species pair), as an unknown bat, or left unclassified by A) three versions of Kaleidoscope Pro and B) two versions of SonoBat. Test data consisted of audio files recorded in 2016 at two locations, Apostle Islands National Lakeshore (n=20,000) and Indiana Dunes National Park (n=20,000).

Table 3. Percent agreement among three versions of Kaleidoscope Pro, calculated as the percent of files given a certain classification by the first program that had the same classification in the second program. Test data consisted of audio files recorded in 2016 at Apostle Islands National Lakeshore (n=20,000) and Indiana Dunes National Park (n=20,000). MYSO, NYHU, and PESU were excluded because their ranges do not extend to Apostle Islands. No Indiana Dunes files were classified as MYSE by Version 4.3.0.

Species Classification	**Percent Agreement 3.1.0 and 4.3.0**		**Percent Agreement 3.1.0 and 5.1.0**		**Percent Agreement 4.3.0 and 5.1.0**	
	Apostle Islands	Indiana Dunes	Apostle Islands	Indiana Dunes	Apostle Islands	Indiana Dunes
EPFU	57.78	88.21	60.89	89.57	84.00	96.30
LABO	54.75	44.02	62.23	38.88	89.94	56.57
LACI	98.25	90.09	98.70	90.59	98.66	92.40
LANO	68.37	37.07	64.81	37.66	84.70	84.93
MYLU	75.76	17.39	72.24	18.63	85.42	68.92
MYSE	18.86	0.00	13.59	14.29	29.55	n/a
MYSO	n/a	6.25	n/a	0.00	n/a	25.00
NYHU	n/a	39.16	n/a	59.61	n/a	69.30
PESU	n/a	33.69	n/a	34.22	n/a	57.02
No ID	21.51	31.28	24.73	33.65	50.67	47.91
Noise	96.22	98.01	96.30	98.03	100.00	100.00
Total	***74.12***	***68.16***	***73.48***	**68.27**	***82.39***	**79.80**

Kaleidoscope Pro results for the Indiana Dunes dataset were more mixed (Tables D2, D4, and D6). Very high agreement (96-99%) was observed between Versions 3.1.0 and 4.3.0, and between Versions 3.1.0 and 5.1.0 for big brown bat and hoary bat. Agreement for the four remaining non-*Myotis* species generally ranged from 50-65% for these two comparison pairs, while the three *Myotis* species mostly had even lower levels of agreement. The Indiana Dunes dataset showed high levels of agreement (80-100%) for all nine species when comparing Versions 4.3.0 and 5.1.0.

Frequently, there was consistency in the disagreements among certain pairs of species. For example, in both datasets, files identified as northern long-eared bat or Indiana bat by Kaleidoscope Pro Version 3.1.0 were very often identified as little brown bat by the later versions. Likewise, files identified by the earlier version as little brown bat, evening bat, or tri-colored bat were all commonly classified as red bat by the later versions.

Table 4. Percent agreement between two versions of SonoBat, calculated as the percent of files given a certain classification by the first program that had the same classification in the second program. Test data consisted of audio files recorded in 2016 at Apostle Islands National Lakeshore (n=20,000) and Indiana Dunes National Park (n=20,000). Although Apostle Islands is not within documented ranges for MYSO, NYHU, and PESU these species were included because the classifier's species list is not customizable. No Indiana Dunes files were classified as MYSO by version 3.2.2.

Species Classification	**Percent Agreement**	
	Apostle Islands	Indiana Dunes
EPFU	68.64	80.26
LABO	14.47	39.27
LACI	87.20	27.33
LANO	45.43	30.70
LUSO	1.26	0.00
MYLU	36.02	37.50
MYSE	23.40	0.00
MYSO	0.00	n/a
NYHU	7.18	9.85
PESU	21.21	44.12
High Frequency Bat	25.40	27.20
Low Frequency Bat	10.25	8.77
Noise/Not Classified	33.78	93.21
Total	***28.04***	***52.07***

For SonoBat, I observed high levels of agreement (85-100%) between Versions 3.2.2 and 4.3.1 for seven of nine species in the Apostle Islands dataset and six of nine species in the Indiana Dunes dataset (Tables D7 and D8). For both datasets, evening bat and the "LUSO" species pair (indicating the file was indistinguishable between little brown bat and Indiana bat) had much lower percent agreement. The majority of files classified as evening bats by Version 3.2.2 were instead identified as red bats by Version 4.3.1 and similarly, the majority of files classified as LUSO by Version 3.2.2 were attributed to little brown bats by Version 4.3.1.

Software Accuracy Test

Results from the software accuracy test are presented both in terms of correct classification rates (Table 5) and true positive rates (Tables D9-D13). Correct classification rate includes only the subset of files that were given a species-level identification, while true positive rate accounts for the entire dataset. The test yielded higher average correct classification rates for Kaleidoscope Pro and lower average correct classification rates for SonoBat when compared to published performance data provided by the software companies.

Table 5. Percent correct classification achieved by three versions of Kaleidoscope Pro and two versions of SonoBat. Correct classification rate is equal to the number of true positive identifications divided by the number of files identified to the species level. Test data consisted of audio files recorded from bats of known species (n= 254 for Kaleidoscope Pro, n=312 for SonoBat). LANO was excluded for Kaleidoscope Pro due to incompatible file format. A classification of "LUSO" by SonoBat was considered correct for true MYLU files. No MYSE files were identified to the species level by SonoBat 4.3.1. For comparison, correct classification rates published by software companies are also shown.

