What does a detection mean? Spatial and behavioural context improves the use of passive acoustic monitoring for the conservation of a wide-ranging bird

by

Elly C. Knight

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Ecology

Department of Biological Sciences
University of Alberta

Abstract

The culture of ecology is shifting towards collaborative, integrative approaches that use ‘big data’ to solve big problems. Passive acoustic monitoring (PAM) has the potential to play a role in this new paradigm because it uses in-situ autonomous recording units (ARUs) to collect a permanent archive of audio recordings. PAM research groups across the world are collecting vast amounts of acoustic data that could be integrated to understand ecological phenomena at a global scale; however, there are several hurdles that must be overcome. First, accurate algorithms that automatically scan acoustic recordings (hereafter, “recognizers”) are required to efficiently determine the species detected within the recorded soundscapes. Second, an understanding of the context of these recognizer-processed datasets is essential for data integration and can improve how recognizer data is used in ecology. Unlike survey data collected by human observers, recognizer data is typically treated as a binomial dataset with minimal context beyond date and time of observations. In this thesis, I demonstrate the importance of spatial and behavioural context of recognizer data using the common nighthawk (*Chordeiles minor*) as a model species. The common nighthawk is a crepuscular and highly mobile bird species that consumes aerial insects and is declining in most parts of its breeding range across North America. First, I demonstrate a fundamental principle that is crucial to using and integrating recognizer data. I show that the classification probability of species detections reported by a recognizer is related to detection distance, and that the classification threshold applied thus defines the survey area. Understanding this spatial context is necessary for estimating density from recognizer data and for integrating multiple datasets. Next, I show that the behavioural context of recognizer data can provide important insight into ecological analyses. I use VHF telemetry to show that the wing-boom display of the common nighthawk is a territorial signal, which I then use to study behaviour-specific habitat use of this species in the boreal forest. I use the wing-boom to separate territorial from home range behaviour and show that the scale that most strongly predicts habitat use corresponds to the movement range of that behaviour. I then use the wing-boom to separate territorial from extraterritorial behaviour to confirm that the common nighthawk is a ‘disturbance specialist’ species in the boreal forest, but only for nesting territories. Finally, I combine these spatial and behavioural contexts to demonstrate a novel method for density estimation that can be applied to single ARUs at broad spatial scales. The goal of this approach is to improve regional, national, or range-wide population estimates, especially for regions that are poorly covered by human surveys or for species that have large home ranges. Together, this density modelling approach and the spatial principles my thesis presents will facilitate future integration of PAM datasets collected with varying methodologies as well as with other data types. The behavioural context component of my thesis encourages PAM users to ensure they put their recognizer data into appropriate ecological context, particularly for wide-ranging species. Collectively, the ecological inferences in my thesis provide a major advance in understanding common nighthawk ecology in the boreal forest and the tools developed will help future research and conservation of this enigmatic species.

Preface

This thesis is an original work by Elly Knight. The common nighthawk tracking work in this thesis received approval from the University of Alberta Care Use Committee under AUP00001523, “The Migratory Connectivity Project”, 11 May 2015 (and subsequent renewals).

Most of the research for this thesis was conducted as part of various collaborations. All chapters consist of my own original work, supported by co-authors as listed below. Unless otherwise indicated, I conducted all conceptualization, data processing and analysis, and writing, with feedback on writing by co-authors. Acoustic recordings were collected by a wide variety of sources, including some of the co-authors mentioned below. The contributions of my supervisors Dr. Erin Bayne, Dr. Mark Brigham, and other co-authors are reflected with the use of plural pronouns in Chapters 2-7.

Chapters 2 and 3 of this thesis have been published in peer-reviewed journals.

Knight EC, Hannah K, Foley G, Scott C, Brigham RM, Bayne EM (2017) Recommendations for acoustic recognizer performance assessment with application to five common automated signal recognition programs. Avian Conservation and Ecology 12:14.

*The convolutional neural network recognizer used in this chapter was trained by Chris Scott.

Knight EC, Bayne EM (2018) Classification threshold and training data affect the quality and utility of focal species data processed with automated audio recognition software. Bioacoustics 28:539-554.

Chapter 4,5, and 6 are under review at peer-reviewed journals.

Knight, EC, Brigham RM, Bayne EM (In review) The Big Boom Theory: Common Nighthawks use wing-boom displays to defend nesting territories. Ornithology ORNITH-21-054.

Knight, E.C., P. Sòlymos, R.M. Brigham, and E.M. Bayne (In review) Scale of effect corresponds to movement range for the combined effects of multiple environmental variables but not for individual variables. Landscape Ecology LAND-S-21-00410.
*The offsets for availability for detection in this chapter were estimated by Peter Sòlymos.

Knight, E.C., R.M. Brigham, and E.M. Bayne (In review) Specialist or generalist? It depends. Context-dependent habitat relationships provide insight into forest disturbance effects for a boreal bird species. Forest Ecology and Management FORECO-S-21-01661.

Chapter 7 is being prepared for peer-reviewed journal submission.

Knight, E.C., P. Sòlymos, S. Haché, R.F. Pankratz, and E.M. Bayne. Deriving density estimates from single-point recordings processed with automated recognition.
*Chapter 7 is a collaboration, with initial design and code for the density modelling approach performed by Peter Sòlymos, data processing for the northern region study areas performed by Rhiannon Pankratz, and study site selection performed by Samuel Haché.

Three first-author papers that were published during my PhD contributed to the development of the thesis ideas but were not part of core thesis itself. These papers are cited in the thesis but are not available in this document.

Yip DA*, Knight EC*, Audet EH, Wilson SJ, Charchuk C, Scott CD, Sòlymos P, Bayne EM (2020) Sound level measurements from audio recordings provide objective distance estimates for distance sampling wildlife populations. Remote Sensing in Ecology and Conservation 71:2759–15.

*Equal contributions

Knight EC, Sòlymos P, Scott C, Bayne EM (2020) Validation prediction: a flexible protocol to increase efficiency of automated acoustic processing for wildlife research. Ecological Applications 30:e02140.

Knight EC, Ng JW, Mader CE, Brigham RM, Bayne EM (2018) An inordinate fondness for beetles: first description of common nighthawk (Chordeiles minor) diet in the boreal biome. The Wilson Journal of Ornithology 130:525–531.

Two first-author papers, two co-authored papers, and one first author national protocol were also published during my PhD and provide important background information on common nighthawk behaviour, ecology, and conservation but were not directly related to the thesis on passive acoustic monitoring.

Knight EC., Harrison A, Scarpignato AL, Wilgenburg SLV, Bayne EM, Ng JW, Angell E, Bowman R, Brigham RM, Drolet B, Easton WE, Forrester TR, Foster JT, Haché S, Hannah KC, Hick KG, Ibarzabal J, Imlay TL, Mackenzie SA, Marsh A, McGuire LP, Newberry GN, Newstead D, Sidler A, Sinclair PH, Stephens JL, Swanson DL, Tremblay JA, Marra PP (2021) Comprehensive estimation of spatial and temporal migratory connectivity across the annual cycle to direct conservation efforts. Ecography 44:665-679.

Knight Elly C, Smith AC, Brigham RM, Bayne EM (2021) Combination of targeted monitoring and Breeding Bird Survey data improves population trend estimation and species distribution modeling for the common nighthawk. Ornithological Applications 123:duab005.

Nebel S, Casey J, Cyr M-A, Kardynal KJ, Krebs EA, Purves EF, Bélisle M, Brigham RM, Knight EC, Morrissey C, Clark RG (2020) Falling through the policy cracks: implementing a roadmap to conserve aerial insectivores in North America. Avian Conservation and Ecology 15:23.

Ng JW, Knight EC, Scarpignato AL, Harrison AL, Bayne EM, Marra PP (2018) First full annual cycle tracking of a declining aerial insectivorous bird, the common nighthawk (Chordeiles minor), identifies migration routes, nonbreeding habitat, and breeding site fidelity. Canadian Journal of Zoology 96:869–875.

Knight EC, Hannah KC, Brigham RM, McCracken J, Falardeau G, Julien M-F, Guénette J-S, Manthorne A, Sidler A, Foley GJ, McKnight J, Laurent KS, Sinclair P, Pankratz R (2019) Canadian Nightjar Survey Protocol. http://wildresearch.ca/wp-content/uploads/2019/05/National-Nightjar-Survey-Protocol-WildResearch-2019.pdf

One additional first-author paper was published during my PhD and provides further reading on emerging technologies in automated signal recognition but was not directly related to the thesis on acoustic and behavioural context.

Knight EC, Hernandez SP, Bayne EM, Bulitko V, Tucker BV (2019) Pre-processing spectrogram parameters improve the accuracy of bioacoustic classification using convolutional neural networks. Bioacoustics 29:337-355.

Dedication

I dedicate my thesis to the hundreds of community scientists who volunteered their time to help build what has now become the Canadian Nightjar Survey. Your contributions not only provide the foundation for future nightjar conservation, but set me on my current path, for which I am eternally grateful. May the nightjars be with you.

"*There's something magical about nighthawks, the sounds they make, the heights they achieve in the sky, their aberrant timing compared to every other spring migrant. We seem to be losing them, and within my own, brief lifetime. This not something we should countenance; we need to do everything we can to maintain them, if only as a shadow of their former presence.*"

~Doug Wilson, Community Scientist

Acknowledgements

To my supervisors, Erin Bayne & Mark Brigham: thank you for sharing your wisdom and investing in me. You gave me freedom to grow but have been cheerleaders and advocates when I needed it, each in your own way. Erin, one of the most impactful things you ever told me was to be myself; thank you for giving me the confidence to be a scientist with a personality. Thank you as well to Evelyn Merrill for your feedback throughout my degree, and to you and Mark and Erin for your encouragement when I shared that I was pregnant. Special thanks to Nancy Mahony, who has been a role model and mentor to me for over a decade; I cannot express how much I appreciate your kindness and leadership as a woman and parent in science.

I acknowledge that my research was conducted on Treaty 10 territory and am grateful to the Dene and Cree people who continue to influence those lands, their traditional territory, through their vibrant history and culture. I am lucky to have spent almost four summers in the sandy pine forest and incredible patterned fen that is the McCelland Lake area of northeastern Alberta. Thanks to my first field partner in the Bayne lab, Captain Dan Yip, for introducing me to the McClelland Lake area and showing me the ropes (literally) of passive acoustic monitoring. I am deeply grateful to Matthew Timpf, Azim Shariff, Orla Osborne, Timothy Forrester, Martin Kueblbeck, and Jeremiah Kennedy for the countless long nights spent chasing nighthawks in the sand. Thank you to Matt especially for sharing your eastern whip-poor-will wisdom.

Back in Edmonton, the members of the Bayne lab were a lifeline. To Julia Shonfield and Janet Ng, my original officemates: you taught me what it was to be a PhD student. Your generosity, leadership, and friendship helped shape my thesis and my work-life balance. To Natalie Sanchez Ulate, Juan Andrés Martinez Lanfranco, and Cesar Estavo, my second-generation officemates: I am grateful for the many conversations about science, mental health,

and everything in between that also shaped this thesis. Thanks as well to Jeremiah Kennedy, Emily Upham-Mills, Richard Hedley, Michelle Knaggs, Natasha Annich, Alex MacPhail, and Elene Haave-Audet for your friendship. And to all the many members of the Bayne lab over the years; I have enjoyed working with each one of you.

I am extremely fortunate to have had much financial support for my research. Funding for my passive acoustic monitoring research was provided by the National Science and Engineering Research Council of Canada, the Northern Scientific Training Program, the UAlberta North program, the Alberta Conservation Association, and many organizations that funded the collection of the passive acoustic monitoring data used in my thesis. Thank you to the organizations that have financially supported my family and I during my research: the Natural Sciences and Engineering Research Council of Canada, the University of Alberta, the Killam Foundation, the Canadian Federation of University Women, the Alberta Chapter of the Wildlife Society, the Alberta Government, the Alberta Society for Professional Biologists, and the Alberta Chapter of the Wildlife Society. Your generous support gave me the time and freedom to build a substantial body of research.

I have had the privilege of working with a great many collaborators during my PhD. To Péter Sólymos, whose brilliance has pushed my quantitative skills and changed the way I think about data. To Chris Scott, whose contributions in deep learning improved the impact of my recognizer research. To Adam Smith, who was critical in helping turn the Canadian Nightjar Survey into a long-term monitoring program. To Kevin Hannah, whose love of nightjars continues to be a common ground for great collaboration. To Elora Grahame, Virginia Noble, Logan Parker, and the rest of the nightjar community, thank you for your shared enthusiasm of these amazing birds. To the many other co-authors I have worked with during my PhD, I have enjoyed working with each and every one of you; thank you.

Although it didn’t make it into the thesis, I cannot express how lucky I am to have the relationships I built and experiences I gained doing full annual cycle tracking of common nighthawks with the Smithsonian Migratory Bird Center and Environment and Climate Change Canada. To Peter Marra and Steven Van Wilgenburg: thank you for putting your trust in me. Your mentorship has greatly broadened my abilities as a scientist. To Autumn-Lynn Harrison and Amy Scarpignato: you are both brilliant and I am lucky to call you friends and mentors. I am so excited to continue working with all of you!

My fascination with common nighthawks started in 2010 during my MSc degree when I stumbled upon nighthawk nests while doing point counts in the sagebrush hills of the Okanagan Valley. Shortly thereafter, my good friend Mike Boyd passed on the beginnings of a community/citizen science project that monitored nightjars. Over a decade later, that community science project has become a formal long-term monitoring project, the Canadian Nightjar Survey, and the spin-off is this PhD thesis. I am eternally grateful to all the community scientists who participated in the nightjar survey over the years. And to the board and members of WildResearch, the organization that housed the original nightjar survey. Of course, many thanks to Mike Boyd for getting me started, and to Megan Harrison, my field partner in 2010 who was encouraging of my nightjar enthusiasm in those early days.

The other crucial person in my nighthawk journey is Janet Ng. The stars aligned to put my PhD desk next to Janet’s. Janet developed the technique to catch common nighthawks, and it was that technique and Janet’s nighthawk wisdom that facilitated most of my research. Janet has patiently spent dozens of hours catching nighthawks in the field with me and talking nighthawks in the office with me. Janet has been there to shake her fist at the nighthawks when they were difficult to study and to be fascinated when we learned something new. Thank you, Janet. And thank you as well Maurice, for being the handsome decoy you know you are.

Thank you to my parents, Rob and Bev Knight, who have been unwaveringly supportive my whole life. I have never felt discouraged to do anything but follow my heart, and that is the greatest gift indeed. Thank you to my Dad, for sharing your love of nature and showing me how fun it is to be wildlife biologist. I am so grateful to share that bond and those memories with you, from river float surveys for steelhead to online nightjar data management. Thank you to my Mom: your contributions are perhaps less thematically tangible, but no less impactful. You have given me the empathy, communication skills, and confidence that are central to being a good scientist and a good person. Thank you to my extended family, the DeMoors, for welcoming me with open arms and being my home away from home.

Finally, thank you to my family. You are the reason I work hard and seek work-life balance. Harris and Juniper, thank you for your patience when I'm working, for your interest in my research, but mostly thank you for your love and support. Ardeya, thank you for being the quirky, funny, vibrant little soul that you are. And for finally learning how to sleep; this thesis would not be complete if you hadn't. Jonathan, this thesis is more than just a little bit yours. You have contributed so much through your support and feedback, but also through your love, which has given me the balance and emotional security to do this work. Thank you for being my sweet spot.

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

Collaboration & Integration: A Shifting Paradigm in Conservation

There is no denying the current ecological crisis. Global biodiversity is declining, with approximately 1 million species facing extinction (Cardinale et al. 2012; IPBES 2019). There are three billion fewer birds in North America than 40 years ago (Rosenberg et al. 2019). Increasing land use pressure and declining knowledge threaten the sustainable land management by indigenous peoples (Garnett et al. 2018; Schuster et al. 2019). Evidence of massive insect declines suggests we are amid an insect apocalypse (Didham et al. 2019; Wagner 2020; Seibold et al. 2019). 85% of global wetland area has been lost (IPBES 2019). The list goes on. To add to this, climate change exacerbates all these issues, often with complex and unpredictable effects (IPCC 2014). As ecologists and conservationists, we have our work cut out for us. Our increasingly global world presents us with challenges at a global scale, and so we need solutions that operate at the same spatial scale.

The response to this challenge has been a culture shift; we cannot answer complex ecological questions across political boundaries without working together (Palmer 1993; Hampton et al. 2013). Due to funding structures and lack of incentive for data sharing, ecological research has historically been dominated by individual scientists conducting local and short-term research (Hampton et al. 2013; Heidorn 2008). But times are changing with the advent of ecological "big data" (Farley et al. 2018; Hampton et al. 2013), the push for open data and reproducibility (Powers & Hampton 2019; Tenopir et al. 2015; Reichman et al. 2011), and the technology to store ecological data from a variety of sources (Farley et al. 2018). For example, Movebank is a free, online database of animal tracking data with over 5 billion datapoints from

3,000 collaborators (Kranstauber et al. 2011). The recent development of e-collaboration tools (Linnes 2020; Kock 2007) and research on the dynamics of effective collaboration (Guerrero et al. 2015; Reeves et al. 2017) has increased the ease and effectiveness of working together across oceans and time zones towards a common goal. In the last year and a half, the global pandemic has provided further impetus for the growth of remote collaboration and calibrated its effectiveness. The pause in human activity due to the pandemic was accompanied by rapid initiation of several global research groups to study the effects of this "anthropause" on wildlife (Rutz et al. 2020). We see quantitative evidence of the shift towards collaboration in the increasing number of authors on peer-reviewed publications (Harrison 2006; Adams et al. 2019).

A paradigm shift towards large-scale collaboration would not be possible without the concurrent development of new statistical tools to integrate multiple data types from varying sources (Farley et al. 2018). Different data types (e.g., structured vs unstructured) have strengths and weaknesses that when integrated, can complement each other and lead to improved statistical outcomes (Isaac et al. 2019). Bayesian hierarchical models are particularly well-suited for data integration because they are flexible and can characterize uncertainty in data and estimates (Farley et al. 2018). For example, recent integration of 10 sources of population estimates, 7 sources of population trajectories, and weather radar data showed that 3 billion birds, or 29% of North America's avifauna have been lost over the past 40 years (Rosenberg et al. 2019). Integrated modelling approaches are available for population trend monitoring (Zipkin et al. 2009; Bowler et al. 2019), species distribution modelling (Miller et al. 2018; Jr. et al. 2019; Simmonds et al. 2020), integrated population modelling (Rushing 2019; Schaub & Abadi 2011), and make use of varying datatypes like citizen science, remote sensing, long-term monitoring stations, and of course in-situ sensor networks like autonomous recording units (ARUs).

Passive Acoustic Monitoring

Passive acoustic monitoring (PAM) is a relatively new form of ecological data collection that holds great promise for contributing to ecology at all scales. PAM uses autonomous recording units (ARUs) to collect longitudinal datasets of acoustic recordings following a pre-set schedule. PAM is useful for surveying taxa that rely on acoustic communication, of which there are many (e.g., birds, amphibians, insects, fish, mammals), as well as monitoring for disturbance and anthropogenic effects via the overall soundscape (Shonfield & Bayne 2017; Gibb et al. 2018; Sugai et al. 2018). PAM is resource-efficient because it surveys the entire acoustic community, can collect multi-visit data, and does not require highly skilled observers. PAM can also produce more objective datasets because human observers are subjective and can influence species observations via their presence and disturbance.

PAM has the potential to contribute substantially to large-scale integrative analyses because acoustic recordings provide a permanent archive of soundscapes that can be reanalyzed at any point to address new objectives. Acoustic archives have already reached big data proportions in many parts of the world (Gibb et al. 2018; Sugai et al. 2018); together, these datasets hold enormous potential for answering scalable questions. For example, PAM data collected by 19 organizations was compiled to track range-shifts of the endangered North Atlantic right whale in the Atlantic Ocean (Davis et al. 2017). Thanks to their pre-programmable nature, ARUs can also be used to fill gaps in datasets of other types, particularly for remote areas or rare species. For example, the Boreal Optimal Sampling Strategy integrates PAM and human point counts from many sources to survey the large and historically under-monitored ecoregion that is Canada’s boreal forest (Van Wilgenburg et al. 2020; Roy et al. 2019; Barker et al. 2015).

Finally, the high temporal resolution of PAM data makes it an ideal source of information on species detectability that is required to integrate data types.

Automated Recognition

There are disadvantages to PAM, but one that has received substantial attention in recent years is the necessity to turn sound files into ecological data (hereafter, "process"). Typically, sound files are listened to by a human observer and converted into species detection data; however, the time and expertise required for this process is prohibitive and hinders the application of PAM to large-scale ecological questions (Gibb et al. 2018; Priyadarshani et al. 2018; Sugai et al. 2018).

Automated processing approaches have been developed to reduce the time required to process ARU recordings. There are two broad categories of automated processing algorithms (Gibb et al. 2018). The first category is "recognizers", which are trained with clips of a focal species (or a species group) and produce a species detection record (Priyadarshani et al. 2018; Xie et al. 2016). The second category is acoustic indices, which convert the entire soundscape into a single metric (Lawrence et al. 2019). I focus on the first type of algorithm, recognizers, for the remainder of this thesis.

A multitude of recognizer approaches have been developed to improve the efficiency of processing audio recordings. Researchers have built recognizers for a wide variety of taxa (Sugai et al. 2018). Some of the earlier algorithms tested include support vector machines (Acevedo et al. 2009; Armitage & Ober 2010), hidden Markov models (Chu & Blumstein 2011), and band-limited energy detection (Charif et al. 2010). More recently, deep learning approaches like convolutional neural networks have been found to outperform other algorithms (Knight et al.

2019b; Stowell et al. 2018). Despite the automation of recognizers, they still require human oversight because they produce false positive detections that require removal from the dataset prior to analysis (hereafter, “validate”) (Priyadarshani et al. 2018). Previously, the time required to validate rendered automated recognition no more efficient than human processing (Borker et al. 2014; Joshi et al. 2017), but as the precision (i.e., proportion of false positives) of recognizers continues to improve, so does their efficiency. There are also post-processing approaches to weed out false positives (Knight et al. 2020; Balantic & Donovan 2019) and statistical approaches to incorporate false positives (Chambert et al. 2017; Doser et al. 2021; Barré et al. 2019), both of which decrease the validation workload.

What Does a Detection Mean?

One of the other disadvantages of PAM, and especially automated recognition, that has received less attention is the lack of context. In other words, what does a detection mean? With traditional human point counts, observers can collect information on distance, direction, movement, behaviour, and abundance (Ralph et al. 1993). In contrast, recognizer-processed data are typically binomial in nature with presence or absence of the focal species in each recording. Covariates for time and date are available from the timestamp of the recording, and environmental covariates can be sourced from remotely sensed datasets or on-the-ground surveys, but there is no information available about events during the time of recording, other than the acoustic environment.

In this thesis, I explore the value of spatial and behavioural context for ecological application of recognizer data. First, without information on the location of the vocalizing individual, recognizer data cannot be used for density estimation because the survey area is unknown (Pérez-Granados & Traba 2021). The survey area of human point counts is typically determined via distance estimation at the time of survey (Sòlymos et al. 2013; Matsuoka et al. 2012), and distance can also be estimated on acoustic recordings by human listeners (Darras et al. 2018); however, the survey area of recognizers is unknown.