True Species		Kaleidoscope Pro 3.1.0	Kaleidoscope Pro 4.3.0	Kaleidoscope Pro 5.1.0	SonoBat 3.2.2	SonoBat 4.3.1
EPFU	Test	82	79	85	58	87
	Publ	69	72	85	98	99
LABO	Test	85	98	92	80	78
	Publ	60	38	62	98	100
LACI	Test	73	91	86	87	83
	Publ	72	75	82	100	100
LANO	Test	n/a	n/a	n/a	100	100
	Publ	80	72	85	95	100
MYLU	Test	72	92	95	85	50
	Publ	72	80	91	94	94
MYSE	Test	77	92	67	88	n/a
	Publ	65	83	90	100	100
PESU	Test	89	100	88	93	95
	Publ	77	91	96	100	100
Average	***Test***	***80***	***92***	***85***	***84***	***83***
	Publ	***71***	***73***	***84***	***98***	***99***

Among the three Kaleidoscope Pro versions, Version 4.3.0 had the highest average correct classification rate, and the highest rate for four of the seven individual species (Table 5). More files were classified to the species level by Version 3.1.0 (88%) than by the other two versions (both about 73%), however the true positive rates were almost identical across all three versions, at 65-70% (Tables D9-D11). Individual true positive rates were generally similar among the three versions, with one notable exception being that Version 5.1.0 had a much lower rate for the northern long-eared bat.

For SonoBat, the two versions had similar average correct classification rates and at the species level, only two species were significantly different (Table 5). SonoBat Version 4.3.1 had a much higher correct classification rate than Version 3.2.2 for the big brown bat, but a much lower rate for the little brown bat. SonoBat Version 3.2.2 gave a species identification to 68% of files while Version 4.3.1 gave a species identification to only 37% of files, significantly reducing the overall true positive rate for the latter (Tables D12 and D13). Individual true positive rates for five of the seven species were also substantially lower in Version 4.3.1.

Discussion

Developers of bat call identification software are assumed to be modifying the underlying call classification algorithms for each successive version of their product, resulting in differences among different versions. For both SonoBat and Kaleidoscope Pro, I found that later versions were more conservative in their identifications, i.e. they classified a smaller number of audio files to a specific species while more files were labeled as unknown bats or noise. Both programs allow the user to adjust settings to increase or decrease the likelihood of a file being classified to the species level. In SonoBat, I set “Acceptable Call Quality” and “Decision Threshold” to the same values for both program versions and still observed the later version to be more conservative, indicating that this difference was not due simply to these settings. In contrast, my results for Kaleidoscope Pro are more uncertain due to the settings I selected. Kaleidoscope Pro offers three options: “More Sensitive”, “Balanced”, and “More Accurate”; however, according to the manufacturer’s software release notes, the meanings of these options are not the same for all program versions (Wildlife Acoustics Inc. 2019). I used the “More Sensitive” setting for Version 3.1.0 and the “Balanced” setting for Versions 4.3.0 and 5.1.0 because these were thought to be most similar based on the release notes. However it is possible that this choice of settings is at least partially responsible for the differences I saw between the earlier version and the two later versions. Previous work by Lemen et al. (2015) demonstrated that a smaller number of files were identified to the species level when Kaleidoscope settings were adjusted towards higher accuracy. Brabant et al. (2018) tested all three Kaleidoscope sensitivity settings and found differences in the number of species-level classifications and in classification accuracy. Testing all three settings would be a useful way to extend my study to help clarify what impact this user-defined option has on the results.

Level of agreement on classifications varied widely depending on which software versions were being compared. When including only files classified to the species-level by both versions, average percent agreement between the two SonoBat versions was 75%, while between the three Kaleidoscope Pro versions it was 63-94%. Other researchers comparing different bat call identification programs have also used a common test dataset with recordings of unknown species (Janos 2013, Lemen et al. 2015). The two programs tested by Janos (BCID and EchoClass) agreed on only 50% of species classifications (only including files with species-level identifications by both programs). Likewise, Lemen et al. tested all four programs (BCID, EchoClass, Kaleidoscope Pro, SonoBat) and observed average species-level agreement rates of only 26-58%. Given that these two studies made comparisons among completely different programs, while my study compared among different versions of the *same* program, it is not surprising that I found higher average agreement.

In testing for accuracy, my first aim was to determine how my results compared to published performance data from the software developers. I found higher average correct classification rates for Kaleidoscope Pro and lower average correct classification rates for SonoBat when compared to data provided by the software companies. In both cases, this may be related to the list of species included in the tests. For Kaleidoscope Pro Versions 3.1.0 and 4.3.0, the software tests performed by Wildlife Acoustics appear to include all possible North American species together (26 or 29 species, respectively) rather than smaller regional subsets. Because the settings I used eliminated many of these species from consideration, the software had fewer opportunities for misclassifications, resulting in higher average correct classification rates. Meanwhile, for Kaleidoscope Pro Version 5.1.0, Wildlife Acoustics tested a more limited set of only 12 species, and in this case their average correct classification rate agreed much more closely with mine. For SonoBat, I ran my data with the “Midwest” classifier which includes two species that were not present in the test dataset (Indiana bat and evening bat). The addition of these extra species likely allowed for more misclassifications, resulting in somewhat lower correct classification rates than suggested by the software developers.