More importantly, an unknown survey area limits integration with other data sources because differences in survey effort between sources cannot be accounted for without paired datasets collected at the same time (Van Wilgenburg et al. 2017). Some integration approaches overlook this step by assuming equal survey areas (Doser et al. 2021); however, the risk of this approach is that if survey areas are unequal, estimates will be biased according to the distribution of data types in the analysis. Understanding the spatial context of recognizer-processed datasets is therefore critical for tackling large-scale ecological questions via integration of PAM and other data types.

Second, the behavioural context of recognizer data is also lacking because observers are not present in the field to observe and interpret visual cues alongside aural ones. Ignoring the behavioural context and thus functional significance of detections can lead to erroneous conclusions about habitat because function-specific habitats can have opposing attributes that nullify, change the strength of, or distort the signal of habitat use (Roever et al. 2013; Manly et al. 2002). For example, some habitats like drinking locations may be used less often than others, and so the attributes of those habitats may be masked without separate consideration (Boyce & McDonald 1999). Further, it can be important for conservation to understand use of a particular function-specific habitat if that habitat contributes disproportionately to fitness or survival (Beyer et al. 2010; Law & Dickman 1998).

Focal Species: Common Nighthawk

Birds are one of the taxa for which PAM holds much potential because birds rely on sound for communication (Catchpole & Slater 2008). Automated recognition is being increasingly used to build detection datasets for bird species (Priyadarshani et al. 2018). Understanding the benefits of spatial and behavioural context for analysis of recognizer data should improve ecological inferences and conservation outcomes for birds. Currently, recognizer data from PAM is used for a variety of applications including occupancy estimation (Chambert et al. 2017), studying phenology (Ulloa et al. 2016), density estimation (Sebastián-González et al. 2018; Doser et al. 2021), and surveying for rare species (Swiston & Mennill 2009), to name a few.

In my thesis, I use the common nighthawk as a model and focal species for automated recognition. The common nighthawk is a nightjar (Family *Caprimulgidae*) that breeds across North America and overwinters in South America (Ng et al. 2018; Knight et al. 2021a). On the breeding grounds, nighthawks occupy a wide range of vegetation types with open to semi-open structures, for example pine forests, sand dunes, grasslands, and flat, gravel rooftops (Brigham et al. 2011). The common nighthawk's use of open habitats is likely for nesting because this species lays its eggs directly on the ground, as do other nightjar species (Figure 1.1 B-D). Common nighthawk selection for open-structured habitats may also facilitate foraging because common nighthawks forage for aerial insects on the wing (Figure 1.1 A). As a nightjar, the common nighthawk is a crepuscular bird, with mottled brown plumage that camouflages it during diurnal roosting and nesting (Figure 1.1 B, D). Nighthawks (genus *Chordeiles*) are typically more crepuscular than other more nocturnal nightjar species, however, likely because they forage during constant flight as opposed to sallying from a perch and therefore rely heavily on vision for foraging (Brigham & Barclay 1995). This foraging strategy also renders nighthawks more mobile than other nightjar species; the sparse available tracking data for the common nighthawks suggests home ranges of upwards of 40 km^2 (*unpublished data*).

Figure 1.1. A. Male common nighthawk on the wing. B. Recently hatched semi-precocial common nighthawk nestling and soon-to-hatch egg. C & D. Female common nighthawk incubating a nest.

Perhaps due to their aerial insectivorous diet, the common nighthawk is a species of conservation concern. As a group, the species that eat flying insects, the “aerial insectivores” (swallows, swifts, flycatchers, nightjars), are declining faster than any other group of birds in Canada (Nebel et al. 2010; Smith et al. 2015; Michel et al. 2015). The common nighthawk is no exception; available monitoring data for this species suggests population declines of 68% between 1970 and 2015 (COSEWIC 2018). Due to these declines, the common nighthawk is listed as Threatened in Canada, although it has been recommended for downlisting to Special Concern by the Committee on the Status of Endangered Wildlife in Canada (COSEWIC), largely due to large numbers of nighthawks found in the boreal forest (COSEWIC 2018). Regionally, the common nighthawk is also listed as Special Concern or Threatened in several eastern provinces and considered imperiled or critically imperiled in several eastern states (Environment Canada 2016a). Two subspecies were also recently added to the US Fish and Wildlife's list of Birds of Conservation Concern (US Fish and Wildlife Service 2021). Declines across the aerial insectivore guild suggest threats are linked to aerial insects as a food source. Evidence of insect declines have been reported across the globe (Didham et al. 2019; Wagner 2020; Seibold et al. 2019) and could be linked to nighthawk population declines via a variety of mechanisms including agricultural pesticides, loss of insect-producing habitats, light pollution, and phenological mismatch due to climate change (Spiller & Dettmers 2019). Additional potential causes of decline include loss of nesting habitat (esp. gravel rooftops), vehicle collisions, fire suppression, increased storm frequency, direct effects of pesticides, temperature extremes, and drought (Environment Canada 2016a). These threats could occur anywhere in the western hemisphere because the common nighthawk's full annual cycle stretches from the breeding grounds in Canada's Yukon Territory to wintering grounds in northern Argentina (Knight et al. 2021a).

The long list of potential threats to the common nighthawk is an indicator of how little we know about this species. The combination of its crepuscular nature, extreme mobility, and excellent camouflage makes this a difficult species to study (Figure 1.1), and most research has been short-term and/or anecdotal. The common nighthawk's nocturnal nature makes it unlikely to be detected by most existing bird survey programs. Furthermore, it has one of the largest breeding ranges in the western hemisphere (Bird Conservation International 2017), and so there are large sampling gaps that impede assessment and management (Environment Canada 2016a). The gaps in common nighthawk understanding are likely greatest in Canada's boreal forest, which is thought to provide breeding grounds for a substantial proportion of the population (Haché et al. 2014) but where the species is poorly studied.

The common nighthawk is therefore an ideal focal species for PAM because ARUs can be deployed in remote locations and left to record at night when the species is most active. Furthermore, existing ARU datasets from across the range present an opportunity to build common nighthawk detection datasets to fill knowledge gaps for this species. In fact, two of the three other existing common nighthawk studies in the boreal forest have used PAM to build species detection datasets (Farrell et al. 2017, 2019; Sidler 2017).

The common nighthawk (*Chordeiles minor*) is also an ideal model species for recognizer methods development and testing first principles because its vocal characteristics negate many of the challenges of applying automated recognition to birds (Priyadarshani et al. 2018). First, it is crepuscular, and therefore calls when there are few other species in the soundscape that might mask its acoustic signal. Second, it has a short call, which results in little overlap of calls between individuals. Third, it's short call is simple and easy for a computer to distinguish from other signals. Fourth, it is a non-passerine species with a limited repertoire and minimal variation between individuals or populations. Fifth, it calls from far above the canopy, which renders the vegetation effects on sound attenuation relatively negligible. Sixth, it calls frequently, making it relatively likely it will be detected if it is vocally active.

Finally, the common nighthawk is an ideal species to emphasize the importance of context for recognizer data because it does present some challenges. The foraging strategy and extreme mobility of the common nighthawk means individuals are always in motion. They can move in and out of the survey area both during and between recordings. Understanding the spatial and behavioural context of detections is therefore critical to interpreting and applying PAM techniques to research on and management of the common nighthawk.

Thesis Scope

The goal of my thesis is to understand and demonstrate the importance of spatial and behavioural context of recognizer-processed species detection data for ecological applications. This goal is split into three objectives.

Objective 1. Spatial Context

The first objective of my thesis is to understand the spatial context of passive acoustic monitoring and bridge the gap between computer algorithms and ecological application. All recognizers report a classification probability (hereafter "score") for each sound evaluated, and it is these score values to which a user-defined threshold is applied to classify each sound as a detection of the target species or not. The meaning of score and score thresholds has yet to be interpreted in an ecological context, however, because there is a disciplinary divide between the computing scientists that design automated recognition software and the ecologists that apply those algorithms. Chapters 2 and 3 show how score values and thresholds are central to understanding the spatial context of recognizer data.

In Chapter 2, I demonstrate the effectiveness of automated recognition for the common nighthawk and set the stage for the importance of spatial context. I compare five different computer algorithms for automated recognition of common nighthawks and provide general recommendations for using recognizers to build species detection datasets. I use the recognizer data in a variety of common statistical applications to show that recognizer performance depends on intended use. I apply each evaluation across a range of score thresholds and show that performance varies depending on the threshold used, suggesting that recognizers may report fewer detections than human observers because they can't "hear" as far.

In Chapter 3, I further investigate the potential for score threshold as a proxy for survey area. I trained recognizers with common nighthawk clips recorded at varying distances. I show that thanks to the principle of spherical spreading (i.e., sound amplitude decreases predictably with distance), score has a reliable relationship with distance when recognizers are trained with calls recorded at short range. Yip et al. (2020) and Knight et al. (2020) are reference papers for this chapter.

Objective 2. Behavioural Context

The second objective of my thesis is to use behavioural context to extend the ecological applications of passive acoustic monitoring. The common nighthawk makes two distinct sounds that likely have different behavioural contexts. The wing-boom is a mechanically produced acoustic signal that occurs when the primary feathers are flexed downwards at the bottom of a steep aerial dive (Miller 1925). The resultant sound is a 'vroom', between 0 and 1 kHz (Figure 1.2) and is thought to be a territorial signal, potentially associated with the nest (Rust 1947). The call, on the other hand, is a short (~0.3 second) 'peent' produced at between approximately 2-5 kHz (Figure 1.2). This call is a more general-purpose signal, although it is most frequently produced near the nest location as well (Armstrong 1965; Caccamise 1974; Wedgwood 1973), in part because the wing-boom is always accompanied by the call (*unpublished data*). The accompaniment of the wing-boom by a call means that a recognizer for the call can be used to detect wing-booms by simply noting them during validation; therefore, the spatial context of the call recognizer also applies to those noted wing-boom detections.

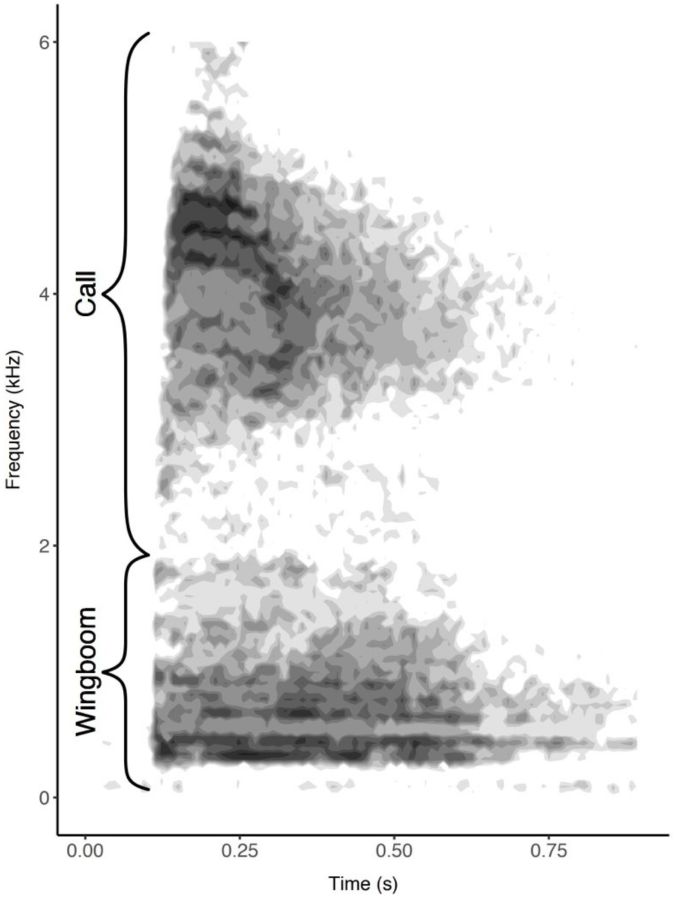


Figure 1.2. Spectrogram of the common nighthawk ‘peent’ call and wing-boom recorded at McLelland Lake, Alberta, Canada on July 13, 2017.

In Chapter 4, I use wing-boom locations of VHF-tagged individual male common nighthawks to study the meaning of the wing-boom signal. I apply kernel density estimation to show that wing-boom areas were exclusively used as a territorial signal. I then use a resource selection function to show that male common nighthawks select areas near the nest to perform wing-boom displays.

In Chapters 5 and 6, I demonstrate application of automated recognition and behavioural context of the wing-boom display at a large spatial scale using an archived dataset of ARU recordings from multiple projects. In Chapter 5, I use the common nighthawk's two types of acoustic signal, wing-boom and call, to model two types of habitat use, territorial and home range, respectively. I use machine learning and a variety of environmental predictors to determine the scales at which the two types of habitat use are best predicted (hereafter, "scale of effect"). Understanding the behavioural context of the common nighthawk's acoustic signals facilitates comparison of scales of effect at different movement ranges and examine a new theory for the mechanism that drives scale of effect.

In Chapter 6, I again use the common nighthawk's two types of acoustic signal, wing-boom and call, but this time to model territorial versus extraterritorial (i.e., locations with calls and no wing-booms) habitat use. I use occupancy modelling to test two competing hypotheses for why common nighthawks use recently disturbed areas in the boreal forest: nesting resources or foraging resources. Understanding the behavioural context of the common nighthawk's acoustic signals allows me to compare the competing hypotheses using three lines of evidence: time since disturbance, disturbance type, and vegetation effects. Unanimous support from all three lines for the nesting resource hypothesis shows that even the emulation of natural disturbance is context dependent. Knight et al. (2018) is an importance reference for this chapter.

Objective 3. Combining Contexts

The third objective of my thesis is to bring together the spatial and behavioural context of common nighthawk recognizer detections to provide a new analytical tool for recognizer-processed datasets.

In Chapter 7, I present a novel approach to density estimation using passive acoustic monitoring. The goal of this new approach is to provide a tool for density estimation at broad geographic scales because effective wildlife management requires accurate population estimates for species assessment and recovery. The approach uses information from three ARU datasets: a species detection dataset, an abundance subset dataset, and a separate known detection distance dataset. The approach incorporates the spatial context of the data by using the effective survey area of the recognizer to convert abundance estimates to density. The approach incorporates behavioural context by focusing on the wing-boom to estimate the population of breeding pairs and to increase the probability of presence during surveys, which reduces the bias and variance of density estimates. I apply the approach to four study areas in Canada’s boreal forest to show that post-wildfire areas support high densities of territorial common nighthawks.

Chapter 2. Recommendations for acoustic recognizer performance assessment with application to five common automated signal recognition programs

Abstract

Automated signal recognition software is increasingly used to extract species detection data from acoustic recordings collected using autonomous recording units (ARUs), but there is little practical guidance available for ecologists on the application of this technology. Performance evaluation is an important part of employing automated acoustic recognition technology because the resulting data quality can vary with a variety of factors. We reviewed the bioacoustic literature to summarize performance evaluation and found little consistency in evaluation, metrics employed, or terminology used. We also found that few studies examined how score threshold, i.e., cut-off for the level of confidence in target species classification, affected performance, but those that did showed a strong impact of score threshold on performance. We used the lessons learned from our literature review and best practices from the field of machine learning to evaluate the performance of five readily available automated signal recognition programs. We used the common nighthawk (*Chordeiles minor*) as our model species because it has simple, consistent, and frequent vocalizations. We found that automated signal recognition was effective for determining common nighthawk presence-absence and call rate, particularly at low score thresholds, but that occupancy estimates from the data processed with recognizers were consistently lower than from data generated by human listening and became unstable at high score thresholds. Of the five programs evaluated, our convolutional neural network (CNN) recognizer performed best, with recognizers built in Song Scope and MonitoR also performing well. The RavenPro and Kaleidoscope recognizers were moderately effective

but produced more false positives than the other recognizers. Finally, we synthesized six general recommendations for ecologists who employ automated signal recognition software, including what to use as a test benchmark, how to incorporate score threshold, what metrics to use, and how to evaluate efficiency. Future studies should consider our recommendations to build a body of literature on the effectiveness of this technology for avian research and monitoring.

Introduction

Autonomous acoustic sampling is a popular method of data collection for ecological research and monitoring because many species use sound as a primary method of communication (Catchpole & Slater 2008; Shonfield & Bayne 2017). In avian research, autonomous recording units (ARUs) are used to collect acoustic recordings, which can then be used for monitoring population trends (Furnas & Callas 2015), behavioral studies (Ehnes & Foote 2014), modeling habitat associations (Campos-Cerqueira & Aide 2016), and detecting rare or inconspicuous species (Holmes et al. 2014; Sidie-Slettedahl et al. 2015). ARUs provide a variety of benefits over traditional human point counts, including the ability to collect data over repeat visits (Drake et al. 2016) and the flexibility to collect data at any time of day or year (Goyette et al. 2011). Additionally, recordings provide a permanent record that can reduce observer bias (Haselmayer & Quinn 2000; Campbell & Francis 2012), be used to verify identification of rare species (Swiston & Mennill 2009; Holmes et al. 2014), and analyzed later for other objectives (Luther & Derryberry 2012). ARU technology has also been widely used to study marine mammals, bats, insects, and frogs.

One of the challenges of using ARUs for ecological research and monitoring is the time required to extract target species detections from recordings (Shonfield & Bayne 2017). In response, automated signal recognition programs have been developed (Oliveira et al. 2015; Katz et al. 2016; Nicholson 2016). Automated acoustic species recognition is the process of training a computer to detect, recognize, and evaluate the acoustic signature of a target species' vocalization. The computer runs the resultant detection algorithm (hereafter "recognizer") on recordings and evaluates the fit of the acoustic signal in the recording using a moving window. Some programs employ a single step process that runs the algorithm against every window (hereafter "moving window recognizer") while others use a two-step process that first conducts signal detection with a moving window, and then runs the algorithm only on detected signals (hereafter "signal detection recognizer"). For each window or signal evaluated, the recognizer assigns a score value, which can be interpreted as a measure of confidence that a target vocalization match has been found. The recognizer then registers a "hit" for each signal with a score above a user-defined threshold. Choosing a high score threshold will minimize false positives, i.e., false identifications, but also results in false negatives, i.e., missed detections. If the score threshold is set low by the user, this will minimize false negatives, but create many false positives. Choosing a score threshold is generally a subjective process based on the priorities of the user (Katz et al. 2016). The results of automated signal recognition are often manually validated by the user to separate true positives from false positives. Many approaches to automated acoustic species recognition or classification have been employed including random forest (Aide et al. 2013; Campos-Cerqueira & Aide 2016), Hidden Markov models (HMMs) (Skowronski & Harris 2006; Potamitis et al. 2014; Oliveira et al. 2015) and/or Gaussian mixture models (GMM) (Ganchev et al. 2015; Heinicke et al. 2015), binary point matching (Katz et al. 2016), spectrogram cross- correlation (Katz et al. 2016), artificial neural networks (Jennings et al. 2008; Tachibana et al. 2014; Nicholson 2016), decision trees (Digby et al. 2013), and band-pass filters (Charif et al. 2010). There are annual and one-time machine learning competitions that drive the development of new birdsong recognizer methods (Stowell et al. 2016, Goëau et al. 2017) with current state-of-the-art approaches using deep machine learning models such as convolutional neural networks to recognize multiple species from soundscape recordings (Koops et al. 2014, Joly et al. 2016, Salamon and Bello 2017). Some of these approaches are commercially or freely available, while others have been custom- built for specific research projects.

The number of tools available for automated signal recognition are rapidly increasing, yet there remains a need for a set of general recommendations for recognizer development and performance evaluation in ecology (Blumstein et al. 2011). Many authors have compared individual automated signal recognition programs to human processing to substantiate their use in ecological monitoring and research; however, authors have used a variety of metrics for evaluation, making it difficult to compare across studies. In other acoustic signal processing disciplines such as music analysis, speech classification, and machine learning, there are established best practices that ecologists can draw on to develop standardized evaluation methods (Salzberg 1997; Sokolova & Lapalme 2009; Raffel et al. 2014; Mesaros et al. 2016). Recognizer evaluation is particularly important because the quality of the species detection data produced can depend on a variety of factors including score threshold (Brauer et al. 2016), signal complexity of target species, quality of training data, spectrogram conversion, e.g., FFT size (Crump & Houlahan 2017), and recognition approach (Stowell et al. 2016). Ultimately, the appropriateness of automated acoustic species recognition will depend on the objective of the research or monitoring.

In response to this need for guidance, our goal was to provide general recommendations for recognizer performance comparison and evaluation. First, we review the literature for bioacoustic recognizer evaluation studies to confirm the need for such recommendations and identify the most commonly used metrics. Next, we conduct a recognizer evaluation based on the different approaches used in the literature to compare five common nighthawk (*Chordeiles minor*) recognizers: MonitoR (Katz et al. 2016), convolutional neural networks (CNN) (Abadi et al. 2015), Song Scope (Wildlife Acoustics 2011), Kaleidoscope (Wildlife Acoustics 2016), and RavenPro (Charif et al. 2010). Finally, we use our literature review, results from our evaluation, and best practices from other disciplines to synthesize general evaluation recommendations for ecologists who want to use automated acoustic recognition for data processing.

Literature Review of Evaluation Tools

Methods

We searched for ecological journal articles, technical reports, and conference proceedings that have evaluated the performance of automated signal recognition software to scan audio recordings for species detections. We searched the literature using Web of Science and combinations of the keywords “acoustic,” “classif*,” “recogn*,” “autom*,” and “song.” We found and reviewed 68 papers that used computers to automatically scan audio recordings and identify detections of target species, including birds, frogs, and mammals (Supplementary Materials Appendix 1). We performed an initial review of these papers to determine recognizer type (single or multiple species), and evaluation data type (clip or recording; Supplementary Materials Appendix 1, Table 1). We excluded multispecies recognizers from further review because multiclass evaluation generally employs a different set of metrics than single species evaluation (Sokolova & Lapalme 2009). We also excluded papers that did not use a test dataset of unedited field recordings (Potamitis et al. 2014) to evaluate their recognizer. The final subset included 12 single- species recognizer papers with a real-world evaluation (Supplementary Materials Appendix 1, Table 1).

Results

Benchmark: Eleven papers used human data processing as the benchmark for recognizer evaluation, and one was unclear about the benchmark used. Of the 11 that specified the benchmark, 8 used detections that had been annotated during human listening, 2 used events that had been annotated during visual spectrogram scanning, and 1 used events that had been annotated during listening and visual spectrogram scanning, i.e., two benchmarks. One paper also included a decibel level threshold as part of their benchmark (Katz et al. 2016).

Score threshold: Score threshold is a user-selected parameter that is the minimum score of any given hit reported by the recognizer. Of the 12 papers reviewed, 7 described the score threshold selected. Of those seven, four papers reported selecting a single score threshold after tests such as Youden's J statistic (Youden 1950; Swiston & Mennill 2009; Ganchev et al. 2015; Ulloa et al. 2016; Crump & Houlahan 2017), two reported choosing low thresholds that allowed for analysis of metrics across score values (Digby et al. 2013; Katz et al. 2016), and one reported a comparison of three score thresholds (Brauer et al. 2016). Two of those seven papers also reported receiver operating characteristic (ROC) metrics (Katz et al. 2016; Ulloa et al. 2016), which incorporate scores from 0 to 1 implicitly. Of the other five papers that did not report score threshold, four mentioned score but did not report threshold used (Waddle et al. 2009; Bardeli et al. 2010; Potamitis et al. 2014; Jahn et al. 2017) and one did not mention score at all (Duan et al. 2013).