Similar studies have evaluated software accuracy using known-species calls of European bats (Rydell et al. 2017, 2018, Brabant et al. 2018). Rydell et al. (2017, 2018) tested three programs and observed overall accuracies of 54-90%, while Brabant et al. (2018) tested four programs and observed overall accuracies of 31-77%. In both studies, accuracy was measured as true positive rate for species-level classifications. The true positive rates I obtained were comparable, at 32-70%, and like mine, their results showed wide variation among species.

My second goal with the accuracy test was to determine whether the software is actually improving, i.e. did more recent software versions have higher accuracy rates than older versions? Surprisingly, I did not find this to be true using either metric, correct classification rate or true positive rate. For both SonoBat and Kaleidoscope Pro, all versions demonstrated similar levels of accuracy. Yet, newer versions were more conservative, identifying fewer files to the species level. This may still be considered an advancement if poorer quality recordings (which are more likely to produce erroneous identifications) are appropriately left unclassified more frequently. Given that my test dataset included only seven species, this question warrants further research with larger, more diverse datasets to verify if my findings are broadly applicable.

Overall, my findings suggest that users of automated bat call identification software should take care to understand how their choices of software programs, versions, and settings may impact their results. It has been well demonstrated that different programs and settings provide different classifications (Janos 2013, Lemen et al. 2015, Rydell et al. 2017, Brabant et al. 2018) and my work extends this idea to different versions of the same program, which are more similar but still not in perfect agreement. Therefore, when examining bat acoustic data from different locations or time periods, it is critical to analyze all datasets using exactly the same analysis software and settings, to ensure they are comparable. Furthermore, it is recommended that prior to processing data, software users take the time to locate and review performance data provided by the software companies, in order to have appropriate expectations of how the selected program will perform for their species of interest. For their part, software developers could assist users by offering easily-accessible performance data with clear explanations of how testing was conducted and how metrics were calculated.

Despite its limitations, automated bat call identification software can be an incredibly valuable tool when properly applied. The recommendations presented here will be relevant to many bat researchers using acoustic monitoring methods and are intended to help make informed decisions that will maximize the efficiency and accuracy of data analysis.

Literature Cited

Adams, A. M. 2013. Assessing and Analyzing Bat Activity with Acoustic Monitoring: Challenges and Interpretations. The University of Western Ontario.

Barclay, R. M. R. 1999. Bats are not birds--A cautionary note on using echolocation calls to identify bats: A comment. Journal of Mammalogy 80:290–296.

Brabant, R., Y. Laurent, U. Dolap, S. Degraer, and B. J. Poerink. 2018. Comparing the results of four widely used automated bat identification software programs to identify nine bat species in coastal Western Europe. Belgian Journal of Zoology 148:119–128.

Clement, M. J., K. L. Murray, D. I. Solick, and J. C. Gruver. 2014. The effect of call libraries and acoustic filters on the identification of bat echolocation. Ecology and Evolution 4:3482–3493.

Ford, W. M., E. R. Britzke, C. A. Dobony, J. L. Rodrigue, and J. B. Johnson. 2011. Patterns of acoustical activity of bats prior to and following white-nose syndrome occurrence. Journal of Fish and Wildlife Management 2:125–134.

Frick, W. F. 2013. Acoustic monitoring of bats, considerations of options for long-term monitoring. THERYA 4:69–78.

Fritsch, G., and A. Bruckner. 2014. Operator bias in software-aided bat call identification. Ecology and Evolution 4:2703–2713.

Froidevaux, J. S. P., F. Zellweger, K. Bollmann, and M. K. Obrist. 2014. Optimizing passive acoustic sampling of bats in forests. Ecology and Evolution 4:4690–4700.

Goerlitz, H. R. 2018. Weather conditions determine attenuation and speed of sound: Environmental limitations for monitoring and analyzing bat echolocation. Ecology and Evolution 8:5090–5100.

Gorresen, P. M., A. C. Miles, C. M. Todd, F. J. Bonaccorso, and T. J. Weller. 2008. Assessing bat detectability and occupancy with multiple automated echolocation detectors. Journal of Mammalogy 89:11–17.

Janos, G. 2013. Utilizing acoustic methods to classify bat species, and to assess their habitat use and perception of owls. Graduate College of Bowling Green State University. Bowling Green State University.

Knight, E. C., K. C. Hannah, G. J. Foley, C. D. Scott, R. M. Brigham, and E. Bayne. 2017. Recommendations for acoustic recognizer performance assessment with application to five common automated signal recognition programs. Avian Conservation and Ecology 12.

Kurta, A. 2017. Mammals of the Great Lakes Region. Third Edit. University of Michigan Press, Ann Arbor.

Lemen, C., P. W. Freeman, J. A. White, and B. R. Andersen. 2015. The problem of low agreement among automated identification programs for acoustical surveys of bats. Western North American Naturalist 75:218–225.

Russo, D., L. Ancillotto, and G. Jones. 2018. Bats are still not birds in the digital era: Echolocation call variation and why it matters for bat species identification. Canadian Journal of Zoology 96:63–78.

Russo, D., and C. C. Voigt. 2016. The use of automated identification of bat echolocation calls in acoustic monitoring: A cautionary note for a sound analysis. Ecological Indicators 66:598–602.

Rydell, J., S. Nyman, J. Eklöf, G. Jones, and D. Russo. 2017. Testing the performances of automated identification of bat echolocation calls: A request for prudence. Ecological Indicators 78:416–420.

Rydell, J., S. Nyman, J. Eklöf, G. Jones, and D. Russo. 2018. Corrigendum to “Testing the performances of automated identification of bat echolocation calls: A request for prudence.” Ecological Indicators 84:273.

Skalak, S. L., R. E. Sherwin, and R. M. Brigham. 2012. Sampling period, size and duration influence measures of bat species richness from acoustic surveys. Methods in Ecology and Evolution 3:490–502.

Tharwat, A. 2018. Classification assessment methods. Applied Computing and Informatics.

United States Fish and Wildlife Service. 2018. Range-wide Indiana Bat Survey Guidelines.

Wildlife Acoustics Inc. 2019. Kaleidoscope Release Notes. <https://www.wildlifeacoustics.com/download/release-notes/498-kaleidoscope-release-notes>. Accessed 28 Mar 2019.

APPENDIX A: SUPPLEMENTARY TABLES FOR CHAPTER 1

Table A1. Linear mixed effects models were used to explore the effects of species, park, and sample site on the difference in bat call files per deployment night in 2016 vs. 2017. The name of each candidate model is shown above the model components, with random effects listed within the parentheses, e.g. (1|Site); model selection results are also shown. A) We tested five candidate models using *Park*, *Species*, and their interactions as fixed effects, all with sample *Site* as a random effect. B) The final model included *Species* as a fixed effect and sample *Site* nested within *Park* as a random effect.

(A) **Model**	**df**	**AICc**	**ΔAICc**	**Weight**	**Log Likelihood**
Park + Species Park + Species + (1|Site)	18	13520.71	0.00	0.84	-6742.07
Species Only Species + (1|Site)	11	13524.11	3.39	0.15	-6750.94
Park*Species Park*Species + (1|Site)	55	13535.34	14.63	0.00	-6710.05
Park Only Park + (1|Site)	10	13593.94	73.23	0.00	-6786.88
Null 1 + (1|Site)	3	13597.15	76.43	0.00	-6795.56
(B) **Model**	**df**	**AICc**	**ΔAICc**	**Weight**	**Log Likelihood**
Species Only Species + (1|Park/Site)	12	13522.96	0.00	1.00	-6749.35
Null 1 + (1|Park/Site)	4	13595.71	72.75	0.00	-6793.84

Table A2. Survey effort by park, 2016-2017. Valid nights had a minimum of four hours of recording. Valid deployments had a minimum of four recording nights between June 1 and August 15. Long deployments (>16 nights) were truncated to exclude nights after the 16th night. If multiple valid deployments were completed for a single site in the same year, only the first deployment was included. Table includes all valid deployments, even those with no valid corresponding deployment in the other year. Park abbreviations are provided on page vii.

Park	Number of Sites Surveyed in 2016	Number of Deployment Nights in 2016	Number of Sites Surveyed in 2017	Number of Deployment Nights in 2017
APIS	28	273	24	247
GRPO	17	152	18	177
INDU	27	241	30	244
ISRO	27	263	26	209
MISS	28	262	27	194
SACN	22	190	28	235
SLBE	35	298	35	335
VOYA	18	216	18	180
Total	**202**	**1895**	**206**	**1821**
Mean Nights Per Site		**9.38**		**8.84**

Table A3. Number of call files manually vetted, number of call files where reviewer agreed with the Kaleidoscope Pro species classification, and percent agreement. Manual reviewer was permitted to assign call files to a species group (e.g. EPFU/LANO or MYLU/MYSO) or unknown group (e.g. unknown high frequency). Here, “agreement” indicates manual reviewer identified the call file to the same exact species, not to a group. Kaleidoscope Pro classification used the Bats of North America Classifier Version 3.1.0 with the “More Sensitive” setting. Data shown are from 2016 and 2017 combined.

Species	Number of Files Vetted	Number of Agreements	Percent Agreement
EPFU	1543	543	35.19
LABO	1214	795	65.49
LACI	759	559	73.65
LANO	1157	620	53.59
MYLU	2266	1866	82.35
MYSE	251	59	23.51
MYSO	20	0	0.00
NYHU	81	32	39.51
PESU	123	74	60.16
Total	**7414**	**4548**	**61.34**

Table A4. A) Number of sample sites with detections of each species in 2016, by park. B) Number of sample sites with detections of each species in 2017, by park. Tables include only valid deployments that had a corresponding valid deployment in the other year. Park and species abbreviations are provided on page vii.