All papers that examined the performance of the recognizer across score values reported that the performance improved with increasing score. Digby et al, (2013) found that recall (true negative rate) varied from nearly 100% at high scores to 0% at low scores. Similarly, Katz et al. (2016) showed that recall and specificity (the proportion of true negatives) ranged from 0 to 1 depending on the chosen score threshold. Brauer et al. (2016) compared three different score thresholds, “low” (minimized false negatives), “medium” (balanced false negatives and positives), and “high” (minimized false positives), and found that the total error of the recognizer ranged from 30% for the low threshold to 18% for the high threshold.

Metrics: In total, 11 different metrics were used across the 12 papers reviewed (Table 2.1). The most frequently used metrics were recall and precision. Among the metrics used, we found a lack of standardization and clarity in the 12 papers reviewed. There was variation in the terminology used for the metrics, with synonyms for 4 of the 12 metrics, and up to 4 synonyms per metric. In particular, the term “accuracy” was used to describe precision and accuracy; however, the formula for accuracy used in the papers we reviewed differs from the formula defined in the classifier evaluation literature (Sokolova & Lapalme 2009; Sokolova et al. 2006). Furthermore, “accuracy” was undefined in one of the papers reviewed (Duan et al. 2013), so we assigned it the same mathematical formula as the other two papers that did define accuracy. Two of the papers reviewed (Bardeli et al. 2010; Brauer et al. 2016) did not cite or define the metrics used, including “total error,” which is not a widely used classifier metric, so we back- calculated the mathematical formula or assigned the metric to the common name used in the paper. The remaining nine papers either provided the mathematical formula for the metrics used, explained the metric in plain language, or provided a citation for the metric formula.

Table 2.1. Recognizer performance metrics used in single-species recognizer studies that assessed recognizer performance on real-field recordings. TP = true positive; FP = false positive; TN = true negative; FN = false negative; β = weighting factor used to balance the weighted average of precision and recall.

Metric	Equation	Synonyms	Papers used
Accuracy	$\frac{TP - FP}{TP + FN}$		3
F-score	$\frac{(\beta^2 + 1)TP}{(\beta^2 + 1)TP + \beta^2 FN + FP}$		1
False negative rate	$\frac{FN}{TP + FP + TN + FN}$	"missed"	3
False positive rate	$\frac{FP}{TP + FP + TN + FN}$		4
Negative predictive value	$\frac{TN}{TN + FP}$		1
Precision	$\frac{TP}{TP + FP}$	"positive predictive value"; "accuracy"	7
Recall	$\frac{TP}{TP + FN}$	"correct"; "sensitivity"; "scanning comprehensiveness"	9
ROC curve			3
Total error	$\frac{FP + FN}{TP + FP + TN + FN}$		1
True negative rate	$\frac{TN}{TP + FP + TN + FN}$	"specificity"	2
True positive rate	$\frac{TP}{TP + FP + TN + FN}$		3

Recognizer Comparison Using Common Nighthawk

Methods

We used a standardized training dataset to allow for a comparison of four commercially or freely available recognizer programs. We also included one custom recognizer program to compare the other programs to the current state-of-the-art. To make this comparison useful to ecologists with minimal bioacoustic experience, we used an "out-of-the-box" approach by relying on the advice given by the program developer for recognizer construction and allowed ourselves 8–12 hours of learning time for each program. The exception was the custom CNN recognizer, which required us to write a Python script to carry out model training and evaluation.

Species: We used the common nighthawk as a model species to test single- species automated acoustic recognition software because this species has simple and consistent calls that have minimal acoustic masking from other species because nighthawks vocalize primarily at dusk and before dawn (Figure 2.1). Further, the common nighthawk vocalizes frequently, making it an ideal candidate with which to evaluate recognizer error rates in detectability and calling rate. The development of a high quality common nighthawk recognizer is also a conservation priority because this species is listed as Threatened under Canada's Species at Risk Act, and there are limited data for the species because of its crepuscular nature (Environment Canada 2016a).

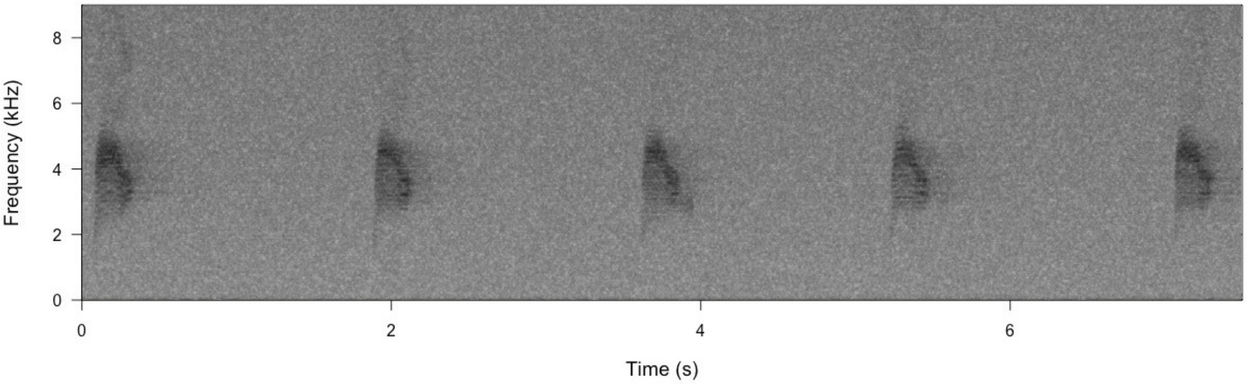


Figure 2.1. Spectrogram of common nighthawk vocalizations constructed with a 2048 FFT window size and Blackman-Harris window type.

Training Dataset: We built common nighthawk recognizers for five automated signal recognition programs using vocalizations from a standardized training dataset. The standardized training dataset consisted of 400 minutes of audio data processed by human listeners: 200 minutes of audio data with common nighthawk detections and 200 minutes of audio data with no common nighthawks. The data were collected from 11 locations in south central British

Columbia, Canada during the breeding season from 12 June to 14 July 2014 and 2015 at dawn or dusk. The absence data were collected from the same locations, but during times of year and day when common nighthawks are not active. Although common nighthawks produce relatively simple and consistent calls, there is variation between individuals (Armstrong 1965), so we hand-selected recordings to incorporate variation in call frequency, duration, and strength. All recordings were made using SM2+ or SM3 recorders (Wildlife Acoustics Inc.) with a bit depth of 16 bits, and a 16 kHz (2014) or 48 kHz (2015) sampling rate.

Song Scope Recognizer: Song Scope is a signal detection recognizer that uses Hidden Markov models (HMMs) to maximize the probability of the arrangement of individual syllables, based on the spectral feature vectors of those syllables. We built the Song Scope recognizer iteratively, following advice available in the software manual (Wildlife Acoustics 2011). First, we extracted 100 “high-quality” calls evenly distributed across 11 locations (9–10 calls from each location). We defined “high-quality” calls as calls that were produced near the microphone, i.e., had little attenuation, and were not masked by any other acoustic signals, e.g., other birds or weather. We included approximately 0.1 seconds of silence preceding and following the vocalization. We then converted the clips to Song Scope annotations and loaded them into the Song Scope software as a single class. Common nighthawk calls have frequencies below 8 kHz, so we set the sample rate at 20 kHz to exceed the Nyquist frequency (double the highest frequency of interest in the signal) with some headroom. We set the frequency minimum, range, max syllable, max syllable gap, max song, and dynamic range at values that maximized the detection of the 100 training annotations in the logarithmic scale with signal detection view (Supplementary Materials Appendix 2, Table A2.1). All other settings were left at default values. We reviewed each of the 100 training annotations to determine how much of each annotation was detected by Song Scope and removed any annotations where the full call was not completely detected. We replaced annotations with new annotations from the same location and reviewed those for detection completeness without adjusting the settings. We repeated this process until all 100 calls were completely detected in the logarithmic scale with signal detection view, and then generated the recognizer with the Song Scope software. The resultant recognizer had a cross training value of 77.32 +/- 5.87% (mean +/-SD) and a total training value of 77.22 ± 4.87% (Wildlife Acoustics 2011).

Kaleidoscope Recognizer: Similar to Song Scope, Kaleidoscope is a signal detection recognizer that builds a classification algorithm by running individual call syllables through HMMs that maximize the probability of detecting the entire call structure. Kaleidoscope differs from Song Scope in that it uses K-means clustering of Fisher scores from a 12-state HMM to cluster all the signals detected into different classes, as opposed to only identifying the signals that match the algorithm above a user-set score threshold. We built the Kaleidoscope recognizer using the cluster analysis function following the tutorial video available from the software manufacturer for “Converting Song Scope Recognizers to Kaleidoscope Cluster-based Classifiers” (Wildlife Acoustics 2016). We exported the annotation information from the 100 Song Scope annotations into a text file as presence training data. Because Kaleidoscope performs cluster analysis, it requires at least two classes to build a recognizer, so we created an absence training class by scanning our 200-minute absence dataset with Song Scope and exporting the highest scored 100 detections into the same text file. As per the training video, we then used the Kaleidoscope software to rescan the training dataset with the training clips to create a Kaleidoscope recognizer. We set maximum cluster distance to the maximum allowable value to simulate a minimum score threshold (Supplementary Materials Appendix 2, Table A2.2). We adjusted the clustering parameters to create a two-cluster recognizer with a presence class and an absence class (Supplementary Materials Appendix 2, Table A2.2). We then processed the test dataset with the Kaleidoscope recognizer using similar signal detection parameters to the Song Scope recognizer (Supplementary Materials Appendix 2, Table A2.2). We validated only those detections that were classified as presence by the Kaleidoscope recognizer and used only hits from channel 1 to prevent duplicate hits.

MonitoR Recognizer: We used the binary-point matching function in MonitoR instead of the cross-correlation approach because our initial tests suggested it was more effective for common nighthawk calls. The binary-point matching function in MonitoR is a template-based approach, where the program converts each cell of the spectrogram of a clip to a 1 or 0 using an amplitude cut-off. As a moving window recognizer, MonitoR then processes audio data by comparing this single-call template to each moving window of the data and scores how many cells the window has in common with the template. Multiple calls can be used to train MonitoR recognizers, but the program creates a template for each training call and scans the data once with each template, as opposed to other programs that aggregate the training calls and scan the data only once. We built the MonitoR recognizer following the training vignette (Hafner & Katz 2015). We used the MakeBinTemplate function to inspect each of the 100 training clips from the Song Scope training dataset, and adjusted the time limit, frequency limits, and amplitude cut-off manually for each template to ensure each call was completely detected (Supplementary Materials Appendix 2, Table A2.3).

CNN Recognizer: Convolutional neural networks (CNNs) are a class of machine learning models that have been successfully applied in a range of domains including speech recognition and visual object recognition (LeCun et al. 2015). CNNs are a type of artificial neural network (ANN) that use moving window convolutional layers to extract features from their inputs, which makes CNNs particularly suited to acoustic detection as they can be applied directly to variable length raw audio, spectrogram inputs, or other representations of sound. ANNs have previously been used for automated acoustic signal recognition, but require that call parameters are first extracted from each acoustic signal before being passed to the ANNs for classification (Jennings et al. 2008), whereas CNNs can scan and classify the spectrograms directly. In general, the filters in convolutional layers are used to detect acoustic features while sliding over the spectrogram, or other visual input. To train a CNN as a moving window recognizer, we used a simple architecture that had multiple convolutional layers, but output a single convolutional feature map (detection function) in the final layer (Supplementary Materials Appendix 2, Table A2.4). During model training we presented short clips to the network, typically with a single common nighthawk call either present or absent. We used the maximum value of the detection function to classify presence/absence, which forced the model to learn a discriminative detection function. We used the TensorFlow framework and the Python API to define and train our CNN model (Abadi et al. 2015). As input to our model, we used log-power mel-scaled spectrograms calculated using librosa (McFee et al. 2017). We used rectified linear units (ReLUs) as the activation function in all layers of the network except the last, which used a sigmoid function. We trained the network for 100 epochs with a cross-entropy cost function, using minibatch stochastic gradient descent with batch size 64 and Adam optimization (Kingma & Ba 2014) with learning rate of 0.001. During model evaluation on continuous recordings, the full time- series output of the detection function was used as the recognizer score. A simple threshold-based peak-picking method was then used to extract a list of discrete detections. The CNN model required fixed length inputs during training, so we created a dataset by manually extracting 100 clips of 2-s duration from across the presence dataset and the same number from the absence dataset.

RavenPro Recognizer: RavenPro uses band-pass filters, a band-limited energy detector, and an amplitude detector, to perform signal detection and identify calls of the appropriate duration within the frequency range of the target species. We followed the RavenPro 1.4 manual to configure our RavenPro recognizer (Charif et al. 2010). We extracted 100 high-quality calls (defined as above) and measured target signal parameters, i.e., frequency, duration, and separation, for each common nighthawk vocalization. We used the default setting for most noise estimation parameters, with adjustments made to those that increased the true positive rate (Supplementary Materials Appendix 2, Table A2.5).

Test Dataset: To test the generalizability of our recognizers, we used a test dataset from a different geographic region than the training dataset. Our test dataset comprised 117 recordings of 5-min duration (2.28 GB) from 45 study sites in northwestern Ontario, Canada. The recordings were made on 13 June and 25 June 2012 at 21:00 and 22:00 when common nighthawks are acoustically active, and there were between 1 and 4 recordings for each of the 45 study sites. The individual recordings within the test dataset were selected randomly from a larger pool of samples, though the resulting dataset represented a gradient of low to high common nighthawk call density. All recordings were collected as 16-bit 16 kHz WAV files using SM2+ Songmeters (Wildlife Acoustics Inc.).

Automated Processing The test dataset was processed with each recognizer. We chose low score thresholds for each of the recognizers so that we could evaluate performance across a gradient of score thresholds (Supplementary Materials Appendix 2). We set the score threshold at 0 for the signal detection recognizers (Song Scope, Kaleidoscope, RavenPro) to allow for full analysis of the score threshold gradient. We then ran the moving window recognizers (MonitoR and CNN) with a similarly low threshold and selected the highest scored 6750 hits, which was the maximum number of hits detected by any of the signal detection recognizers (Song Scope). Without this hit threshold, both moving window recognizers would have produced as many hits as moving windows, i.e., hundreds of thousands (Figure 2.2) because they have no signal detection process. We ran each recognizer with the same MacBook Pro (late 2013) with a 2.3 GHz Intel Core i7 and 16 GB 1600 MHz DDR3 of RAM. We timed the processing duration of the test dataset while no other software was running.

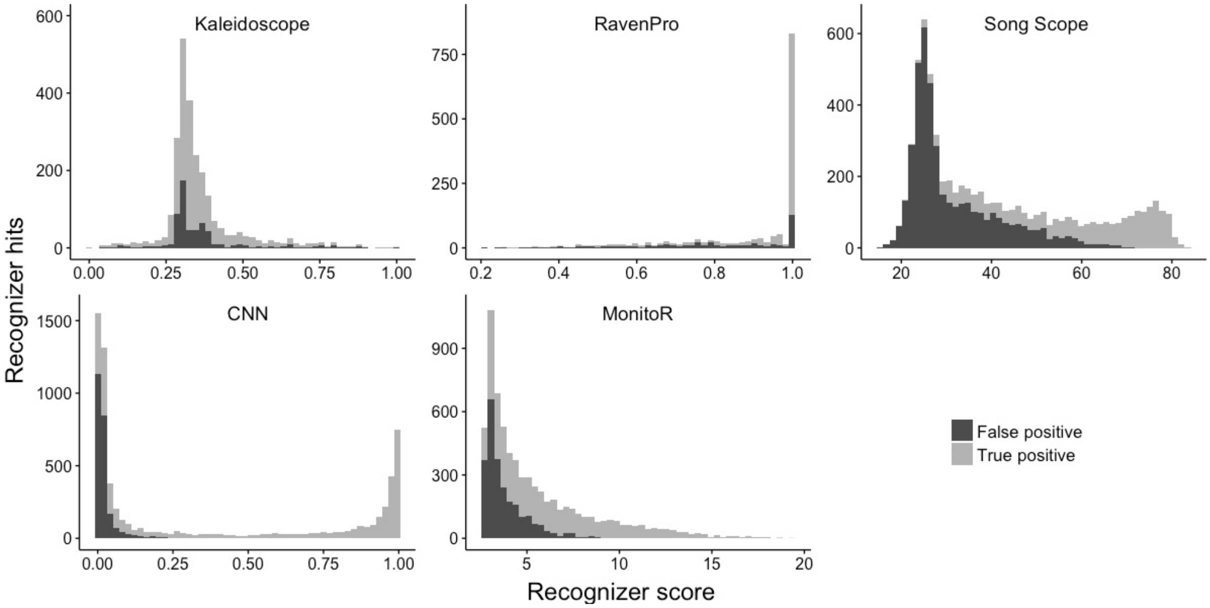


Figure 2.2. Distribution of true positive and false positive recognizer hits relative to score for common nighthawk recognizers in five different programs. The top row programs are signal detection recognizers and the bottom row programs are moving window recognizers. Recognizer scores are the raw scores reported by the programs and are unstandardized. Kaleidoscope score is the inverse of the distance metric.

Benchmark Development: We compared our recognizers to human listening and used the maximum number of true detections by any method as our benchmark because the recognizers detected the presence of common nighthawks in several recordings that human listeners had missed. Using a human listening benchmark would have decreased the presence-absence recall of those recognizers because the comparison would have been to a benchmark that included false negatives. To develop the human listening dataset, two human observers viewed and simultaneously listened to each 5-min recording in its entirety using sound visualization software and time-stamped each common nighthawk vocalization using a Microsoft Access data entry form.

Statistical Analysis: We referred to existing best practices in the machine learning literature and other acoustic signal detection disciplines to develop our evaluation approach (Davis & Goadrich 2006; Sokolova & Lapalme 2009; Raffel et al. 2014). We evaluated the overall performance of each of the five common nighthawk recognizers relative to the benchmark. We also evaluated the applied performance of each of the recognizers including presence-absence recall, occupancy modeling, and call rate correlation. All analyses were conducted in R version 3.3.1 (R Core Team 2016) with the base package, the PRROC package (Grau et al. 2015), and the ROCR package (Sing et al. 2005).

Prior to analysis, we standardized the score of each hit for each recognizer on a scale from 0 (lowest score) to 1 (highest score) to enable comparison between recognizers. We standardized the score of each hit by dividing it by the maximum score for that recognizer minus the minimum score for that recognizer. Kaleidoscope does not directly report a score, but instead uses a clustering approach to report distance between detections, so we used the inverse of the distance to cluster center as a surrogate for score. We included score threshold in our evaluation by applying a score threshold in 0.01 increments to the dataset for each recognizer before calculating each metric.

To evaluate overall performance of the recognizers, we calculated precision, recall, F-score, and area under the curve (AUC) because these metrics are suitable for one-class classifiers (recognizers trained only with examples of the target species, e.g., Song Scope, MonitoR, RavenPro) and binary classifiers (recognizers trained with examples of both the target species and nontarget species, e.g., CNN, Kaleidoscope; Sokolova et al. 2006). Precision is the proportion of recognizer hits that are true detections of the target species (Table 2.1). Recall is the proportion of target species vocalizations detected as hits by a recognizer (Table 2.1). F-score incorporates precision and recall, and allows the user to weight the relative importance of precision versus recall by setting the β value (Table 2.1). For AUC, we plotted precision-recall as well as ROC curves for each of the recognizers because some authors suggest precision-recall is more appropriate for recognizer performance evaluation (Davis & Goadrich 2006). We did not apply a score threshold for this evaluation because AUC incorporates score implicitly. We did not include human listening in AUC calculation because human listening detections do not have score values.

We then evaluated the applied performance of the recognizers and human listening in a presence-absence study because presence-absence data are used for a variety of applications in ecological research and monitoring. To simulate a presence- absence study and to balance sampling effort across study sites, we subsampled our test recording dataset to the first recording for each of the 45 study sites. We then determined whether the recognizer or listener accurately determined the presence or absence of a common nighthawk for each score threshold increment of 0.01, and then modeled this presence-absence recall with a binomial logistic regression for each processing approach. For each approach, we constructed null, first-order, second-order, and third-order polynomial models with score threshold as the covariate. We compared the four models for each approach using Akaike Information Criteria (AIC; Burnham and Anderson 2002) and selected the model with the lowest AIC score.

We also evaluated the performance of the recognizers and human listening for occupancy modeling. Occupancy modeling is a widely used application of presence/absence data that uses repeated visits to account for imperfect detection of the target species (MacKenzie et al. 2002). ARU data are particularly well- suited for occupancy modeling because they collect multiple time- series recordings that can be used as repeat-visit data (Shonfield & Bayne 2017). We modeled common nighthawk detection and occupancy for each of the recognizers and human listening using a single season occupancy model framework (MacKenzie et al. 2002) with each 5-min recording used as a separate "sampling occasion." Prior to modeling, we removed seven study sites from the dataset for which there was only one recording because occupancy models require at least two recordings, i.e., visits, to estimate the detectability parameter. The resultant dataset comprised 38 sites. We then ran a null occupancy model with the validated recognizer data for each 0.01 score threshold for each recognizer to examine how detectability and occupancy changed with score threshold.

We also evaluated the performance of each recognizer and human listening for measuring call rate. Call rate ARU data have been used for behavioral studies (Ehnes & Foote 2014), and can be used as a proxy for abundance of some species if baseline patterns in call rates or song frequency are well known, which can in turn be used for monitoring population trends (Jeliazkov et al. 2016). We calculated the Spearman correlation coefficient between the benchmark and the call rate for each score threshold increment using the individual recording as the sampling unit.

Finally, we compared the efficiency of each of the five automated acoustic recognition programs and human listening as the time required to learn the software, build the recognizer, scan the test audio dataset, and validate the recognizer results as true or false positives. We limited learning time to 8–12 hours to develop a functional aptitude for each of the programs using our "out-of- the-box" approach. We quantified the time spent to build each recognizer, including a standardized four hours of training dataset compilation time because we used a single compiled training dataset for all five recognizers. We quantified the time required to scan by timing the computer processing of our test dataset. We quantified the time to validate by timing the validation of each of the recognizer hits and taking the mean validation time per hit. To compare the efficiency of the five recognition programs to human listening, we calculated processing time in hours per hour of audio data for a 10 hour audio dataset and a 1000 hour audio dataset. We calculated processing time as the time required to learn and build the recognizer plus time to validate the recognizer results. We did not include scanning time in our efficiency calculation because this part of the process does not require human supervision. For time to validate, we calculated the time it would take to validate the recognizer when run with a score threshold for the peak of the precision-recall curve, i.e., the maximum value of precision + recall. Finally, we calculated the audio dataset size at which the efficiency of recognizer processing becomes faster than human listening, assuming 1 hour of listening per 1 hour of audio data and 1 hour of initial learning.

Results

A total of 5556 common nighthawk calls were detected across the 117 five-minute recordings (mean = 152 per recording, SD = 196), which was used as the benchmark for recognizer evaluation. Common nighthawks were detected in 85 of the 117 recordings, and at 38 of 45 sites from northwestern Ontario, Canada.