(A) Park	Total Sites	EPFU	LABO	LACI	LANO	MYLU	MYSE	MYSO	NYHU	PESU
APIS	24	20	23	22	23	24	23	--	--	--
GRPO	17	16	16	16	17	17	17	--	--	--
INDU	26	26	26	26	26	26	13	23	26	26
ISRO	24	21	22	21	22	23	24	--	--	--
MISS	24	24	24	23	24	24	22	--	--	23
SACN	18	18	17	18	18	18	18	--	--	16
SLBE	34	33	33	33	33	33	20	--	--	--
VOYA	18	17	18	18	18	18	18	--	--	--
All Parks*	**185**	**175**	**179**	**177**	**181**	**183**	**155**	**23**	**26**	**65**

(B) Park	Total Sites	EPFU	LABO	LACI	LANO	MYLU	MYSE	MYSO	NYHU	PESU
APIS	24	15	16	20	18	19	19	--	--	--
GRPO	17	15	11	17	17	17	10	--	--	--
INDU	26	26	26	26	26	26	12	18	26	26
ISRO	24	16	20	21	24	21	21	--	--	--
MISS	24	24	24	24	24	24	21	--	--	24
SACN	18	18	18	18	18	18	15	--	--	14
SLBE	34	34	34	34	34	34	24	--	--	--
VOYA	18	13	14	17	17	17	13	--	--	--
All Parks*	**185**	**161**	**163**	**177**	**178**	**176**	**135**	**18**	**26**	**64**

*All Parks includes all parks for which that species was given as an option for the classification software, e.g. for MYSO & NYHU it was only INDU.

Table A5. A) Percent of sample sites with detections of each species in 2016, by park. B) Percent of sample sites with detections of each species in 2017, by park. Tables include only valid deployments that had a corresponding valid deployment in the other year. Park and species abbreviations are provided on page vii.

(A) Park	EPFU	LABO	LACI	LANO	MYLU	MYSE	MYSO	NYHU	PESU
APIS	83.33	95.83	91.67	95.83	100.00	95.83	--	--	--
GRPO	94.12	94.12	94.12	100.00	100.00	100.00	--	--	--
INDU	100.00	100.00	100.00	100.00	100.00	50.00	88.46	100.00	100.00
ISRO	87.50	91.67	87.50	91.67	95.83	100.00	--	--	--
MISS	100.00	100.00	95.83	100.00	100.00	91.67	--	--	95.83
SACN	100.00	94.44	100.00	100.00	100.00	100.00	--	--	88.89
SLBE	97.06	97.06	97.06	97.06	97.06	58.82	--	--	--
VOYA	94.44	100.00	100.00	100.00	100.00	100.00	--	--	--
All Parks*	**94.59**	**96.76**	**95.68**	**97.84**	**98.92**	**83.78**	**88.46**	**100.00**	**95.59**

(B) Park	EPFU	LABO	LACI	LANO	MYLU	MYSE	MYSO	NYHU	PESU
APIS	62.50	66.67	83.33	75.00	79.17	79.17	--	--	--
GRPO	88.24	64.71	100.00	100.00	100.00	58.82	--	--	--
INDU	100.00	100.00	100.00	100.00	100.00	46.15	69.23	100.00	100.00
ISRO	66.67	83.33	87.50	100.00	87.50	87.50	--	--	--
MISS	100.00	100.00	100.00	100.00	100.00	87.50	--	--	100.00
SACN	100.00	100.00	100.00	100.00	100.00	83.33	--	--	77.78
SLBE	100.00	100.00	100.00	100.00	100.00	70.59	--	--	--
VOYA	72.22	77.78	94.44	94.44	94.44	72.22	--	--	--
All Parks*	**87.03**	**88.11**	**95.68**	**96.22**	**95.14**	**72.97**	**69.23**	**100.00**	**94.12**

*All Parks includes all parks for which that species was given as an option for the classification software, e.g. for MYSO & NYHU it was only INDU.

APPENDIX B: SETTINGS FOR SOFTWARE TESTS

Kaleidoscope Pro

- Files processed in zero cross (.zc) format
- Signal Parameters
  - 8-120 kHz
  - 2-500 ms
  - Maximum inter-syllable gap = 500 ms
  - Minimum number of pulse = 2
  - Advanced signal processing used
- Used “-1 More Sensitive (Liberal)” option for Version 3.1.0 and “0 Balanced (Neutral)” option for Versions 4.3.0 and 5.1.0 based on software release notes.
- Species group individually selected for each dataset. For Apostle Islands, this included: EPFU, LABO, LACI, LANO, MYLU, and MYSE. For Indiana Dunes, this included the Apostle Islands list plus MYSO, NYHU, and PESU. For Hooton & Adams dataset, this included the Apostle Islands list plus PESU.

SonoBat

- Files processed in full spectrum (.wav) format
- Acceptable Call Quality = 0.80
- Sequence Decision Threshold = 0.90
- Maximum Number of Calls to Consider Per File = 16
- Autofilter applied at 5 kHz
- Default Detector Sample Frequency = 256 kHz
- Used the same species group (“Midwest”) for all datasets. This group includes nine species: EPFU, LABO, LACI, LANO, MYLU, MYSE, MYSO, NYHU, and PESU.

APPENDIX C: FILE CLASSIFICATION CATEGORIES

Kaleidoscope Pro

- Species: The file contains bat calls and the overall consensus for the sequence is a single species.
- No ID: The file contains bat calls but the overall consensus for the sequence cannot be narrowed down to a single species, therefore it is considered an unknown bat.
- Noise: The file does not contain bat calls that can be detected by the classifier, therefore it is considered to be non-bat noise.