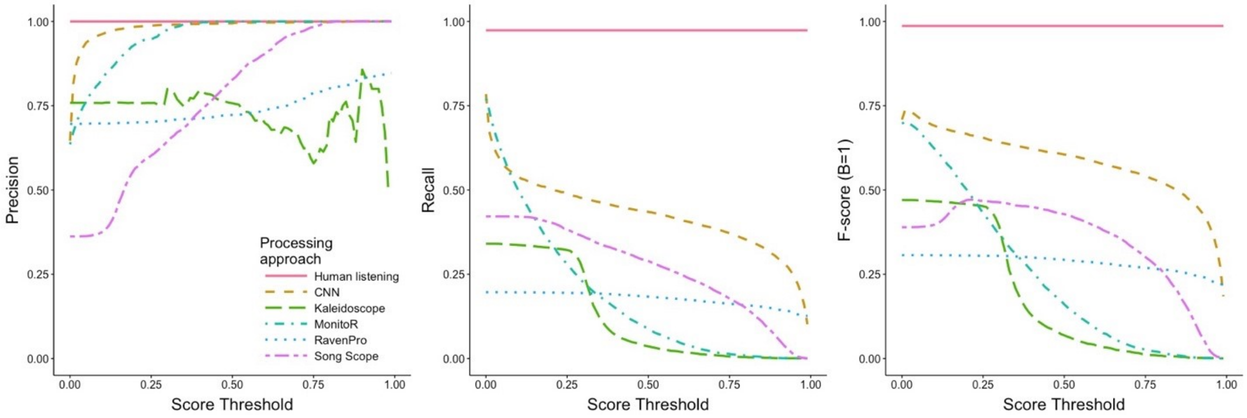


Figure 2.3. Precision, recall, and F-score of common nighthawk call detection for automated acoustic recognition programs at varying score thresholds. Precision, recall, and F-score of human listening is provided for comparison. Precision is the proportion of recognizer hits that are true detections of the target species. Recall is the proportion of target species vocalizations detected by the recognizer. F-score combines precision and recall into a single evaluation metric.

Precision, Recall, and F-score: As expected, recall and F-score decreased and precision increased with increasing score threshold for all recognizers (Figure 2.3). Score threshold had a minimal impact on precision and recall of the RavenPro recognizer, with impacts seen only at score thresholds above 0.7. Precision for the CNN, MonitoR, and Song Scope recognizers neared 1.0 at high score thresholds, with few false positives reported by the two moving window recognizers (CNN and MonitoR) except at low thresholds. The Kaleidoscope and RavenPro recognizers both had a precision of approximately 0.7 across most score thresholds. The CNN had the highest recall across all score thresholds, with the MonitoR recognizer also reaching a high recall of 0.75 at low score thresholds. The Song Scope recognizer recall decreased steadily from 0.42 at the lowest threshold, while the Kaleidoscope recognizer recall of 0.34 dropped off rapidly above a score threshold of 0.3. The RavenPro recognizer had relatively low recall of approximately 0.2 across all score thresholds. The F-scores of the five recognizers were similar to the recall values, with the exception of a lower F-score for the Song Scope recognizer below a score of 0.2. Human listening precision was 1.0 because we assumed that every human listener detection was a true positive; however, human listening recall was 0.97 because human listeners missed 146 of the 5556 common nighthawk calls detected in the test dataset.

Area Under the Curve: The CNN recognizer had the highest precision-recall curve AUC (0.94), followed by MonitoR (0.88), Song Scope (0.87), RavenPro (0.82), and Kaleidoscope (0.77; Figure 2.4). The ranking of the top two recognizers from the ROC curve AUC was different than the precision-recall curve AUC; the SongScope recognizer had an AUC of 0.90, while the CNN had an AUC of 0.88. The ranking of the other recognizers was the same between the two AUC measures; however, the ROC AUC of the Kaleidoscope recognizer (0.53) was much lower than the precision-recall AUC (0.77).

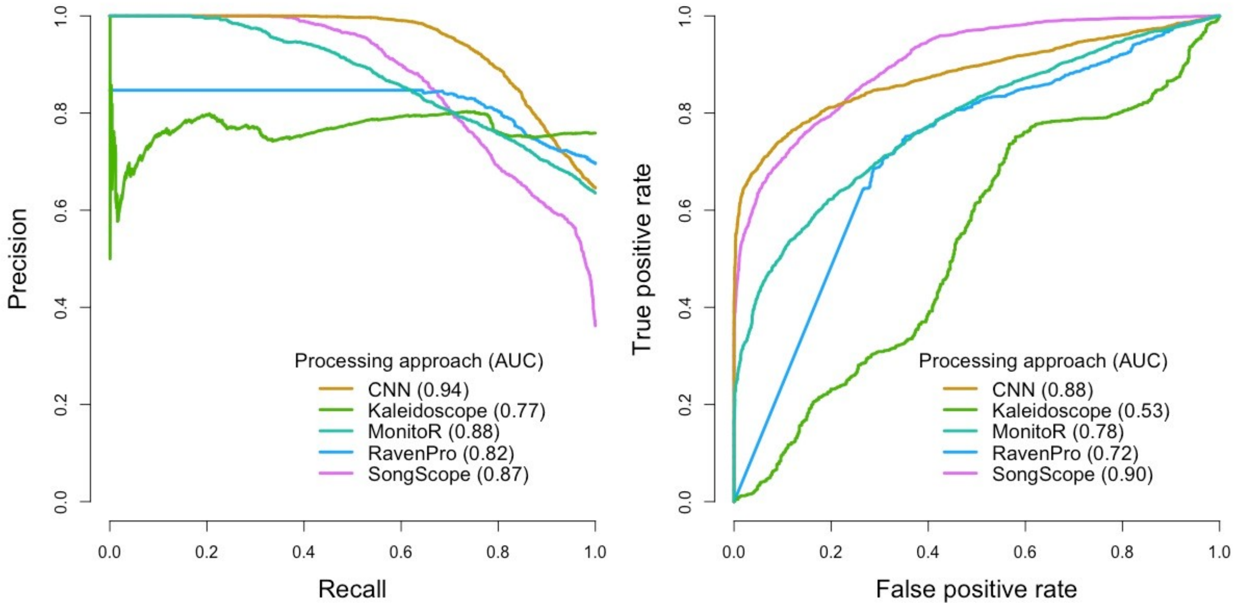


Figure 2.4. Precision-recall curve (left) and receiver operating characteristic (ROC; right) curve of common nighthawk call detection for automated acoustic recognition programs. AUC is area under the curve for each program.

Presence-absence: At low score thresholds, the CNN, Song Scope, and MonitoR recognizers determined common nighthawk presence-absence with similar recall as a human listener (95.4%; Figure 2.5). At high score thresholds, only the CNN and RavenPro recognizers detected common nighthawk presence-absence with greater than 50% recall. As with precision and recall, score threshold had little impact on the presence-absence recall of the RavenPro recognizer. The CNN recognizer had the highest presence-absence recall of the five programs across the score threshold gradient. The CNN (AIC weight = 0.95), Kaleidoscope (AIC weight = 0.92), and Song Scope (AIC weight = 0.97) recognizers were modeled as third-order polynomials, and the MonitoR recognizer (AIC weight = 0.69) was modeled as a second-order polynomial (Supplementary Materials Appendix 3, Table A3.1). The null model with the lowest AIC score for the RavenPro recognizer was the null model (AIC weight = 0.43), suggesting that score threshold had no effect on presence-absence recall.

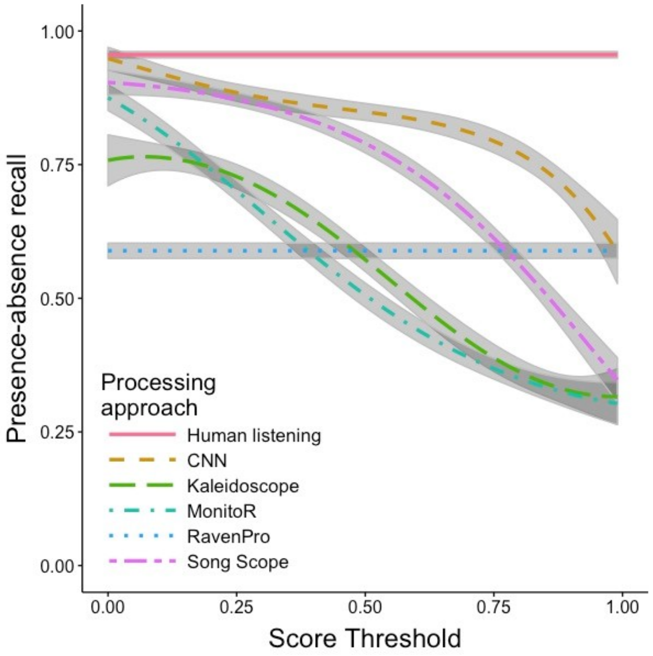


Figure 2.5. Recall of five automated acoustic recognition programs for detecting common nighthawk presence per recording at varying score thresholds. Recall of human listening is provided for comparison. Shaded areas indicate 95% confidence intervals.

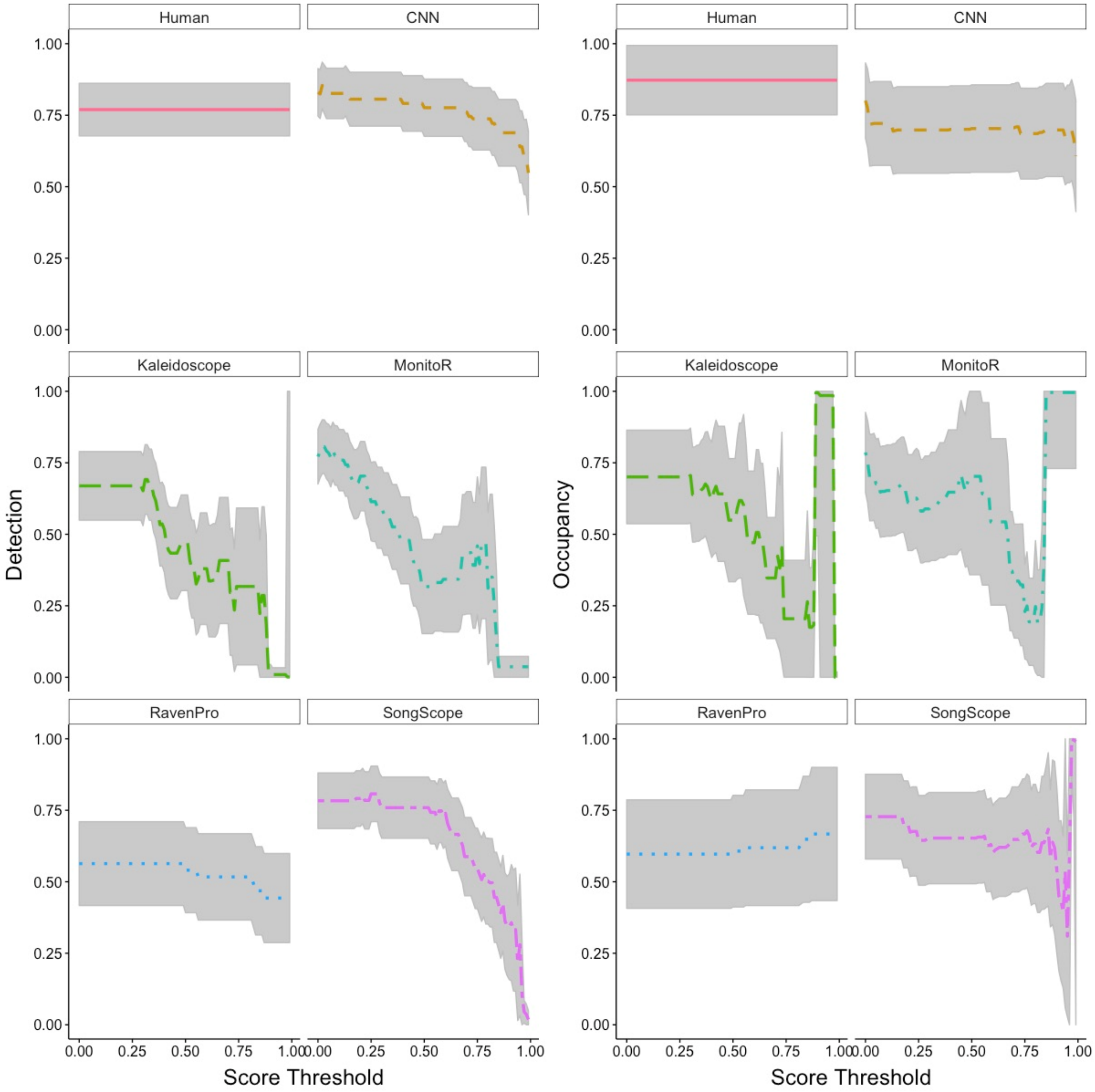


Figure 2.6. Common nighthawk occupancy and detection in null occupancy models for automated acoustic recognition programs at varying score thresholds. Occupancy and detection of human listening is provided for comparison. Shaded areas indicate 95% confidence intervals.

Occupancy: Naive occupancy of the 110 visits, i.e., recordings, included in occupancy modeling was 0.89 (34 of 38 sites). The occupancy estimate from human listening was 0.87 (SE = 0.06; Figure 2.6). In general, the occupancy estimates from recognizer data were lower than the estimate from human listening, although the occupancy estimate from the CNN recognizer

(0.80) was not significantly so. The occupancy estimates from the Kaleidoscope, MonitoR, and Song Scope recognizers decreased with increasing score threshold as detection also decreased, and at high score thresholds, the estimates became unstable, varying between 0 and 1. The occupancy estimates from the CNN and the RavenPro recognizers were more stable across score thresholds, although the RavenPro estimate was much lower (0.60).

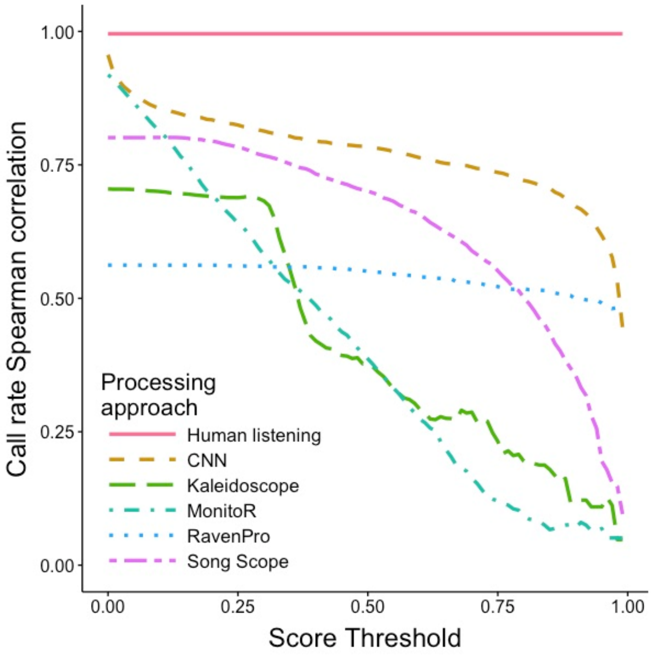


Figure 2.7. Spearman correlation of common nighthawk call rate between automated acoustic recognition programs across varying score thresholds. Correlation of call rate from human listening is provided for comparison.

Call Rate: At low score thresholds, the CNN and MonitoR call rate was similar to human listening (0.96 and 0.91 correlation, respectively); however, call rate correlation of the MonitoR recognizer decreased rapidly and linearly to near 0 with increasing score threshold, while the CNN recognizer call rate correlation decreased slowly before dropping steeply at a score threshold of 0.9 (Figure 2.7). The Song Scope recognizer call rate correlation was between 0.7 and 0.8 at moderate score thresholds. Call rate correlation for the RavenPro recognizer varied minimally across score thresholds (max = 0.56, min = 0.48). The Kaleidoscope call rate correlation was 0.7 and decreased steadily but irregularly after a score threshold of approximately 0.3.

Efficiency: All five of the automated signal recognition programs became faster than human listening for datasets larger than 36 hours of audio (Table 2.2). The CNN recognizer had the largest initial time investment, and thus had the highest processing time per hour of audio data for a small dataset (10 hours audio). For a large audio dataset (1000 hours audio) the differences between the recognizers were due primarily to differences in the number of hits at maximum precision-recall between recognizers. The Song Scope recognizer was the most efficient, while the Kaleidoscope recognizer was the slowest. Although not included in the processing time calculations, scanning time should also be included in efficiency considerations. The CNN and Kaleidoscope recognizers were the fastest to scan our test dataset, while the MonitoR recognizer was two orders of magnitude slower because this program scanned the audio dataset separately through each of the 100 templates.

Table 2.2. Time in hours (h) spent to learn each of the automated acoustic recognition programs, build a recognizer, scan audio recordings with the recognizer, and validate the recognizer output. Total times and dataset size were calculated using the number of hits produced by each recognizer when the score threshold is set to maximize accuracy.

Recognizer	Learn time	Build time	Scan time per h audio	Validate time per h audio	Total time per h audio (10 h dataset)	Total time per h audio (1000 h dataset)	Dataset size (h) where recognizer is faster than human listening
Human listening	1	0	0	1	1.1	1.00	NA
CNN	24	8	0.003	0.11	3.31	0.14	36
Kaleidoscope	8	4	0.001	0.16	1.76	0.17	19
MonitoR	8	8	0.32	0.52	2.20	0.22	25
RavenPro	8	2	0.03	0.13	1.50	0.12	16
Song Scope	12	8	0.03	0.11	2.48	0.11	26

Evaluation Recommendations

Based on our analysis, we suggest that ecologists who use automated acoustic recognition for processing acoustic recordings follow six general recommendations. These suggestions are drawn largely from best practices in machine learning and other acoustic signal processing disciplines (Salzberg 1997; Sokolova et al. 2006; Sokolova & Lapalme 2009; Raffel et al. 2014), as well as our literature review of evaluation methods in ecology and lessons learned during our common nighthawk recognizer evaluation. We also suggest that ecologists familiarize themselves with general machine learning practices because there is great potential for interdisciplinary research, but a known lack of communication between the two disciplines (Thessen 2016).

Recommendation 1: Benchmark

Recognizer evaluation should employ a test dataset that differs from the training dataset to avoid "overly optimistic" results (Salzberg 1997). Within the test dataset, it is important to establish a benchmark of known target species detections to evaluate recognizer performance. We recommend human listening as a comparison benchmark; however, we remind readers that human listening is also subject to error (Bart & Schoultz 1984; McClintock et al. 2010; Brauer et al. 2016). If any false negatives in human detections are discovered during the process of reviewing recognizer detections, we recommend instead using the maximum number of target species detections detected by any method, i.e., human processing or a recognizer, as the benchmark. In our performance evaluation, there were 146 common nighthawk calls (2.63% of total) detected by a recognizer that were missed by human listeners. Brauer et al. (2016) also reported a 2% error rate in human identification of anuran calls, while Rydell et al. (2017) found error rates ranging from 9–22% for bat species identified by human listeners. If the target species vocalizations are susceptible to false positive identification by human observers, we recommend using a dependent double observer method when developing the benchmark to reduce the probability of misidentification (Forcey et al. 2006). Acoustic signals at farther distances (Skowronski & Fenton 2009), lower sound pressure (Jahn et al. 2017), or with low signal-to-noise ratios, i.e., high levels of background noise, will be difficult to detect for both humans and recognizers, and therefore should not be excluded when preparing a benchmark (Skowronski & Harris 2006). Human listening can also be subject to observer bias (Sauer et al. 1994). Jennings et al. (2008) found that human observers with less than a single year of experience performed worse at classification than recognizers. Human annotation error can also be reduced by using the consensus from multiple observers as the benchmark dataset (Drake et al. 2016).

Recommendation 2: Score Threshold

We strongly recommend that the influence of score be included in recognizer evaluation because our review showed it has a fundamental impact on recognizer performance, no matter what metric was used. Following Katz et al. (2016), we further recommend the use of score threshold instead of the reported raw scores of each detection in recognizer evaluation so that ecologists can use their evaluation results to select an optimal score threshold for data processing. We found in both our own recognizer evaluation and in our review of the literature that performance varied widely with score threshold. Furthermore, not all papers that used recognizers reported how they selected their score threshold despite the importance of this decision. Factors such as project objective, recording quality, call complexity, and signal clarity influence the choice of score threshold and the subsequent performance metrics. In our evaluation, the exception was the RavenPro recognizer, whose performance was largely unaffected by score threshold, perhaps because RavenPro is a band limited energy detector that identifies signals based only on a frequency range specification. It is possible that score threshold may be particularly important for programs with more complex classification approaches. Inclusion of a gradient of score thresholds in evaluation will facilitate selection of an appropriate score threshold for further analysis, which can be chosen based on the objectives of the project (Katz et al. 2016). We also found that some papers did not report score threshold, and we argue that it is crucial that score thresholds are explicitly reported within papers that use automated signal recognition.

Recommendation 3: Metrics

We suggest ecologists use metrics that are considered best practice in other signal processing disciplines (Sokolova & Lapalme 2009). Specifically, we suggest that four metrics always be reported for single species recognizer evaluation: (1) precision, (2) recall, (3) F-score, and (4) area under the curve (AUC). These metrics are regularly reported during classifier evaluation in other disciplines and will also allow ecologists to compare evaluation results with state-of-the-art studies in machine learning and elsewhere. Ecologists can also calculate these statistics across multiple datasets or partitioned datasets so that variance in metrics can be evaluated (Salzberg 1997) and statistical tests to compare recognizer performance can be applied (Dietterich 1998; Demšar 2006).

Precision and Recall: Precision is the proportion of recognizer hits that are true detections of the target species and is calculated as

$$Precision = \frac{tp}{tp + fp}$$

where *tp* is the number of true positives (detections of target species) and *fp* is the number of false positives (recognizer hits that were mislabelled as the target species).

Recall is the proportion of target species vocalizations detected as hits by a recognizer and is calculated as

$$Recall = \frac{tp}{tp + fn}$$

where *fn* is the number of false negatives (detections of the target species in the benchmark dataset that the recognizer missed). Precision and recall were the most commonly used metrics in our literature review and in the classification literature (Raghavan et al. 1989; Provost et al. 1998; Davis & Goadrich 2006). Precision and recall are appropriate for signal recognition evaluation because unlike some metrics, they do not require quantification of true negatives (i.e., other species), which are not reported in single-class recognizers such as Song Scope and MonitoR. In contrast, accuracy focuses on true and false negatives and assumes that false negative and positive errors are equally likely and consequential, which is often a poor assumption in signal recognition (Provost et al. 1998). Precision and recall are also particularly appropriate when the target species is rare, as a recognizer can have a high accuracy by simply predicting the target species is always absent, and the accuracy of a recognizer can be inflated by adding more negative examples to the dataset. Using precision and recall allows for direct comparison of recognizer performance with other published studies. Across the studies we reviewed, the mean recall was 0.60 and the mean precision was 0.71 (Swiston & Mennill 2009; Bardeli et al. 2010; Digby et al. 2013; Potamitis et al. 2014; Ganchev et al. 2015; Jahn et al. 2017) With the exception of the Kaleidoscope recogizner and the Song Scope recognizer at low score thresholds, the precision of our common nighthawk recognizers was above 0.71. The recall of our MonitoR and CNN recognizers reached 0.60 at low score thresholds, but the other recognizers did not.