SonoBat

- Species: The file contains bat calls and the overall consensus for the sequence is a single species.
- Species combination (LUSO): The file contains bat calls and the overall consensus for the sequence is the closely related species pair *Myotis lucifugus*/*Myotis sodalis*.
- High frequency: The file contains high frequency bat calls but the overall consensus for the sequence cannot be narrowed down to a single species, therefore it is considered an unknown high frequency bat.
- Low frequency: The file contains low frequency bat calls but the overall consensus for the sequence cannot be narrowed down to a single species, therefore it is considered an unknown low frequency bat.
- High/Low frequency: The file contains both high and low frequency bat calls and the overall consensus for the sequence cannot be narrowed down to a single species, therefore it is considered an unknown bat.
- No frequency: The file does not contain bat calls that can be detected by the classifier or does not contain enough call pulses to meet the minimum requirements. Therefore it is likely non-bat noise or a low quality recording.
- Not classified: The program did not attempt to classify the file, therefore it is likely non-bat noise.

- For the purposes of this analysis, files assigned to the “LUSO” combination were considered to have a species-level classification.
- For Figure 3, high frequency, low frequency, and high/low frequency files were combined into a single unknown bat category. Additionally, no frequency and unclassified files were combined into a single noise/not classified category.

APPENDIX D: SUPPLEMENTARY TABLES FOR CHAPTER 2

Table D1. Percent agreement between Kaleidoscope Pro Versions 3.1.0 and 4.3.0, calculated as the percent of files given a certain classification by the first program that had the same classification in the second program. Audio files were recorded in 2016 at Apostle Islands National Lakeshore. Table includes only files that were assigned a species-level classification by both versions of the software (n=14,742).

		Classification by Version 4.3.0					
		EPFU	**LABO**	**LACI**	**LANO**	**MYLU**	**MYSE**
Classification by Version 3.1.0	**EPFU**	76.02	1.75	10.53	11.11	0.58	0.00
	LABO	0.00	77.26	0.08	0.19	22.44	0.04
	LACI	0.00	0.00	100.00	0.00	0.00	0.00
	LANO	5.50	0.12	7.46	86.92	0.00	0.00
	MYLU	0.00	2.54	0.12	0.02	96.89	0.42
	MYSE	0.00	0.00	0.00	0.00	59.74	40.26

Table D2. Percent agreement between Kaleidoscope Pro Versions 3.1.0 and 4.3.0, calculated as the percent of files given a certain classification by the first program that had the same classification in the second program. Audio files were recorded in 2016 at Indiana Dunes National Park. Table includes only files that were assigned a species-level classification by both versions of the software (n=11,827). Asterisk indicates only a single file for the species.

		Classification by Version 4.3.0								
		EPFU	**LABO**	**LACI**	**LANO**	**MYLU**	**MYSE**	**MYSO**	**NYHU**	**PESU**
Classification by Version 3.1.0	**EPFU**	96.24	0.02	2.78	0.96	0.00	0.00	0.00	0.00	0.00
	LABO	0.05	64.86	0.36	0.05	4.29	0.00	0.03	26.57	3.79
	LACI	0.09	0.00	99.82	0.00	0.00	0.00	0.00	0.09	0.00
	LANO	19.71	0.49	26.88	52.72	0.00	0.00	0.00	0.21	0.00
	MYLU	1.29	48.07	0.00	0.86	36.05	0.00	0.86	6.01	6.87
	MYSE*	0.00	0.00	0.00	0.00	100.00	0.00	0.00	0.00	0.00
	MYSO	16.67	16.67	0.00	0.00	50.00	0.00	16.67	0.00	0.00
	NYHU	0.00	48.34	0.27	0.00	0.13	0.00	0.00	50.86	0.40
	PESU	0.00	27.88	0.96	0.96	1.92	0.00	0.00	7.69	60.58

Table D3. Percent agreement between Kaleidoscope Pro Versions 3.1.0 and 5.1.0, calculated as the percent of files given a certain classification by the first program that had the same classification in the second program. Audio files were recorded in 2016 at Apostle Islands National Lakeshore. Table includes only files that were assigned a species-level classification by both versions of the software (n=14,934).

		Classification by Version 5.1.0					
		EPFU	**LABO**	**LACI**	**LANO**	**MYLU**	**MYSE**
Classification by Version 3.1.0	**EPFU**	82.53	0.00	9.64	7.83	0.00	0.00
	LABO	0.03	79.23	0.10	0.21	20.42	0.00
	LACI	0.00	0.00	99.96	0.04	0.00	0.00
	LANO	6.84	0.00	9.33	83.83	0.00	0.00
	MYLU	0.02	5.88	0.12	0.06	93.32	0.59
	MYSE	0.00	0.44	0.00	0.00	69.91	29.65

Table D4. Percent agreement between Kaleidoscope Pro Versions 3.1.0 and 5.1.0, calculated as the percent of files given a certain classification by the first program that had the same classification in the second program. Audio files were recorded in 2016 at Indiana Dunes National Park. Table includes only files that were assigned a species-level classification by both versions of the software (n=11,861). Asterisk indicates only a single file for the species.