F-Score: F-score combines precision and recall into a single metric and is calculated as

$$Fscore = \frac{B^2 * precision * recall}{B^2 * precision + recall}$$

where B is a user-defined metric that allows for prioritization of precision over recall, or vice-versa. Precision and recall are evenly balanced when $B = 1$, precision is favoured when $B > 1$, and recall is favoured when $B < 1$ (Sokolova et al. 2006). We recommend that if ecologists choose to use a value for other than 1, that they also report F-score with $B = 1$ to allow for comparison across studies. Situations where ecologists might consider using $B < 1$ including detection of rare species or situations with legal implications.

Area Under the Curve (AUC): Following other acoustic signal processing disciplines, we recommend reporting the AUC of a precision-recall curve as a univariate method for comparing recognizers. Receiver operating characteristic (ROC) curve AUC is more commonly used in the classifier evaluation literature; however, precision-recall curves are more appropriate for cases with class imbalance such as recognizer evaluation (Davis & Goadrich 2006). In other words, a large quantity of false positives, as is the case for many recognizers at low score thresholds, is more accurately reflected in the AUC of a precision-recall curve than an ROC curve, and our comparison of the two approaches supports this. We therefore recommend a precision-recall AUC; however, ecologists may also want to calculate an ROC AUC for comparison with other published studies.

Recommendation 4: Application Evaluation

Although overall recognizer evaluation is important, the influence of the metrics chosen can depend on the intended application for the data (Stowell et al. 2016). We therefore also recommend evaluation be done for the intended application of the resultant species detection data. Recognizer evaluation for occupancy modelling purposes is particularly important, as our results suggest this approach becomes unreliable for recognizer data with low recall because species detection probability is too low for reliable occupancy estimates at low recall (MacKenzie et al. 2002). We also found that the shape of the curve across the score threshold gradient for all three response variables we examined (presence-absence recall, occupancy estimate, and call rate correlation) was similar to the shape of the recall curve. Future work should investigate whether the relationship between the shape of the score-recall curve is an adequate proxy for all response variables, or whether it varies depending on the detectability, call rate, and occupancy of the target species.

Recommendation 5: Regional Generalizability

Geographic variation in acoustic signal is demonstrated in many bird species (Slabbekoorn & Smith 2002) and other animals which produce sound (Pröhl et al. 2006; Campbell et al. 2010; Sun et al. 2013), which is important to consider during recognizer evaluation (Russo & Voigt 2016; Gillespie et al. 2013). For simplicity, we evaluated the regional generalizability of our common nighthawk recognizer with a test dataset from a different region than the training data; however, in best practice, ecologists should test recognizers across multiple geographic regions. Evaluating with multiple test datasets will help ecologists determine whether a single recognizer is effective or whether regionally-specific recognizers are required for their target species. For example, marine mammal classifiers have been shown to be 14.4% less accurate when tested with data from a different region than the training data (Erbs et al. 2017). For ecologists that plan to use recognizers for a single region, training and test data should be sourced from the region of interest.

Recommendation 6: Efficiency Evaluation

For many ecologists, the purpose of employing an automated signal recognition approach is to increase the efficiency of audio data processing; therefore, we recommend collecting data on time spent to build and run a recognizer and validate the output. The time per hour of audio data can then be compared to other data processing approaches, including human listening. For our recognizers, we found that human listening became less efficient with datasets larger than 36 hours of audio; however, we note that using a visual scanning approach (i.e., viewing the spectrogram) instead of listening may have improved the efficiency of our human processing approach. If the automated recognizer used performs poorly, however, the manual post-processing time required may outweigh the advantages of automation due to the time required to validate the results (Stowell et al. 2016). Digby et al. (2013) found that automated recognition (2 minutes per hour of recording) could be at least as or more efficient than manual scanning (2 – 5 minutes per hour of recording) but noted that the efficiency of a recognizer will depend on the species' vocalization characteristics, call rate, and the quality of recognizer. Indeed, human listening may be more efficient than single-species recognizers if multiple species data are needed from audio recordings; however, there are also many multi-species recognizer approaches currently under development (Stowell et al. 2016, Goëau et al. 2017). Ultimately, relative efficiency will depend on a variety of factors including score threshold, with more time required to validate recognizer output if a low score threshold is chosen to prioritize recall over precision.

Discussion

Autonomous recording units are important tools for ecological monitoring and research because they are portable, collect data over extended periods, can be used in remote locations, are not restricted to a particular season, and the data they collect can be archived as a permanent record (Shonfield & Bayne 2017). The use of automated signal recognition for processing ARU data is growing because it can reduce the time required to process the large amounts of data; however, the best practices are needed (Blumstein et al. 2011). In particular, recognizer performance evaluation is a critical step for projects that employ automated signal recognition. All recognizers misclassify detections to some extent, which can have implications for study results and may lead to poor management decisions if the results are not validated (Russo & Voigt 2016; Rydell et al. 2017). In our review of the bioacoustics literature, we found little similarity in recognizer performance evaluation between studies. Some studies reported minimal performance evaluation results, which renders the ecological results of these studies difficult to interpret. In papers that did report performance evaluation, we found an inconsistency in the evaluation terminology used and a lack of reference to the classification literature (Salzberg 1997; Davis & Goadrich 2006; Sokolova & Lapalme 2009). Given the increasing use of recognizers by ecologists, these deficiencies suggest a need for guidance on performance evaluation. We used best practices from other acoustic signal processing disciplines and our own evaluation of automated signal recognition software to provide recommendations for comparing and recognizers.

Using the common nighthawk as a model species, we found that a convolutional neural network (CNN) recognizer outperformed the other recognizers across all evaluations, but that Song Scope and MonitoR recognizers had similar precision and recall rates at some score thresholds. Currently, the construction of CNN recognizers requires programming expertise, but an increasing number of authors have reported success with this method for automated signal recognition (Koops et al. 2014, Salamon et al. 2016, Salamon and Bello 2017). Using our "out-of-the-box" approach, we found MonitoR and Song Scope had similar learning curves, assuming the operator is already familiar with the R programming language. At the time of writing, however, Song Scope was no longer under development or supported by the manufacturer. As the simplest automated signal recognition program, RavenPro was the easiest to learn, but the simplicity of its band-width delimitation classification approach limited its performance. Duan et al. (2013) also compared Raven Pro and Song Scope, and similarly reported a more intuitive user interface. Duan et al. (2013) also found that RavenPro had higher recall but lower precision than Song Scope. The Kaleidoscope recognizer also had low precision and recall relative to the other recognizers, with precision varying erratically across score threshold, likely because we used distance to cluster centre as a surrogate for score. Rydell et al. (2017) similarly found that Kaleidoscope performed worse than other recognizers for bat call classification. We caution that our performance and efficiency evaluation of these five programs was based on a single model species with a simple, diagnostic call and little ambient masking noise and that ecologists should compare these programs for other species before choosing which program to use for audio data processing.

Overall, automated signal recognition was effective for determining common nighthawk presence-absence and call rate, particularly at lower score thresholds, but the occupancy estimates from the data processed with recognizers were consistently lower than derived from human listening, with the exception of the CNN recognizer. Other authors have successfully derived occupancy estimates from recognizer data that are comparable to naïve occupancy (Kalan et al. 2015; Campos-Cerqueira & Aide 2016). Although ARUs can be as effective as human surveyors at detecting occurrences (Holmes et al. 2014; Kalan et al. 2015), the greater number of false negatives from an automated analysis (Brauer et al. 2016) reduces the apparent occupancy estimate for an organism at a location (MacKenzie et al. 2002). It has been suggested that the difference in recall between automated signal recognition and human listening is caused by a smaller detection radius of the recognizer relative to the human listener (Jahn et al. 2017, *unpublished data*), which could be due to both the signal detection and classification components of the recognizer and would explain our reduced occupancy estimates. This may not be an error per se but may instead reflect the fact that more standardization is needed when using ARUs to determine the effective area being sampled (Yip et al. 2017a). We also found that occupancy estimates became unstable at high score thresholds with low recall, and therefore caution against the use of occupancy models produced from recognizer data with low recall recognizers because low recall contributes to low detectability, which biases occupancy estimates (MacKenzie et al. 2002). Future research should investigate the sensitivity of occupancy modelling to this new data type.

Although automated signal recognition is effective for common nighthawks, there is little consensus to date on the overall effectiveness of the existing technology for avian ecological research and monitoring. Future application of our recommendations would be most useful for taxa with more complex acoustic signals, different calling rates, and in environments with varying levels of ambient noise. Thorough performance evaluation in recognizer studies following our general recommendations will contribute to building a body of literature for future meta-analysis on the overall effectiveness of automated signal recognition for wildlife monitoring and research.

Supplementary Materials

Supplementary materials are available at https://www-ace-eco-org.login.ezproxy.library.ualberta.ca/vol12/iss2/art14/.

Chapter 3. Classification threshold and training data affect the quality and utility of focal species data processed with automated audio recognition software

Abstract

Automated recognition is increasingly used to extract information about species vocalizations from audio recordings. During processing, recognizers calculate the probability of correct classification (“score”) for each acoustic signal assessed. Our goal was to investigate the implications of recognizer score for ecological research and monitoring. We trained four recognizers with clips of common nighthawk (*Chordeiles minor*) calls recorded at different distances: near, midrange, far, and mixed distances. We found distance explained 49% and 41% of the variation in score for the near and mixed-distance recognizers, but only 3% and 6% of the variation for the midrange and far recognizers. We calculated detection functions for each of the recognizers at various score thresholds and found that the detection function for the near and mixed-distance recognizers satisfied the assumptions of density estimation for most score thresholds, while the detection function for the midrange and far recognizers did not. The detection functions also showed that score threshold choice is a decision about sampling area, not just about the balance between recall and precision. Overall, we showed that training recognizers with ‘high-quality’ clips that were recorded at close range will improve the utility of the data without affecting how many true positives the recognizer detects.

Introduction

Autonomous recordings units (ARUs) are increasingly used by wildlife ecologists and managers to survey for animal species that communicate with acoustic signals (Shonfield &

Bayne 2017). In particular, ARUs are used to survey for focal species that are of conservation or management concern because they are a non-invasive method that can record time-series data in a cost-effective manner (Drake et al. 2016). ARU recordings can also provide a permanent record that can be used to verify identification of rare species (Swiston & Mennill 2009; Holmes et al. 2014) or reanalysed later to study other species (Derryberry et al. 2018). ARU data can be used for a variety of ecological research applications for focal species including monitoring population trends (Furnas & Callas 2015), behavioural studies (Ehnes & Foote 2014), occupancy modelling (Chambert et al. 2017), density estimation (Marques et al. 2012), and habitat modelling (Campos-Cerqueira & Aide 2016).

Recordings collected by ARUs require processing to extract information about detections of the focal species, which is time consuming to do by listening or visual scanning if large volumes of recordings are collected. Automated recognition can be an efficient way to extract species detection information from large bioacoustic datasets (Stowell et al. 2016; Shonfield & Bayne 2017; Priyadarshani et al. 2018). Automated recognition is the process of training a computer to detect and classify a focal species' vocalization. The training process produces a detection algorithm (hereafter "recognizer") that can then be run over audio recordings to classify the acoustic signals therein. For each signal evaluated, the recognizer assigns a classification probability (hereafter "score"), which can be interpreted as a measure of the probability that the signal being evaluated can be classified as the focal species. The recognizer then registers a detection for each signal with a score above a user-defined threshold (hereafter "score threshold").

A variety of automated recognition approaches have been developed over the past decade and a half. Some of the common approaches include random forests (Aide et al. 2013; Campos-

Cerqueira & Aide 2016), hidden Markov models (Skowronski & Harris 2006; Potamitis et al. 2014; Oliveira et al. 2015) and/or Gaussian mixture models (Heinicke et al. 2015; Ganchev et al. 2015), binary point matching (Katz et al. 2016), spectrogram cross- correlation (Katz et al. 2016), artificial neural networks (Tachibana et al. 2014; Nicholson 2016), decision trees (Digby et al. 2013), and band-pass filters (Charif et al. 2010). Some of the approaches mentioned employ a single step process that runs the algorithm against every window of an audio recording and reports a score value for each window (hereafter "moving window recognizer"). Others use a two-step process that first conducts signal detection with a moving window, and then runs the algorithm only on detected signals (hereafter "signal detection recognizer"). Regardless of the process, all single-species recognizers calculate a score metric and thus require setting a score threshold to separate detections of the focal species from other non-target signals.

Studies using recognizers have typically treated score as a measure of classification probability; thus, score threshold choice is usually described as a subjective decision based on the priorities of the user (Wildlife Acoustics 2011; Katz et al. 2016) to balance precision (i.e., minimizing false positives) and recall (i.e., minimizing missed detections). Some authors have used statistical tests such as Youden's J statistic (Youden 1950) to select a score threshold (Swiston & Mennill 2009; Ganchev et al. 2015; Ulloa et al. 2016; Crump & Houlahan 2017), but score is often selected arbitrarily. Score threshold can have a substantial impact on the performance of a recognizer (Katz et al. 2016; Brauer et al. 2016) and yet is not reported in most papers that use recognizers to extract species detection information from audio recordings (Chapter 2).

The application of signal processing to ARU surveys for ecological applications must be considered in the context of the properties of sound. Sound attenuates with distance in a predictable way in the absence of environmental conditions that affect spherical spreading. As the call of a focal species attenuates, the signal fades and distorts, becoming dissimilar to the training data in a similarly predictable way. We predict that if a recognizer is trained with audio clips of the focal species that were recorded at close range, that recognizer score will have a predictable relationship with the distance at which a sound is recorded. If score has a predictable relationship with distance, then score threshold is a decision about sampling area when recognizers are trained with clips recorded at close range. One of the drawbacks of using recognizers to date is that the perceived inability to determine sampling radius. Without knowledge of sampling radius, ARU data are unsuitable for density estimation (Dawson & Efford 2009).

Our goal was to investigate the implications of score threshold for ecological research and monitoring. We tested whether score has a predictable relationship with distance, and if so, whether it depends on how the recognizer was trained. First, we trained multiple recognizers with clips recorded at known distances. We then modelled score as a function of distance with polynomial regression for each of the recognizers. Next, we explored the generalizability of the relationship between score and distance by including random effects and weather covariates in the top model for the recognizer trained with clips recorded at close range. Finally, we examined how recognizer training data and score threshold affect the probability of detecting a call at various distances, and thus the suitability of the processed data for distance sampling.

Methods

Model Species

We used the common nighthawk (*Chordeiles minor*) as a model species to test our hypothesis because this non-passerine species has a simple and consistent call that is effectively recognized by a variety of automated recognition programs (Figure 3.1; Chapter 2). Understanding recognizer data quality and utility is also a conservation priority because this species is listed as Threatened under Canada's Species at Risk Act, and there are limited data for the species because it is active during crepuscular periods when standard bird surveys are not typically conducted (Canada 2016a).

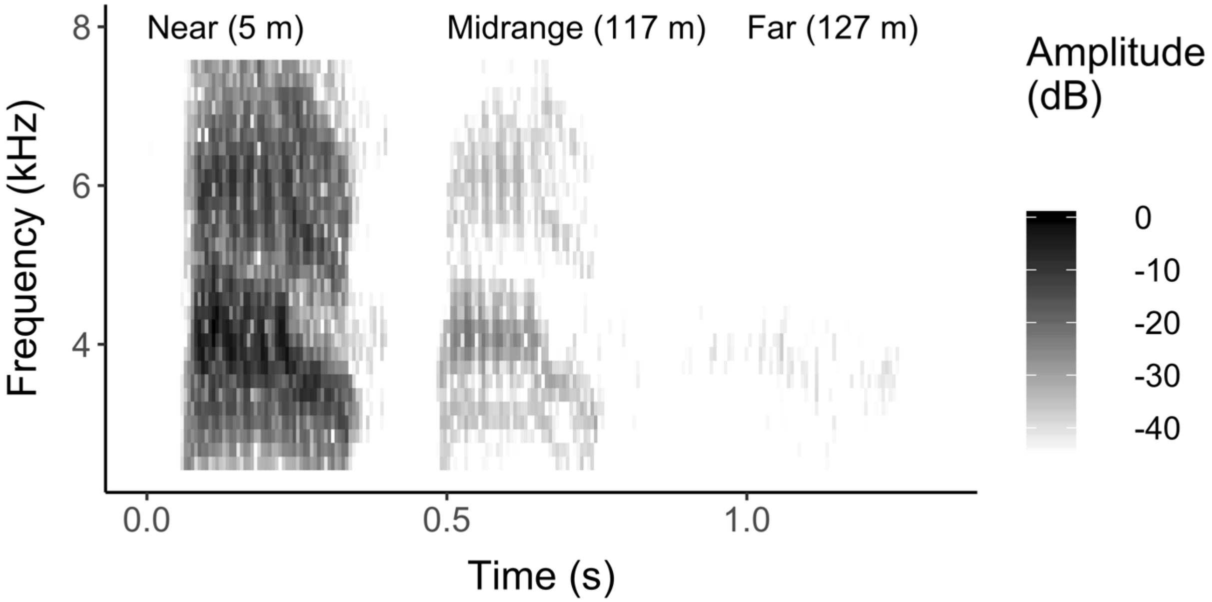


Figure 3.1. Spectrogram of the same common nighthawk vocalization recorded at multiple distances (near, midrange, far) and used to build recognizers. Spectrogram constructed with the same parameters used for recognizer construction.

Audio Data Collection at Known Distances

We collected audio recordings of common nighthawks with known locations by attracting a territorial male to the beginning of a transect of ARUs using conspecific broadcast calls. An observer stood at the beginning of the transect and recorded the time stamp, height, horizontal distance, and bearing of every vocalization from the target individual. We minimized distance estimation error by using the same observer for all observations, by the observer calibrating their distance estimates with a laser range finder prior to every observation period, and by limiting observations to those within 20 m horizontal distance of the observer because human observer distance estimation error is minimized at short distances (Nadeau & Conway 2012). We collected recordings on five transects (i.e., of five individual males) between July 13 and July 20, 2016, starting at 1 hour before sunset and ending at sunset. Each transect consisted of 11 ARUs placed at standardized distances along a linear feature (30 m, 60 m, 90 m, 120 m, 150 m, 180 m, 210 m, 240 m, 300 m, 400 m, 500 m). We measured temperature, wind speed, and humidity during each survey using a Kestrel 3000 (Kestrel Meters, Minneapolis, MN, USA). We played an airhorn at the start of the recording period from the beginning of the transect.

Following acoustic data collection, we clipped each of the recordings at the airhorn to synchronize the target vocalizations therein. Next, we visually confirmed the time stamp of each vocalization of the target individual and identified any vocalizations that were masked by the broadcast call. We then used the seewave package (Sueur et al. 2008) in R to clip each unmasked detection from each of the 11 recordings along the transect as 0.7 s clips. We reviewed the clipped vocalizations and removed any sets of vocalizations where the target individual was masked by other individuals or background noise. The final dataset comprised 64 vocalizations at each of the 11 distances, or 704 clips in total. See Yip et al. (2020) for detailed acoustic data collection methods.

Recognizer Construction & Processing

We used Song Scope software (Wildlife Acoustics, Maynard, MA, USA) to train recognizers with the acoustic clips of known distances. Song Scope is a signal detection recognizer that extracts Mel Frequency Cepstral Coefficients from each detected signal and computes the overall score using Hidden Markov Models (Wildlife Acoustics 2011). Although Song Scope was recently discontinued by its manufacturer, we chose it because it outperformed all other 'out-of-the-box' recognizer programs in a recent comparison for our model species (Chapter 2), including its replacement software Kaleidoscope (Wildlife Acoustics, Maynard, MA, USA). Despite its deprecation, Song Scope remains freely available from the manufacturer and continues to be effectively used for wildlife research and monitoring (Chambert et al. 2017; Venier et al. 2017; Shonfield & Bayne 2017).

We trained each recognizer with 50 clips of common nighthawk vocalizations from the known distance dataset. We trained three single-distance recognizers (near, midrange, far; Figure 3.2) with the same 50 vocalizations recorded at different distances, and one mixed-distance recognizer with the same 50 vocalizations, but randomly and evenly selected from the three single-distance training datasets. We used a standardized set of signal detection parameters for all three recognizers (Supplementary Materials) and ensured that each of the training clips was fully recognized by the signal detection process in Song Scope. We ran each recognizer over the full dataset of known distance clips using a score threshold of 0. The first author validated the recognizer results by reviewing each recognizer hit visually and aurally to confirm the clip was of the target individual for that transect.

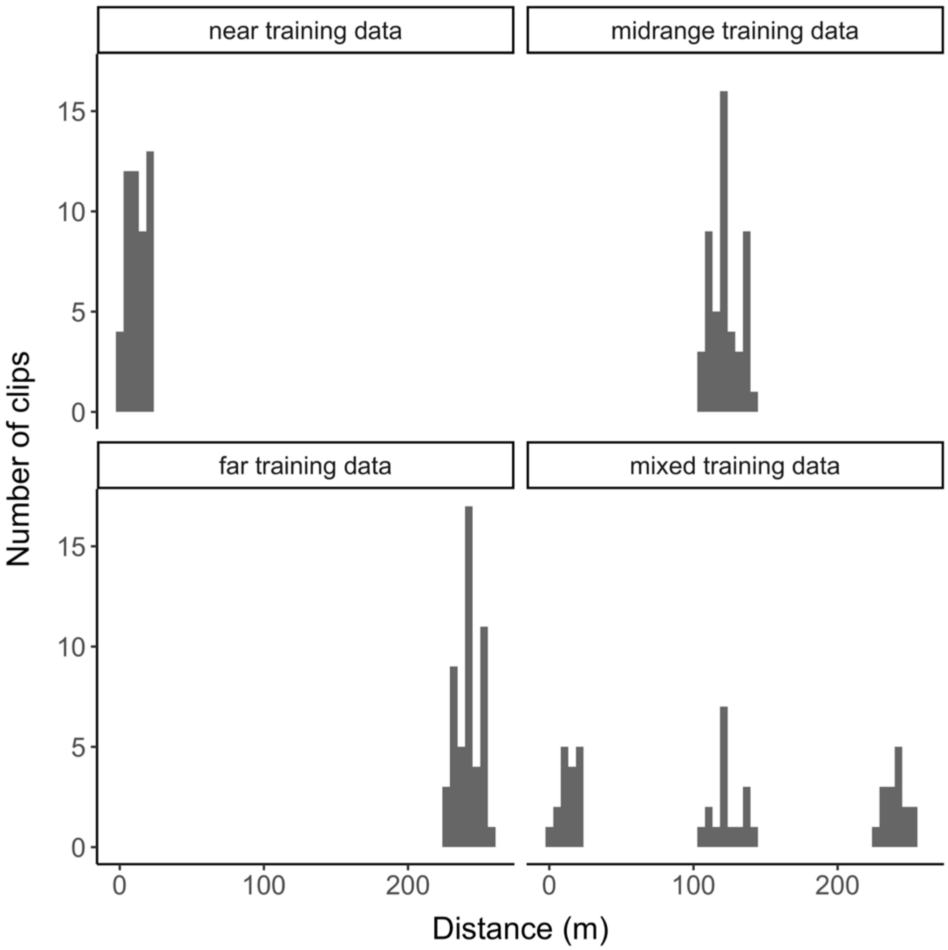


Figure 3.2. Recording distance of training clips used to build four different recognizers to detect common nighthawk calls.

Statistical Analysis

We tested whether distance was a significant predictor of score for each of the near, midrange, far, and mixed-distance recognizers using polynomial linear models. Prior to fitting each model, we removed the fifty clips that had been used in training to avoid including over-fitted score values in the analysis. For each of the four recognizers, we ran a null model and models with first, second, and third order polynomials for distance. We ranked models using small sample size corrected Akaike's Information Criterion (AICc) and selected the most parsimonious model with ΔAICc <2 from the top model as the best fitting model.

We then selected the recognizer that had the strongest relationship between score and distance (the near recognizer) and added weather covariates and random effects to the best fitting model to explore the generalizability of the relationship between score and distance. We included fixed effects for temperature, wind speed, and humidity because weather can affect sound attenuation, and a random effect for vocalization ID nested within individual bird ID because recognizer score can vary between individuals. We tested for significant collinearity between fixed effects by calculating the variance inflation factor (VIF) for each predictor. We retained all predictors for model selection because the maximum VIF was 1.70. We compared a global model to simplified models using AICc and selected the most parsimonious model with ΔAICc <2 from the top model as the best fitting model. We used the sigma estimates from the best fitting model to calculate the interclass correlation (ICC; i.e., the correlation in score value), for the random effects (vocalization and individual bird).

Finally, we modelled the probability of detection as a function of distance (i.e., the detection function) (Buckland et al. 2015) for each of the recognizers to determine the suitability of the processed data for distance sampling (Sòlymos et al. 2013; Buckland et al. 2015). We used a generalized linear model with a binomial response to estimate the detection function because our data were not a random sample of distances, but a series of binary data for non-randomly placed ARUs on the transect (Marques et al. 2009; Buckland et al. 2015). To model the detection function, we followed a half-normal detection function using a generalized linear model with a binomial distribution and a complementary log-log link (“cloglog”) function (Sòlymos et al. 2013). We did not fix the intercept at 1 because we wanted to test whether the probability of detection was near 1 at zero metres, which is a key assumption of distance sampling (Buckland et al. 2015). We used the entire dataset of 704 clips and modelled whether or not the recognizer had reported a hit for each clip as the binomial response variable with the negative squared distance of the clip as the predictor variable. We then created additional models using different score thresholds (40, 45, 50, 55, 60, 65, 70) by removing any hits from each dataset that were below the chosen score threshold. We also calculated the recall for each recognizer for each score threshold as the number of detections divided by the total number of clips processed (704). Recall is the proportion of calls that were detected by the recognizer and is a recommended metric for assessing recognizer performance (Chapter 2, Priyadarshani et al. 2018).

All analyses were conducted in R version 3.4.3 (R Core Team 2017) using the packages lme4 (Bates et al. 2015) and usdm (Naimi et al. 2013).

Results

The best fitting model for the prediction of score with distance was a second order polynomial for the near and mixed-distance recognizers and a third order polynomial for the midrange and far recognizers (Supplementary Materials; Figure 3.3). Distance from all four recognizers were significant predictors of recognizer score (all $P < 0.001$); however, distance explained more of the variation in score for the near and mixed-distance recognizers than for the midrange and far recognizers (near: R2 = 0.49; midrange: R2 = 0.03; far: R2 = 0.06; mixed: R2 = 0.41; Figure 3.3).

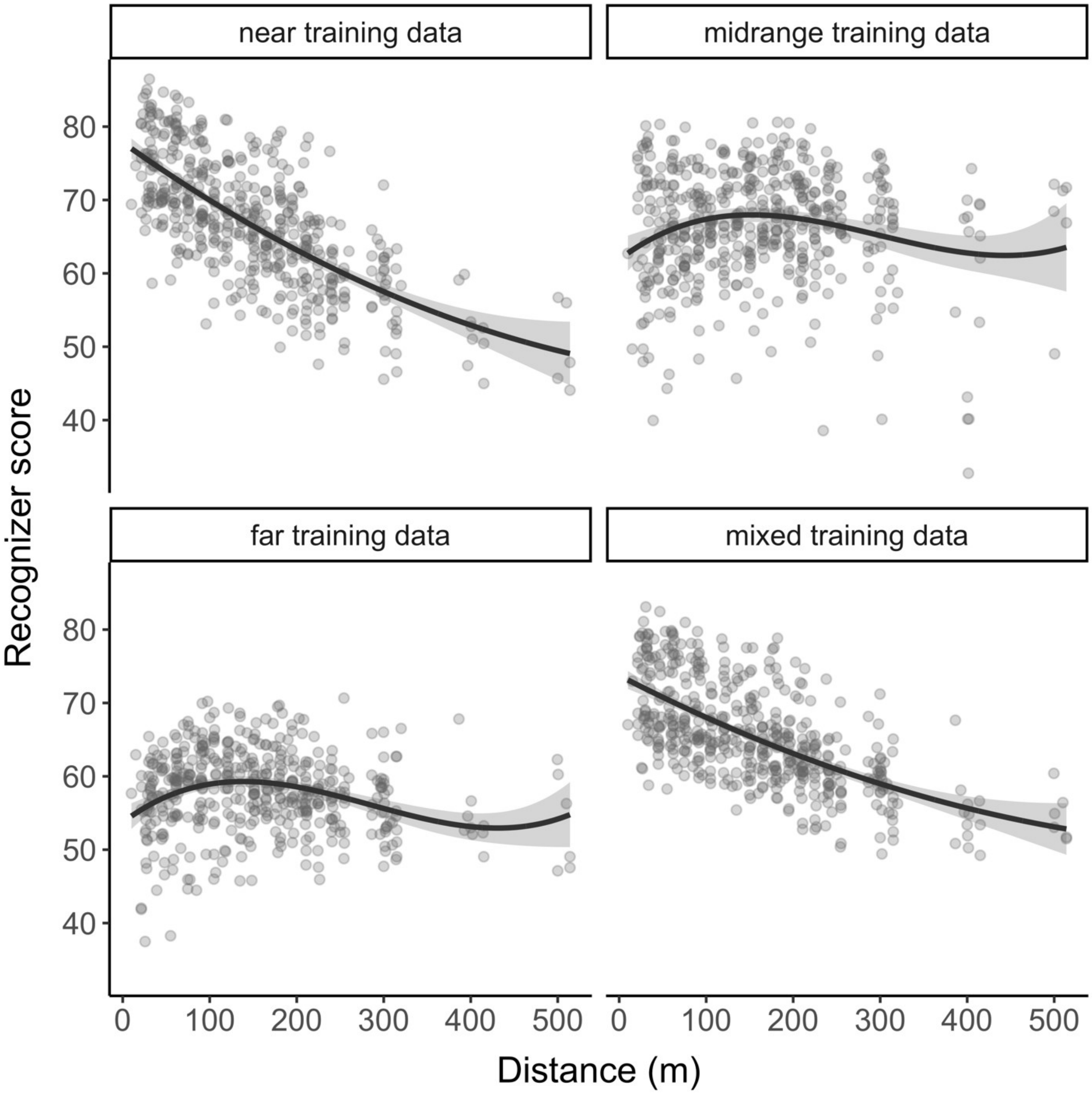


Figure 3.3. Relationship between recognizer score and distance of common nighthawk detections from audio clips for four recognizers built with training data of differing known detection distances. Lines and 95% confidence intervals are model predictions from polynomial models and are plotted against the raw data.

When we added weather covariates and random effects to the second order polynomial model for the near recognizer and compared models with AICc, none of the models with weather covariates had strong support (Table 3.1). The top ranked model was the model with no weather covariates (ΔAICc = 0.57). Both random effects had strong ICCs: the ICC for vocalization was 0.67 and the ICC for individual bird was 0.53 (Figure 3.4).

Table 3.1. AICc ranking of models for prediction of distance with score of common nighthawk detections from acoustic data processed with automated acoustic recognition. All models included a second order polynomial for vocalization distance as a fixed effect and vocalization nested within individual bird as random effects. Bold indicates the model selected as the most parsimonious model with ΔAICc <2 from the top model.

Model	df	logLik	AIC$_c$	ΔAIC$_c$	AIC$_c$ w
score = null	**6**	**-1339.0**	**2690.1**	**0.00**	**0.27**
score = temperature	7	-1338.8	2691.8	1.69	0.12
score = wind speed	7	-1338.5	2691.3	1.20	0.15
score = humidity	7	-1338.2	2690.7	0.57	0.20
score = temperature + wind speed	8	-1338.5	2693.2	3.15	0.06
score = temperature + humidity	8	-1338.0	2692.4	2.28	0.09
score = wind speed + humidity	8	-1338.0	2692.4	2.30	0.09
score = temperature + wind speed + humidity	9	-1337.9	2694.3	4.18	0.03

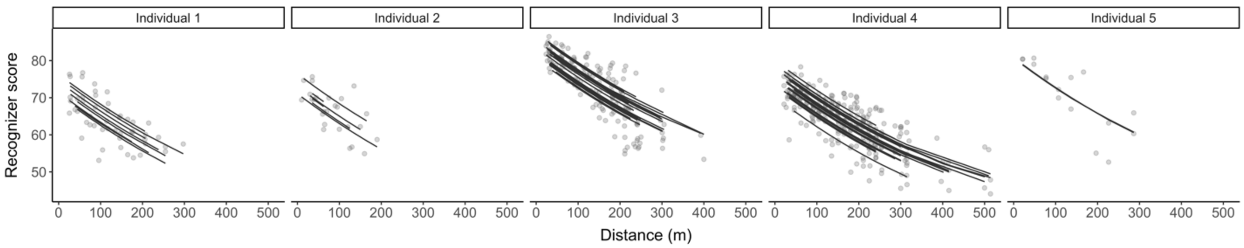


Figure 3.4. Relationship between recognizer score and distance of common nighthawk detections from audio clips of vocalizations of five individual birds recorded at multiple distances. Lines are model predictions for each vocalization from polynomial mixed effect models and are plotted against the raw data.

Each of the recognizers detected vocalizations in approximately 75% of the 704 clips that were processed when there was no score threshold applied (i.e., only the signal detection process was applied). The near recognizer detected 495 vocalizations, the midrange recognizer detected 513 vocalizations, the far recognizer detected 488 vocalizations, and the mixed-distance recognizer detected 510 vocalizations. When score threshold was applied, recall declined faster for the far recognizer than for the other three recognizers, which had similar recall across all score thresholds applied (Figure 3.5).

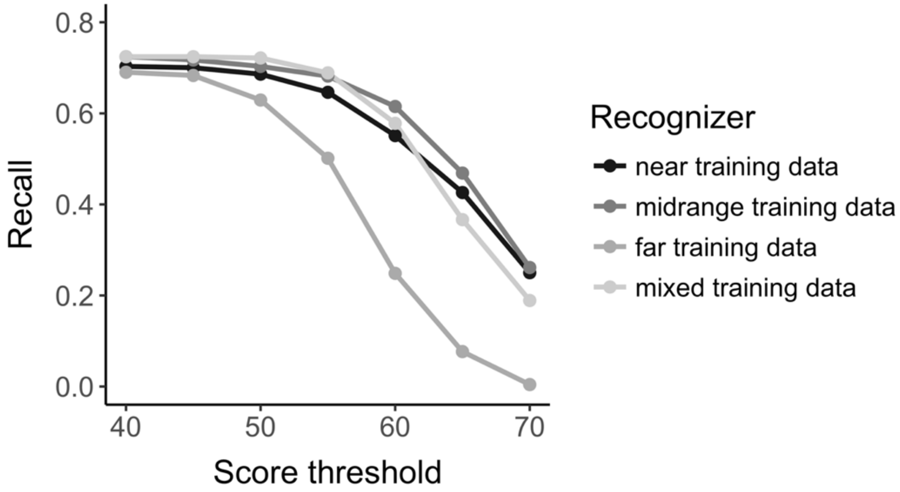


Figure 3.5. Recall of common nighthawk call detection from audio clips using recognizers built with training data of differing known detection distances and run with varying score thresholds. Recall is the number of correctly detected common nighthawk vocalizations divided by the total number of clips scanned with the recognizer.

The probability of detection at distance = 0 was near 1 and the 95% confidence intervals for the detection overlapped for all four recognizers when there was no score threshold applied (i.e., only the signal detection process was applied). When a score threshold was applied to the near recognizer, the probability of detection at distance = 0 was near 1 for all score thresholds less than 70 (Figure 3.6). As score threshold increased, the maximum distance at which the recognizer was able to detect calls decreased. The mixed-distance recognizer behaved similarly to the near recognizer, except that the probability of detection at distance = 0 was less than 1 for score thresholds of 65 and 70. For the midrange and far recognizers, the probability of detection at distance = 0 was near 1 only for low score thresholds (40, 45). For score thresholds of 50 and greater, the probability of detection at distance = 0 decreased with increasing score threshold.

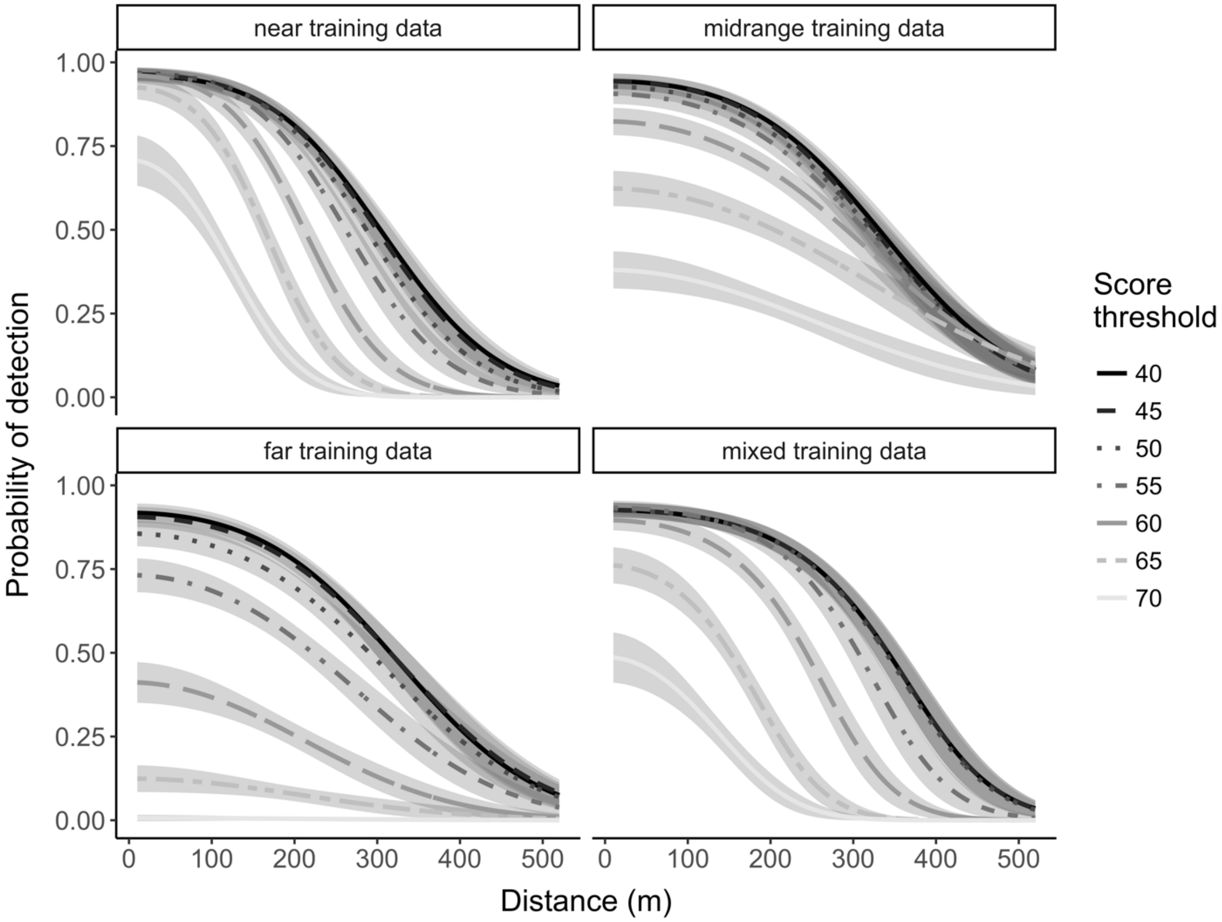


Figure 3.6. Probability of detecting a common nighthawk call from audio clips using recognizers built with training data of differing known detection distances and run with varying score thresholds. Lines and 95% confidence intervals are model predictions from binomial detection data.

Discussion

We showed that the probability of correct classification reported by a recognizer, or score, has a predictable relationship with distance if the recognizer is trained with clips recorded at minimal distances or with clips evenly distributed across a range of distances. In contrast, little of the variation in score values reported by recognizers trained with clips recorded at midrange or large distances were explained by distance. When we applied a variety of score thresholds to the processed results of each of the recognizers and modelled the detection function for each, we found that the maximum distance at which the recognizer was able to detect calls decreased, which shows that applying a score threshold also limits the sampling area for those two recognizers. Overall, our results have two important implications. First, that choosing a score threshold will affect the sampling area. Second, that training data for recognizers affects the quality and utility of the results.

Score threshold impacts the performance of a recognizer due to the trade-off between recall and precision. Both Katz et al. (2016) and Brauer et al. (2016) have shown that low score thresholds produce more false positives, while high score thresholds increase the probability of missing a detection of the focal species. We have also previously shown that high score thresholds may be inappropriate for occupancy modelling due to the low number of detections produced (Chapter 2). Due to the trade-off between recall and precision, choosing a score threshold has previously been regarded as a subjective decision based on the priorities of the user; however, our results suggest that score threshold should also be regarded as a decision about sampling area. Several studies have shown that recognizer recall is lower than that of a human listener unless minimal score thresholds are used (Katz et al. 2016; Chapter 2). Our detection functions show that this difference in recall is caused by a smaller detection radius of the recognizer relative to the human listener, which has also been suggested by Jahn et al. (2017). Lower recall compared to human listening is, therefore, not due to classification errors per se, but reflects the fact that the effective area being sampled should be determined when using ARUs (Yip et al. 2017b). If researchers are interested in quantifying the sampling radius of their recognizer for a particular score threshold, we recommend developing a known distance dataset, modelling the detection function for that score threshold, and calculating the distance at which the number of individual birds detected outside is equal to the number of missed individuals within (Buckland et al. 2005).

The amplitude of a sound at a particular distance is known to be affected by humidity, temperature, and wind because those weather conditions affect the amount of sound attenuation (Harris 1966). We therefore expected that the fade and distortion of the spectral signature of sound would similarly be affected by weather conditions, which would be reflected in the score of the recognizer detection. We may not have detected an effect of weather conditions because we did not sample a wide range of temperature, humidity, and wind over the five days of acoustic data collection. Instead, much of the variation in the relationship between score and distance was attributed to the characteristics of the individual bird and vocalization, with greater than 50% intraclass correlation for both random effects. Individual variation in vocalizations has been shown in many bird species (Linhart & Šálek 2017) and has also been suggested for the common nighthawk (Armstrong 1965). The mono-syllabic nature of the common nighthawk call likely limits the extent of variation compared to other species, especially passerines that learn their songs and often have individual-specific repertoires of more than one song (Catchpole & Slater 2008); therefore, the relationship between score and distance will likely be more variable for those species. Researchers should ensure that known distance datasets reflect the same amount of variation in song as the recognizer to accurately determine the relationship between score and distance. We did not test for an effect of vegetation type but note that vegetation will likely alter the relationship between score and distance via sound attenuation and thus should also be considered when building known distance datasets.

We showed that the relationship between score and distance facilitates density estimation, which is a key objective for wildlife management (Buckland et al. 2005, 2015). ARUs have previously been criticized because of the inability to determine sampling area (Dawson & Efford 2009); however, we showed here that the detection function of data processed with near or mixed-distance recognizers satisfies the assumption that the probability of detection is near one at zero metres. This assumption is important for density estimation methods that model probability of detection as a function of distance sampling (Sòlymos et al. 2013; Buckland et al. 2015). The detection functions of the recognizers built with midrange and far clips did not satisfy this assumption when score values were greater than 50. When score thresholds were applied to both these recognizers, hits at all distances were classified as negative detections because there was no linear, negative relationship between score and distance, which resulted in a probability of detection of less than 1 at zero meters. While the relationship between score and distance for the near and mixed-distance recognizers results in data that is suitable for density estimation, we do not recommend using score to estimate distance to a particular individual because the variation introduced by any given individual may introduce bias in results. Instead, we have previously shown that relative sound level, or the loudness of a particular sound, is a robust proxy for distance (Yip et al. 2020).

Due to the variation in approaches to automated recognition, some of our results are likely generalizable to other automated recognition approaches and some are not. We suggest that the relationship between distance and score is likely generalizable because it is based on the principle of spherical spreading and sound attenuation, but that the shape of the relationship may depend on the feature extraction and classification methods. The shape of the detection function, however, may depend more on the signal detection performance of the recognizer. We showed that when no score threshold was applied (i.e., all detected signals were included), the detection function for all four recognizers was suitable for density estimation. The detection function for moving window recognizers may not follow the same pattern. The detection function will also be less ideal for recognizers applied to full-length field recordings because the recall rate of the recognizers will be reduced by sound masking from other noises in the recording (Priyadarshani et al. 2018). Future research should investigate the extent of generalizability for the relationship between score and distance, particularly for moving-window recognizers.

In general, recognizers are trained with ‘high quality’ audio clips (Venier et al. 2017; Shonfield et al. 2018). ‘Quality’ is a subjective and undefined term; however, most authors imply it has two components: amount of background noise and recording distance (Priyadarshani et al. 2018). Previous studies have investigated the impact of background noise and found that developing a training dataset with minimal background noise improves the precision of the recognizer (Wildlife Acoustics 2011, Duan et al. 2013). We are unaware of any studies that have investigated the distance component of quality. We found that the recall of the recognizer was not affected by the distance of the training data unless far training data was used. We note, however, that we did not investigate the precision, or false positive rate of the recognizers. We also found that the implications and utility of the data differ between the distance of the recognizer training data. Although both the near and mixed-distance recognizers showed a relationship between distance and score, the relationship was strong for the near recognizer, and thus it satisfied the assumptions of distance sampling for a larger range of score thresholds than the mixed-distance recognizer. We are unsure whether the relationship between score and distance for mixed-distance recognizers is specific to the balanced training data design we used. Our results therefore confirm the existing standard to train recognizers with clips recorded at close range, or ‘high quality’ data (Priyadarshani et al. 2018).

The continued accrual of large bioacoustic datasets in many parts of the world emphasizes the increasing importance of automated methods to extract focal species information from those datasets. Many approaches to automated recognition have been attempted and their performance compared, but the processed data is rarely evaluated for specific ecological applications. Species detection data can be used for a wide variety of ecological applications, ranging from habitat models (Campos-Cerqueira & Aide 2016) to studying phenology (Willacy et al. 2015) to detecting rare species (Sidie-Slettedahl et al. 2015); however, each of those applications comes with a separate set of assumptions. We showed here for the first time that the distance of training data and the score threshold applied to focal species recognizers can impact the meaning and utility of the processed data, particularly with respect to density estimation. We therefore encourage practitioners to think carefully about the desired data application before choosing how to train their recognizer. For most applications, training recognizers with 'high quality' clips that were recorded at close range will improve the utility of the data without affecting the recall performance of the recognizer.

Supplementary Materials

Additional Methods & Results

Table 3.A.1. Parameter settings for recognizers built in Song Scope software.