		Classification by Version 5.1.0								
		EPFU	**LABO**	**LACI**	**LANO**	**MYLU**	**MYSE**	**MYSO**	**NYHU**	**PESU**
Classification by Version 3.1.0	**EPFU**	96.21	0.00	2.89	0.90	0.00	0.00	0.00	0.00	0.00
	LABO	0.25	58.44	0.50	0.50	7.90	0.00	0.03	29.18	3.20
	LACI	0.27	0.00	99.36	0.18	0.00	0.00	0.00	0.18	0.00
	LANO	23.02	0.07	24.92	51.92	0.00	0.00	0.00	0.07	0.00
	MYLU	1.15	58.46	0.38	0.38	34.62	0.00	0.38	1.15	3.46
	MYSE*	0.00	0.00	0.00	0.00	0.00	100.00	0.00	0.00	0.00
	MYSO	14.29	14.29	0.00	0.00	71.43	0.00	0.00	0.00	0.00
	NYHU	0.00	16.05	0.14	0.00	0.29	0.00	0.00	83.52	0.00
	PESU	0.00	33.03	0.92	0.92	2.75	0.00	0.00	3.67	58.72

Table D5. Percent agreement between Kaleidoscope Pro Versions 4.3.0 and 5.1.0, calculated as the percent of files given a certain classification by the first program that had the same classification in the second program. Audio files were recorded in 2016 at Apostle Islands National Lakeshore. Table includes only files that were assigned a species-level classification by both versions of the software (n=13,215).

		Classification by Version 5.1.0					
		EPFU	**LABO**	**LACI**	**LANO**	**MYLU**	**MYSE**
Classification by Version 4.3.0	**EPFU**	98.00	0.00	2.00	0.00	0.00	0.00
	LABO	0.10	98.11	0.00	0.10	1.70	0.00
	LACI	0.03	0.00	99.97	0.00	0.00	0.00
	LANO	1.39	0.00	1.23	97.38	0.00	0.00
	MYLU	0.00	1.14	0.00	0.00	98.67	0.19
	MYSE	0.00	2.00	0.00	0.00	20.00	78.00

Table D6. Percent agreement between Kaleidoscope Pro Versions 4.3.0 and 5.1.0, calculated as the percent of files given a certain classification by the first program that had the same classification in the second program. Audio files were recorded in 2016 at Indiana Dunes National Park. Table includes only files that were assigned a species-level classification by both versions of the software (n=10,262). No Indiana Dunes files were classified as MYSE by Version 4.3.0. Asterisk indicates only a single file for the species.

		Classification by Version 5.1.0								
		EPFU	**LABO**	**LACI**	**LANO**	**MYLU**	**MYSE**	**MYSO**	**NYHU**	**PESU**
Classification by Version 4.3.0	**EPFU**	99.74	0.00	0.22	0.04	0.00	0.00	0.00	0.00	0.00
	LABO	0.20	80.51	0.05	0.05	1.47	0.00	0.00	17.24	0.49
	LACI	0.39	0.00	99.55	0.06	0.00	0.00	0.00	0.00	0.00
	LANO	3.48	0.00	0.00	96.52	0.00	0.00	0.00	0.00	0.00
	MYLU	0.00	8.95	0.00	0.00	91.05	0.00	0.00	0.00	0.00
	MYSE	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	MYSO*	0.00	0.00	0.00	0.00	0.00	0.00	100.00	0.00	0.00
	NYHU	0.00	9.64	0.00	0.00	0.00	0.00	0.00	90.36	0.00
	PESU	0.00	9.21	0.00	0.00	0.66	0.00	0.00	1.97	88.16

Table D7. Percent agreement between SonoBat Versions 3.2.2 and 4.3.1, calculated as the percent of files given a certain classification by the first program that had the same classification in the second program. Audio files were recorded in 2016 at Apostle Islands National Lakeshore. Table includes only files that were assigned a species-level classification by both versions of the software (n=4,267). No Apostle Islands files were classified as MYSO by Version 3.2.2.

		Classification by Version 4.3.1									
		EPFU	**LABO**	**LACI**	**LANO**	**LUSO**	**MYLU**	**MYSE**	**MYSO**	**NYHU**	**PESU**
Classification by Version 3.2.2	**EPFU**	97.59	1.20	0.00	0.00	0.00	1.20	0.00	0.00	0.00	0.00
	LABO	0.00	85.47	0.00	0.56	2.79	0.56	0.00	6.70	3.91	0.00
	LACI	0.09	0.09	99.28	0.36	0.00	0.18	0.00	0.00	0.00	0.00
	LANO	3.09	0.77	5.93	89.69	0.00	0.26	0.00	0.00	0.26	0.00
	LUSO	0.00	0.73	0.00	0.15	9.45	85.32	0.00	2.91	1.45	0.00
	MYLU	0.00	0.36	0.18	0.00	0.72	97.67	0.00	0.00	1.08	0.00
	MYSE	0.00	0.00	0.00	0.00	0.00	7.69	84.62	7.69	0.00	0.00
	MYSO	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	NYHU	0.00	73.55	0.83	0.00	0.00	2.48	0.00	0.00	23.14	0.00
	PESU	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00

Table D8. Percent agreement between SonoBat Versions 3.2.2 and 4.3.1, calculated as the percent of files given a certain classification by the first program that had the same classification in the second program. Audio files were recorded in 2016 at Indiana Dunes National Park. Table includes only files that were assigned a species-level classification by both versions of the software (n=5,409). No Indiana Dunes files were classified as MYSE or MYSO by Version 3.2.2.