Parameter	Setting
FFT size	256
FFT overlap	½
Frequency minimum	40
Frequency range	80
Amplitude gain (dB)	0
Background filter (s)	1
Max syllable (ms)	432
Max syllable gap (ms)	0
Max song (ms)	432

Parameter	Setting
Dynamic range (dB)	35
Algorithm	2.0
Maximum complexity	32
Maximum resolution	8
Score threshold	0
Quality threshold	20

Table 3.A.2. AICc ranking of polynomial models for prediction of distance with score of common nighthawk detections from acoustic data processed with automated acoustic recognition. Recognizers were trained with vocalizations recorded at three different distances. “Mixed” indicates that the recognizer was trained with vocalizations from all three distances. Bold indicates the model selected as the most parsimonious model with ΔAICc <2 from the top model.

Recognizer	Model	df	logLik	AIC$_c$	ΔAIC$_c$	AIC$_c$ w
Near	Distance = null	1	-1662.6	3329.1	313.33	0.00
Near	Distance = score	2	-1507.5	3021.0	5.17	0.05
Near	**Distance = score + I(score^2)**	**3**	**-1503.9**	**3015.8**	**0.00**	**0.69**
Near	Distance = score + I(score^2) + I(score^3)	4	-1503.8	3017.8	2.02	0.25
Midrange	Distance = null	1	-1667.9	3339.9	11.81	0.00
Midrange	Distance = score	2	-1667.7	3341.5	13.41	0.00
Midrange	Distance = score + I(score^2)	3	-1662.0	3332.0	3.95	0.12
Midrange	**Distance = score + I(score^2) + I(score^3)**	**4**	**-1659.0**	**3328.1**	**0.00**	**0.88**
Far	Distance = null	1	-1439.8	2883.6	27.21	0.00
Far	Distance = score	2	-1437.6	2881.3	24.93	0.00
Far	Distance = score + I(score^2)	3	-1430.4	2826.9	12.50	0.00
Far	**Distance = score + I(score^2) + I(score^3)**	**4**	**-1423.1**	**2856.4**	**0.00**	**1.00**
Mixed	Distance = null	1	-1622.7	3249.4	253.41	0.00
Mixed	Distance = score	2	-1622.7	2999.2	3.21	0.11
Mixed	**Distance = score + I(score^2)**	**3**	**-1493.9**	**2995.9**	**0.0**	**0.56**
Mixed	Distance = score + I(score^2) + I(score^3)	4	-1493.5	2997.1	1.11	0.32

Chapter 4. The Big Boom Theory: Common nighthawks use wing-boom displays to defend nesting territories

Abstract

Research and monitoring of most landbirds uses auditory cues from displaying males for determining habitat relationships. Unfortunately, using cues that signal territory boundaries can underestimate home range size and exclude extraterritorial habitat from critical habitat designation, environmental impact assessments, or other habitat studies. We show that the common nighthawk wing-boom display is a territorial signal associated with the nest location that can be used to differentiate territorial from home range habitat use. The common nighthawk is a poorly understood but widespread species whose populations are declining across most of North America. Precise information about habitat requirements and space use is required to fulfill conservation objectives but is difficult to obtain due to the extreme mobility of this species. We captured, tagged, and tracked 21 male common nighthawks in northeastern Alberta to confirm the biological significance of the wing-boom display and describe common nighthawk territoriality. Mean wing-boom use density (hereafter "area") size was 10.2 ha (SD=11.7 ha). We found minimal overlap in wing-boom area (5 of 15 neighboring male pairs, 0.2%-4.5% overlap), suggesting the wing-boom display represents an exclusive territory. Comparison of wing-boom locations and random points within the wing-boom area confirmed that male common nighthawks select areas near the nest to perform wing-boom displays. There was high wing-boom area overlap for the same individual between years. Differences between years reflected shifts in nest location, suggesting that the wing-boom display is a good indicator of the nest location and territory. Future common nighthawk surveys should record acoustic detection type to differentiate between territorial and extraterritorial habitat use. Failure to

incorporate biological significance in habitat studies could lead to management actions that do not fully satisfy the resource requirements of species of conservation concern.

Introduction

Research and monitoring of most avian landbird species has long focused on using auditory cues from displaying males for a variety of applications, most notably population size, trend estimation and understanding habitat relationships (Ralph et al. 1993; Hudson et al. 2017; Rosenberg et al. 2017). For many species, the song is a signal produced to communicate territory boundaries to conspecifics and attract a mate (Catchpole & Slater 2008). In other words, males sing at the edges of their territories to delineate and defend an exclusive area (Nice 1941; Odum & Kuenzler 1955; Maher & Lott 1995). This defended territory has been considered the 'fundamental unit of space' for passerines and other landbird species because many field ornithology methods like point counts and territory mapping rely on visual and auditory observation of males during conspicuous territory displays (Whitaker & Warkentin 2010).

In recent years, radiotelemetry and GPS tracking have facilitated following the movement of birds undertaking less perceptible behaviors. In contrast to the territory, the home range is the total 'area traversed by an individual in its normal activities of food gathering, mating, and caring for young' (Burt 1943). Nice (1941) facilitated this distinction by providing a classification system that distinguishes territories used for mating, nesting, and feeding young (Type A) from territories used only for reproduction (Types B-D). Since then, tracking individual movements has revealed that many birds use areas outside the song-defended territory (reviewed by Whitaker and Warkentin 2010). Extraterritorial movements are made by many species, often for foraging, but also for extrapair mating opportunities, roosting, access to water, or to gain information on habitat quality in the surrounding area. These extraterritorial movements can result in home range sizes 1.4 to 4 times larger than Type A territories (Leonard et al. 2008; Anich et al. 2009; Streby et al. 2012; Jirinec et al. 2016; Bas et al. 2016; Frantz et al. 2016).

Although unstudied, there are likely consequences of territory and home range mismatch for the understanding of habitat requirements during the breeding season (Whitaker and Warkentin 2010, Streby et al. 2012). Vegetation characteristics can vary between territorial and extraterritorial areas; for example, both Wood Thrush and Sardinian Warblers use denser vegetation outside the territory for roosting (Bas et al. 2016, Jirinec et al. 2016). Failure to incorporate extraterritorial habitat in critical habitat designation, environmental impact assessments, or other habitat studies could lead to management actions that do not fully satisfy the resource requirements of the target species. Understanding the biological purpose of territorial and extraterritorial areas is thus required to understand those resource requirements and provide context to habitat studies.

The nightjars (Family *Caprimulgidae*) are a group for which there is likely a large mismatch in estimates between territory and home range. Most nightjar species have type B or C territories (Nice 1941), where they defend a small area for nesting and sometimes mating but conduct extraterritorial movements for foraging and roosting. Observation of "inconsistencies in expected territorial behavior" for this family date back to Lack's observation of breeding male European Nightjars roosting near each other outside of their territories (1932). Since then, radiotelemetry and GPS tracking have revealed that European Nightjars forage up to ~4 km from the nest, and habitat characteristics vary between the breeding territory and extraterritorial foraging areas (Evens et al. 2018, 2020, 2017). Similar patterns have also been detected for the

Red-necked Nightjar (*C. Camacho, pers. comm.*) and the Egyptian Nightjar (*Yohay Wasserlauf, pers. comm.*).

Among the nightjars, the common nighthawk (*Chordeiles minor*) is likely to have a high degree of mismatch between territory and home range. The common nighthawk is among the most mobile nightjar species; tracking data have revealed home range estimates hundreds of times larger than the breeding territory (*unpublished data*). This extreme mobility impedes the ability to track common nighthawks with anything but GPS tags, and so there is confusion about whether the species is territorial, and if so, what the territory size estimates are (Brigham et al. 2011). As a result, there is rarely a distinction made between territorial and extraterritorial detections in habitat studies (Newberry et al. 2018; Farrell et al. 2019; Viel et al. 2020), which is particularly problematic when results are interpreted in the context of nesting (Hagar et al. 2004). Accurate information about territoriality and individual spacing is critical for the common nighthawk; research and management at biologically appropriate scales is sorely needed for this poorly understood and declining species that is listed as of conservation concern in multiple jurisdictions (Environment Canada 2016a).

Fortunately, the common nighthawk makes two distinct sounds that could be used during research and monitoring to differentiate territorial from home range habitat use. The wing-boom is a mechanically produced acoustic signal that occurs when the primary feathers are flexed downwards at the bottom a steep aerial dive (Miller 1925). The resultant sound is a ‘vroom’, between 0 and 1 kHz, that can be heard at long-range due to its frequency (Figure 4.1). The wing-boom is thought to be a territorial signal, potentially associated with the nest (Gross 1940, Rust 1947; Ng 2009). In constrast, the call is a short (~0.3 second) ‘peent’ between approximately 2-5 kHz that has an effective detection radius of approximately 300 m (Figure

4.1; Chapter 3). This call is a more general-purpose signal, although it is most frequently produced near the nest location as well (Armstrong 1965; Wedgwood 1973; Caccamise 1974), in part because the wing-boom is always accompanied by the call (*unpublished data*).

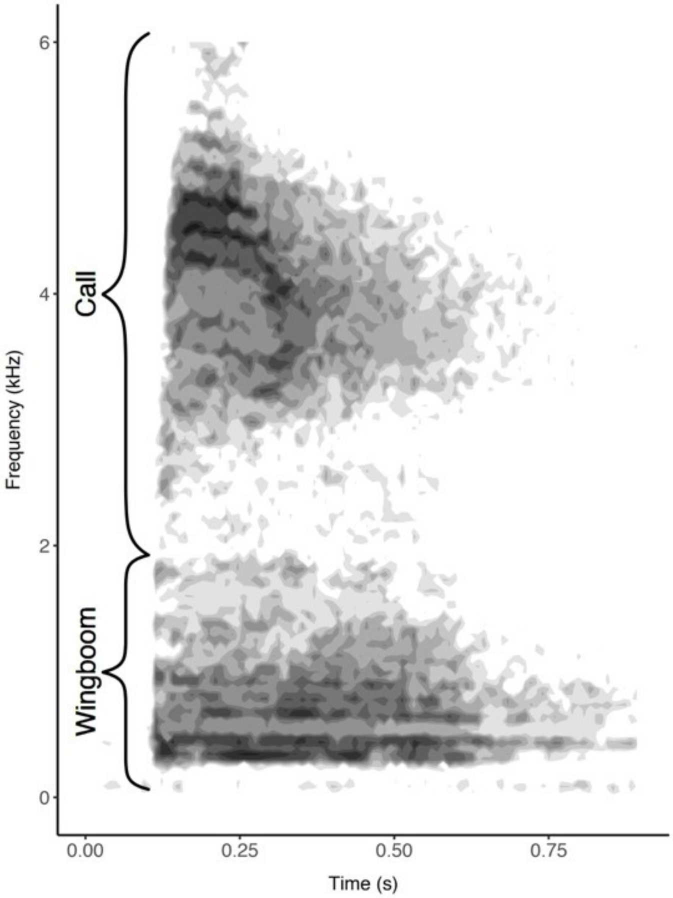


Figure 4.1. Spectrogram of the common nighthawk 'peent' call and wing-boom recorded at McLelland Lake, Alberta, Canada on July 13, 2017.

Our goal was to confirm the biological significance of the common nighthawk wing-boom display to inform and aid management of this species. Our first objective was to determine whether the mechanical wing-boom is analogous to song in that it is a signal that delineates an exclusive, discrete area within the home range. We used kernel density estimation to estimate and map the wing-boom use distribution (hereafter "area") of male VHF-tagged common nighthawks at five study sites in northeastern Alberta, Canada and quantify overlap in wing-boom area between adjacent territories. We also described inter-annual territory fidelity by quantifying overlap of wing-boom area between 2016 and 2017. Our second objective was to confirm that the territory is used for nesting. We used a resource selection function to determine whether male common nighthawks were selecting areas closer to the nest to perform wing-boom displays.

Methods

Study Area

We conducted our study in the boreal forest north of Fort McMurray, Alberta, Canada (Figure 4.2). The area is within the Athabasca Plain natural subregion, which is characterized by sandy and gravelly uplands of glaciofluvial, deltaic, and eolian origins (Natural Regions Committee 2006). The dominant upland vegetation was a mix of surviving and regenerating jack pine (*Pinus banksiana*) after the area burned in 2011. To the southwest lies McLelland Lake and a large wetland complex, and directly east is the Athabasca River. Immediately to the south is Alberta's oilsands region and, thus the area has substantial anthropogenic disturbance including retention clearcuts, transmission and water lines, roads, and open-pit bitumen mines. We chose the study area because the region has a large, dense population of breeding common nighthawks. Additionally, the anthropogenic disturbance and distribution of wetlands in the area facilitated radiotracking of our highly mobile study species.

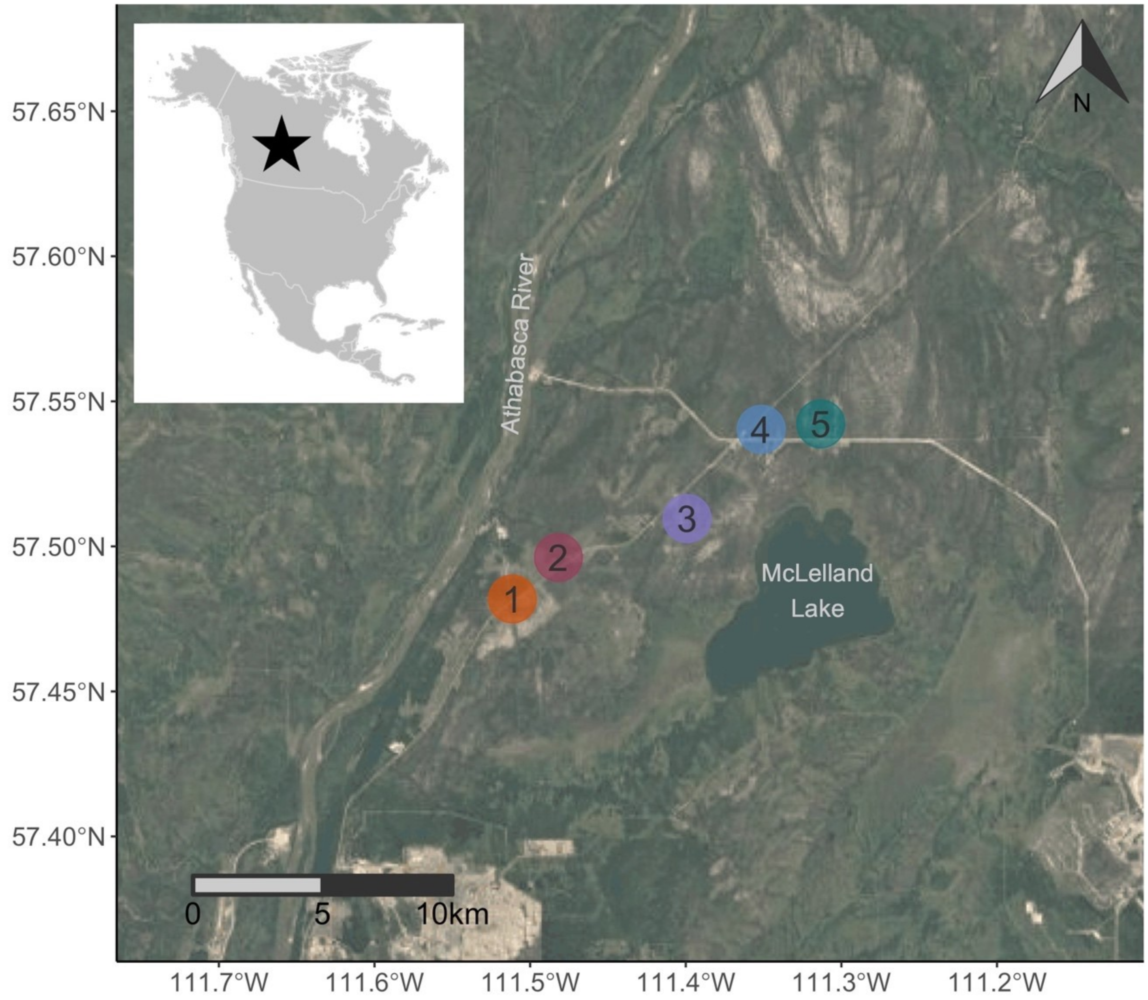


Figure 4.2. Study area north of Ft. McMurray, Alberta, Canada where common nighthawks at five study sites were tracked with VHF tags to study acoustic behavior and territoriality. Google earth imagery accessed via the package ggmap in R version 4.0.3 on May 3, 2021.

Individual Tracking & Observation

We captured and fitted male common nighthawks with VHF tags during the breeding seasons of 2016 and 2017. We used broadcast calls and a decoy to lure individuals into 38 mm mist nets. Individual males captured in 2016 were selected based on presence at our five study sites. In 2017, we attempted to recapture the same individuals from the previous year, as well as additional males at the same study sites. Upon capture, each individual was fitted with either a

0.6 g PicoPip Ag376, 1.1 g Pip Ag392 (Lotek wireless, Newmarket, ON, Canada), 0.35 g BD-2X, or 1.2 g BD-2 (Holohil Systems Ltd., Carp, ON, Canada) VHF transmitter (0.4 – 1.7% body mass [mean=1.0% ± 0.5% SD]). All tags were attached to fall off naturally: BD-2X VHF tags were glued to the bare patch of skin between the feather tracts on the back with super glue gel (Gorilla Glue Company, Cincinnati, Ohio); Pip Ag 392 and BD-2 VHF transmitters were tied and glued to one of the central rectrices. All individuals were banded with a uniquely-numbered metal leg band for individual identification between years. All work was conducted under University of Alberta Animal Care and Use Approval AUP00001523.

We relocated each individual throughout the breeding season and conducted observations of their acoustic behavior. We waited at least one day following capture to allow the birds to acclimate to the presence of the VHF tags. All observations were conducted between June 1 and July 20 when common nighthawks are nesting and between 21:30 and 03:30 when birds were active. Following relocation of the target individual, an observer conducted 10 minutes of focal observation of the bird, marking the location of each wing-boom display on a datasheet with distance measurements. Focal observations were terminated early if the target individual was confused with other individual nighthawks. In a few instances, we also terminated the observation if the focal individual moved out of the observation area (> 200 m); we assume those individuals were not leaving the area to wing-boom in other locations. Following each focal observation, the observer walked to the estimated location of each recorded wing-boom and used a GPS device to record the coordinates.

Nest Location

We attempted to locate the nest of each tracked nighthawk. Nightjar nests are cryptic and difficult to locate because incubating or brooding individuals will not flush until disturbed (Holyoak 2001). We thus relied on eyeshine to locate each nest; nightjar retinas have highly reflective *tapeta lucida* that reflects a yellow-orange luminescence. We searched each study site after dark by walking transects spaced approximately 100 m apart and sweeping back and forth with a high-powered headlamp. We assumed that the identity of the male for each nest was the male performing wing-boom displays around the nest area. On several occasions, we directly observed the male provisioning the chicks at that nest or interacting with the female.

Statistical Analysis

Kernel density estimation: We used kernel density estimation (KDE) to determine the wing-boom area of each tracked nighthawk. We considered all observations as biologically independent because there was more time between them than the interval required to reach any point within the territory (approximately 10 seconds) (Barg et al. 2005). There was variation in the number of wing-boom locations available for each individual (5-110 points), so we tested for an effect of sample size on wing-boom area. We randomly sampled between 5 (the minimum number required to calculate KDE) and 110 (the maximum number available) wing-boom locations for each individual and estimated the wing-boom area from those randomly sampled points. We used the ad hoc method in the adehabitatHR package (Calenge 2006) in R version 4.0.3 (R Core Team 2020) to estimate the smoothing parameter of the bivariate normal kernel with a grid size of 1000 and an extent of 2. We bootstrapped this process 100 times, then fit a nonlinear least squares growth curve to the results, with the area of the 100% isopleth (quantile of kernel estimate) as the response and the number of wing-boom locations as the predictor. We fit this growth curve separately for each individual bird and selected only birds for which the growth curve estimate reached 90% of the model asymptote (Supplementary Materials). In other words, we used only birds for whom the relationship between sample size and home range area appeared to asymptote and removed birds for whom our sample size was inadequate to accurately predict home range size. This approach resulted in a sample size threshold of 30 wing-boom locations, which is also the recommended minimum sample size for kernel density estimation (Seaman et al. 1999). We removed six individuals for which there were fewer than 30 wing-boom locations in a given year (range: 5-23 locations). We then used KDE and the same settings as before to estimate the 29 remaining wing-boom areas (i.e., individual-year combinations).

Objective 1: Confirm territoriality: We determined whether wing-booms were a territorial signal by testing whether there was overlap in the 95% isopleth of the wing-boom area between adjacent male common nighthawks. There were 13 pairs of immediately adjacent neighbors amongst our VHF-tagged focal individuals. In other words, we evaluated whether the wing-boom area was exclusive to a particular male. For each neighbor pair, we used the adehabitatHR package to calculate the mean probability of home range overlap, which takes into account the relative probability of use of the area of overlap (i.e., three-dimensional overlap; hereafter “overlap”) (Fieberg & O’Kochanny 2005). We used the adehabitatHR package and a grid size of 2000 to calculate the volume of each individual’s wing-boom area that was contained within the wing-boom area of its neighbor, and then calculated the mean overlap for each neighbor pair.

We also examined the degree of interannual territory fidelity using the ten males for which we had two years of sufficient wing-boom locations to estimate wing-boom area. We again calculated mean overlap of the 95% wing-boom area between years. We tested for territory fidelity using a one-sample t-test. Although our hypothesis was directional because overlap cannot be less than 0, we used a two-sided test to ensure our conclusions were conservative. We also tested for differences in the area of the 95% isopleth between years using a paired two-tailed t-test. We calculated the distance between nests in successive years for the 5 males with known nest locations in both years. Of those five males, one in 2016 renested after the first nest failed, so we calculated the distance to his 2017 nest from both 2016 nests.

Objective 2: Confirm nest defense: We used a resource selection function (RSF; Manly 2002) to determine whether common nighthawk wing-booms were associated with the nest location. In other words, we tested whether individual male common nighthawks were selecting areas closer to the nest than available to perform their wing-boom displays. We defined the domain of availability as the 95% isopleth of the wing-boom locations. This definition is a highly conservative estimate of home range, as common nighthawks in the study area can have 40 km^2 home ranges (*unpublished data*). We then created 100 random points within the 95% isopleth for each individual (available points). We measured the distance to the nest for each available point and each wing-boom location (used points). We modelled the relative probability of wing-boom selection using mixed-effects logistic regression with a logit link in the lme4 package (Bates et al. 2015). We included distance to nest, the area of the 95% isopleth, and the interaction between the two as predictors. Individual ID was fit as a random intercept. We compared this model to a model with only distance to nest as a predictor, and to a null model using AICc. We centered and scaled both predictors prior to model fitting. We used the merTools package (Knowles and

Frederick 2020) to generate mean predictions and 95% confidence intervals of wing-boom relative selection probability across a range of distances to nest and 95% isopleth areas by drawing 1000 sampling distributions for the random and fixed effects, estimating the fitted value across each distribution, and taking the mean and 95% quantile of those 1000 fitted values.

Results

We captured and tracked 21 male common nighthawks in 2016 and 2017 and estimated the wing-boom use density for 27 bird-year combinations after removing samples with fewer than 30 wing-boom locations (n=6). Fourteen of those wing-boom areas were estimated for males captured and tracked in 2016 and 13 were for males capture and tracked in 2017. Ten of those individuals that were caught and tracked in both years; no effects of tags applied the previous were noted on recaptured individuals. The number of wing-boom locations for each bird ranged from 31 to 110 (mean=60.3, SD=19.7). We found nests for 21 of the 27 bird-year combinations.