		Classification by Version 4.3.1									
		EPFU	**LABO**	**LACI**	**LANO**	**LUSO**	**MYLU**	**MYSE**	**MYSO**	**NYHU**	**PESU**
Classification by Version 3.2.2	**EPFU**	99.30	0.42	0.20	0.03	0.00	0.00	0.00	0.03	0.03	0.00
	LABO	1.21	93.86	0.11	0.00	0.44	0.22	0.00	2.30	1.86	0.00
	LACI	10.96	1.99	85.71	1.00	0.33	0.00	0.00	0.00	0.00	0.00
	LANO	20.67	2.51	3.07	73.74	0.00	0.00	0.00	0.00	0.00	0.00
	LUSO	33.33	0.00	0.00	0.00	0.00	66.67	0.00	0.00	0.00	0.00
	MYLU	0.00	0.00	0.00	0.00	0.00	100.00	0.00	0.00	0.00	0.00
	MYSE	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	MYSO	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
	NYHU	0.41	72.43	0.41	0.00	0.00	0.00	0.00	0.41	26.34	0.00
	PESU	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00

Table D9. Classification accuracy results for Kaleidoscope Pro Version 3.1.0. True positive rate (true positive identifications/total number of files tested for that species) is found in the "Percent Correct" column. Test data consisted of audio files recorded from bats of known species (n=254). LANO was excluded due to incompatible file format.

True Species	Total Percent Given a Species ID	Percent Correct	Percent Incorrect	Percent Not Given a Species ID
EPFU	100.00	82.00	18.00	0.00
LABO	96.30	81.48	14.81	3.70
LACI	91.11	66.67	24.44	8.89
LANO	n/a	n/a	n/a	n/a
MYLU	89.29	64.29	25.00	10.71
MYSE	76.47	58.82	17.65	23.53
PESU	56.25	50.00	6.25	43.75
Overall	***88.19***	***69.69***	***18.50***	***11.81***

Table D10. Classification accuracy results for Kaleidoscope Pro Version 4.3.0. True positive rate (true positive identifications/total number of files tested for that species) is found in the "Percent Correct" column. Test data consisted of audio files recorded from bats of known species (n=254). LANO was excluded due to incompatible file format.

True Species	Total Percent Given a Species ID	Percent Correct	Percent Incorrect	Percent Not Given a Species ID
EPFU	84.00			16.00
		66.00	18.00	
LABO	77.78			22.22
		75.93	1.85	
LACI	75.56			24.44
		68.89	6.67	
LANO	n/a			n/a
		n/a	n/a	
MYLU	67.86			32.14
		62.50	5.36	
MYSE	70.59			29.41
		64.71	5.88	
PESU	53.13			46.88
		53.13	0.00	
Overall	***72.83***			***27.17***
		66.14	***6.69***	

Table D11. Classification accuracy results for Kaleidoscope Pro Version 5.1.0. True positive rate (true positive identifications/total number of files tested for that species) is found in the “Percent Correct” column. Test data consisted of audio files recorded from bats of known species (n=254). LANO was excluded due to incompatible file format.

True Species	**Total Percent Given a Species ID**	**Percent Correct**	**Percent Incorrect**	**Percent Not Given a Species ID**
EPFU	92.00	78.00	14.00	8.00
LABO	88.89	81.48	7.41	11.11
LACI	80.00	68.89	11.11	20.00
LANO	n/a	n/a	n/a	n/a
MYLU	66.07	62.50	3.57	33.93
MYSE	17.65	11.76	5.88	82.35
PESU	53.13	46.88	6.25	46.88
Overall	***73.62***	***65.35***	***8.27***	***26.38***

Table D12. Classification accuracy results for SonoBat Version 3.2.2. True positive rate (true positive identifications/total number of files tested for that species) is found in the “Percent Correct” column. Test data consisted of audio files recorded from bats of known species (n=312). A classification of “LUSO” was considered correct for true MYLU files.

True Species	Total Percent Given a Species ID	Percent Correct	Percent Incorrect	Percent Not Given a Species ID
EPFU	76.00	44.00	32.00	24.00
LABO	74.07	59.26	14.81	25.93
LACI	68.89	60.00	8.89	31.11
LANO	46.55	46.55	0.00	53.45
MYLU	71.43	60.71	10.71	28.57
MYSE	47.06	41.18	5.88	52.94
PESU	90.63	84.38	6.25	9.38
Overall	***68.27***	***56.41***	***11.86***	***31.73***

Table D13. Classification accuracy results for SonoBat Version 4.3.1. True positive rate (true positive identifications/total number of files tested for that species) is found in the "Percent Correct" column. Test data consisted of audio files recorded from bats of known species (n=312). A classification of "LUSO" was considered correct for true MYLU files.

True Species	Total Percent Given a Species ID	Percent Correct	Percent Incorrect	Percent Not Given a Species ID
EPFU	60.00	52.00	8.00	40.00
LABO	16.67	12.96	3.70	83.33
LACI	80.00	66.67	13.33	20.00
LANO	27.59	27.59	0.00	72.41
MYLU	10.71	5.36	5.36	89.29
MYSE	0.00	0.00	0.00	100.00
PESU	59.38	56.25	3.13	40.63
Overall	***37.18***	***32.05***	***5.13***	***62.82***