Territory Size

The mean 95% isopleth of the wing-boom area was 10.2 ha (SD=11.7 ha, min=1.1 ha, max=54.9 ha.; Figure 4.3). The largest wing-boom area was nearly twice the area of the next largest (45.9 ha vs. 26.8 ha) and was for an individual tracked in both years who appeared to be unpaired in 2017. After removal of that wing-boom area, the mean 95% isopleth of the wing-boom area was 8.5 ha (SD=7.7 ha, max=26.8 ha).

Territory Overlap

Although males often traversed each other’s territories, the overlap in the 95% wing-boom area between neighbouring males was minimal (Figure 4.3). We tracked individuals in 13 pairs of directly adjacent territories and found five instances of territory overlap. The amount of overlap in four of those instances was minimal, ranging from 0.2% (SD=0.1%) to 4.5% (SD=1.9%) overlap. The remaining pair of territories had a mean overlap of 57.2% (SD=48.9%) and included one bird whose 2016 territory boundaries shifted north after his first nest failed and his female renested further north (Figure 4.3).

Territory Fidelity

There was significant overlap in the mean 95% wing-boom area between years (t_9=17.24, *P*<0.001; Figure 4.4). We tracked 10 individuals in both years and found overlap in individual wing-boom areas for all ten. The mean overlap between years was 68.4% (SD=12.6%) and ranged from 38.3% to 82.7%. We found nests in both years for five of those males. The mean distance between nests in successive years was 101 m (SD=78 m) and ranged from 22 m to 209 m (Figure 4.4). One male at site 5 had two nests in 2016 after the first nest failed; his 2017 nest was 90 m from the first nest and 209 m from the renest (Figure 4.3). There was no significant difference in wing-boom area between years (t_9 = -1.45, *P* = 0.18).

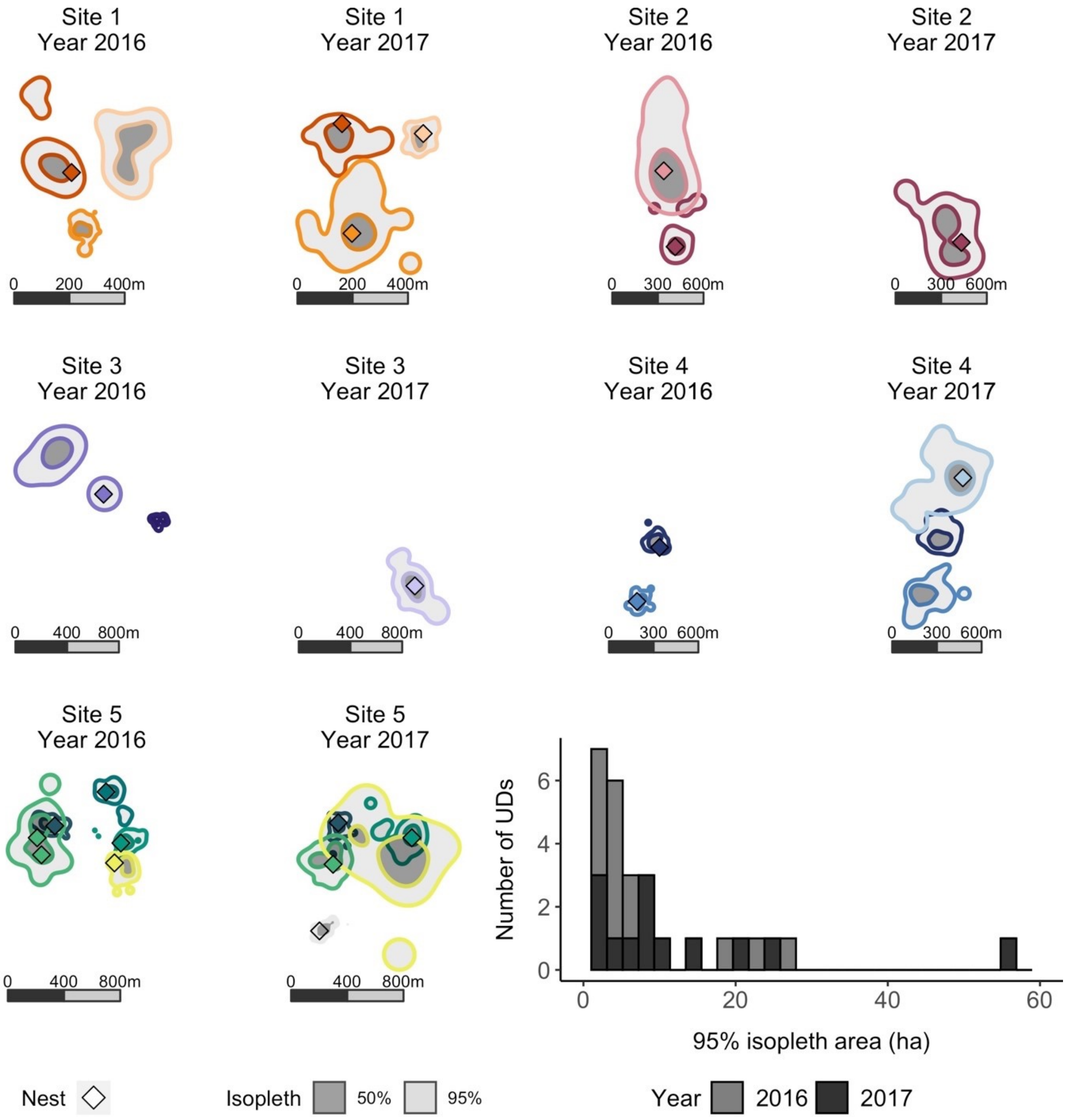


Figure 4.3. 95% and 50% isopleths from kernel density estimation of wing-boom display locations for male common nighthawks breeding in northeastern Alberta, Canada in 2016 and 2017. Individuals were marked with VHF tags and relocated to observe and record the location of their wing-boom displays. Nest sites are plotted when the location was known. Note the two nest locations for one bird at Site 5 in 2016.

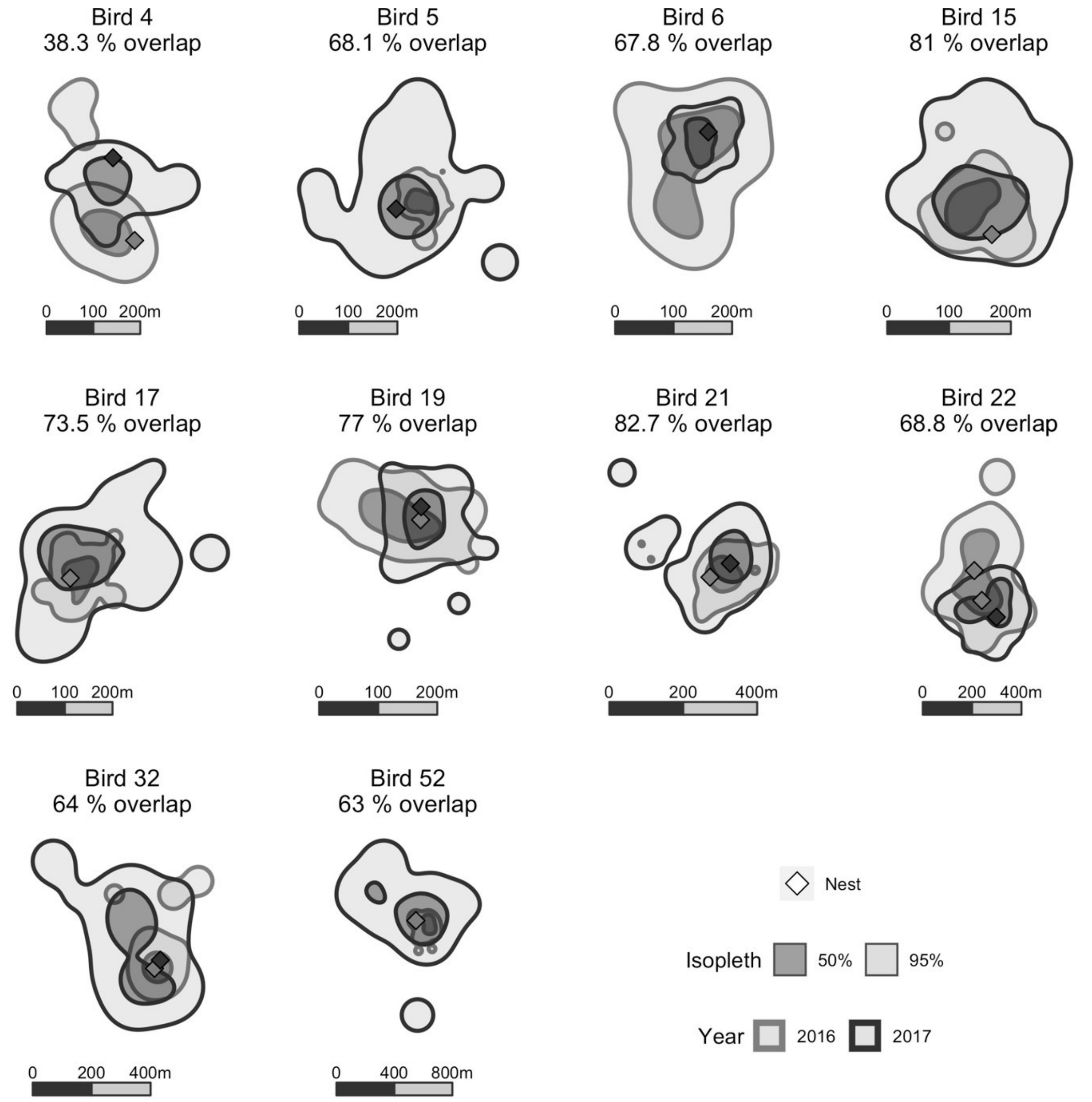


Figure 4.4. Between-year overlap of 95% and 50% isopleths from kernel density estimation of wing-boom display locations for male common nighthawks breeding in northeastern Alberta, Canada in 2016 and 2017. Individuals were marked with VHF tags and relocated to observe and record the location of their wing-boom displays. Nest sites are plotted where nest location was known.

Wing-boom Site Selection

The RSF with distance to nest, KDE area, and the interaction between the two was much more predictive than the null model (distance RSF: ΔAICc=80.4; null RSF: ΔAICc=301.9), suggesting that common nighthawks select areas near the nest to perform wing-boom displays (Figure 4.5). The relative probability of a wing-boom display within the territory declined with distance from the nest. The distance at which the relative probability of wing-boom display approached zero depended on the size of the 95% isopleth.

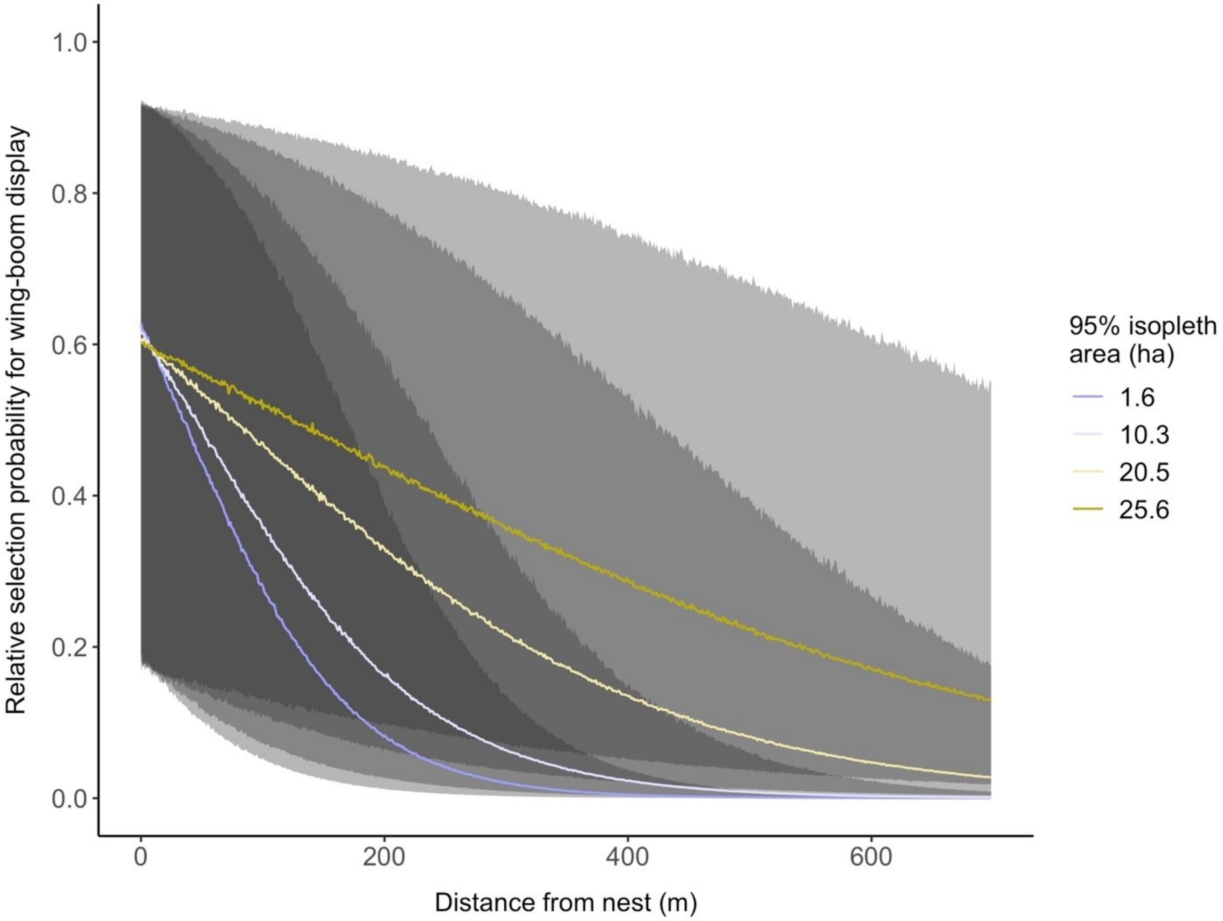


Figure 4.5. Mean and 95% confidence interval of relative selection probability for where common nighthawks perform wing-boom displays relative to the nest location. Individuals were marked with VHF tags and relocated to observe and record the location of their wing-boom displays. Predictions are the mean and 95% quantile of 1000 fitted values from simulated fixed and random effects.

Discussion

We used focal observations of VHF-tagged male common nighthawks to evaluate territoriality by this highly mobile species of conservation concern. There was minimal overlap of the 95% isopleth of wing-boom use density (hereafter “area”) between adjacent neighboring common nighthawks. Resource selection confirmed that the wing-boom signal is associated with the nest location, as the relative probability of a wing-boom display within the territory declined with distance from the nest. For individuals that were tracked in both years of the study, there was high wing-boom area overlap between years, suggesting high interannual territory fidelity. Furthermore, slight differences in the wing-boom area between years as nest location shifted suggested that the wing-boom display is a good indicator of the nest location and territory.

Territoriality has three common conceptual definitions: 1) a defended area, 2) an exclusive area, and 3) site-specific dominance (Maher & Lott 1995). We confirmed the common nighthawk wing-boom signal is territorial signal using the exclusive use definition. Although we do not report behavioral data to confirm defense and site-specific dominance in this study, we observed many instances of wing-boom displays directed at other males and territorial intruders, as have other authors (Armstrong 1965; Caccamise 1974). A complementary behavioral study of aggressive signal should be conducted to confirm common nighthawk wing-boom territories as defended areas with site-specific dominance (Searcy & Beecher 2009). Territoriality can be described as a multidimensional gradient, one axis of which is the resources being defended. Nice’s (1941) territorial classes follow this resource gradient, from Type A territories that are used for foraging, mating, and nesting resources, to Type D territories that are used exclusively as a nest site. We showed that the wing-boom display is specifically associated with the nest location, confirming that common nighthawk territories are used for nesting. Some songbird species with Type A territories also sing more frequently near the nest location (Simpson 1985); however, common nighthawks are well known to have overlapping home ranges (Armstrong 1965; Ng 2009), roost within each other’s home ranges (Ng 2009), and forage communally outside the territory (Brigham & Fenton 1991; Aldridge & Brigham 1991; Brigham et al. 1992; Brigham 1990). These behaviors are congruent with evolutionary expectations because if territoriality is adaptive, it should optimize the trade-off between resource defensibility and limited resource abundance (Peiman & Robinson 2010). Common nighthawk change roosts frequently and roosts vary widely in characteristics; thus, roosts are unlikely to be a limited resource in our postfire study area (Fisher et al. 2004). Aerial insects in the boreal forest are an abundant but difficult to defend resource. In fact, the majority of aerial insectivorous birds do not have Type A territories, which may be in part due to their reliance on a food source that is difficult to defend and their "exceptional powers of flight" (Nice 1941).

Armstrong (1965) and Caccamise (1974) arrived at different conclusions about common nighthawk territoriality; they both reported little excursion outside the territory, and Armstrong specifically concluded common nighthawks nesting in downtown Chicago held type A territories. Varying conclusions about the territoriality of the common nighthawk may be due to plasticity in territorial behavior associated with population density, particularly if the resource being defended is a female mate. Our study area provided breeding grounds for a particularly dense population of common nighthawks at the time of study, and so may differ from other areas. Territoriality typically increases with decreasing population density because the cost of defensibility decreases with fewer rivals; however, this relationship has been reported less often for birds (Maher & Lott 2000), perhaps because the probability of extra-pair copulation increases with density (Westneat & Sherman 1997; Petrie & Kempenaers 1998). Alternatively, varied conclusions about common nighthawk territoriality may be driven by the same mechanism as the omission of extraterritorial habitat in passerine studies: the inability to track a highly mobile and cryptic species outside the territory where it is less conspicuous. Neither Armstrong nor

Caccamise (1965, 1974) had the technology to confirm individual identities or to track individuals when they are not vocalizing on territory. Furthermore, Armstrong reported overlapping home ranges and unidentified nighthawks flying in and out of his study area, suggestive of type B, rather than type A territories. A comparative study of territoriality across the range of common nighthawks using standardized methods is required to confirm whether the species exhibits territorial plasticity.

What remains unclear, however, is whether the resource being defended by male common nighthawks is the nesting habitat, the nest contents, or the female. We suggest nesting habitat is unlikely, at least in this population, because bare sand is not a limited resource within the study area; however, nest site selection research is needed to formally reject this hypothesis. Territoriality in other populations may be more linked to nesting habitat availability; for example, Gross (1940) concluded that nesting sites were the driver of territoriality due to reports of individuals nesting within a few meters of each other. Secondly, defence of nest contents (i.e., eggs or nestlings) against conspecifics is also unlikely because infanticide has not been documented in caprimulgids. Thirdly, mate defense is common across bird species because extra-pair copulation occurs in the majority of species (Petrie & Kempenaers 1998; Griffith et al. 2002). Given the extreme mobility of the common nighthawk and the repercussions of extra-pair copulation for a species that generally fledges two young per year, we suggest the female is the most likely resource being defended. Our conclusion is supported by records of males performing wing-boom displays at females when they are not on the nest (Bowles 1921, Sutherland 1963). The wing-boom signal may also communicate information to the female like food delivery (Halkin 1997) or safety to forage (Boucaud et al. 2017) as common nighthawk parental care is shared between the male and female (Brigham et al. 2011). Future research should investigate the role of the female in common nighthawk territoriality, including nest site selection and density dependence.

Understanding territoriality and regional variation of it is important for management of common nighthawks because it determines whether this species is using separate or overlapping areas for different resource needs. Furthermore, type B territories will have a much greater degree of mismatch than type A territories and thus likely have larger differences in habitat characteristics between territorial and extraterritorial habitat use. We recommend future common nighthawk monitoring and research differentiate between the call and the wing-boom display during breeding surveys (Knight et al. 2019a) to facilitate differentiating territorial and extraterritorial habitat use. Wing-boom surveys can potentially also facilitate occupancy and density estimation of this species. Home range sizes are much too large to satisfy the 'closure' assumption of occupancy models (MacKenzie et al. 2002); however, territory size estimates here (mean 10.2 ha) and elsewhere (10.4 ha, Armstrong 1965; 10.5 ha, Wedgwood 1973) suggest this assumption may be valid for the time of day when male common nighthawks are actively defending territories. Existing occupancy estimates derived from detections of the 'peent' call should instead be interpreted as probability of habitat use (MacKenzie & Royle 2005). Although the call is also thought to have a territorial purpose, it is also emitted outside the home range and during other activities (*unpublished data*). Finally, the wing-boom signal can be used to improve the effectiveness of environmental impact assessment and environmental monitoring by focusing on the presence of common nighthawks during nesting. We caution that common nighthawk nests are mobile after the semi-precocial chicks hatch (Kramer & Chalfoun 2012) and can move up to 48 m from the incubation location. Our study included wing-boom locations from before and after nest hatching, but we did not investigate potential differences between the two nesting periods. We also emphasize the importance of restricting inferences from the wing-boom signal to the breeding season and grounds; however, because there are anecdotal reports of wing-booms produced on migration. Acoustic behavior of common nighthawks on the wintering grounds is unstudied but reported as primarily silent (*K. Cockle pers. comm.*).

In summary, understanding the biological significance of auditory cues is an important prerequisite for using those cues to inform wildlife management, particularly when there is a mismatch between territorial and home range habitat use. Nightjar species are excellent candidates for using auditory cues to differentiate between territorial and home range habitat use because acoustic communication is particularly important to these crepuscular birds. Many nightjars are silent during most of the annual cycle but begin to deliver calls from prominent vantage points at the beginning of the breeding season, indicating territorial significance (Holyoak 2001; Cleere 2010). Using those cues to differentiate between territorial and extraterritorial habitat use is likely important because all known tracking studies have revealed extensive extraterritorial foraging and roosting movements (Evens et al. 2017, Events et al. 2018, Evens et al. 2020, *C. Camacho pers. comm., Y. Wasserlauf pers. comm.*). Links found between habitat loss and demographics of other aerial insectivorous bird species (Spiller & Dettmers 2019) further emphasize the importance of correctly understanding habitat associations for this guild, which is declining faster than any other group of birds in Canada (Canada 2019). Failure to understand how resource requirements are separated spatially across the home range, particularly when using auditory cues to model habitat, could lead to inappropriate management actions.

Supplementary Materials

Sample size threshold for kernel density estimation

Methods: We used kernel density estimation (KDE) to determine the wing-boom use distribution (hereafter "area") of each tracked nighthawk. There was variation in the number of wing-boom locations available for each individual (5-110 points), so we tested for an effect of sample size on size of wing-boom area. We randomly sampled between 5 (the minimum number required to calculate KDE) and 110 (the maximum number available) wing-boom locations for each individual and estimated the wing-boom area from those randomly sampled points. We bootstrapped this process 100 times, then fit a nonlinear least squares growth curve to the results, with the area of the 100% isopleth (quantile of kernel estimate) as the response and the number of wing-boom locations as the predictor. We fit this growth curve separately for each individual bird and selected only birds for which the growth curve estimate reached 90% of the model asymptote. In other words, we used only birds for whom the relationship between sample size and home range area appeared to asymptote and removed birds for whom our sample size was inadequate to accurately predict home range size.

Results: We used a sample size threshold of 30 wing-boom locations, which is also the recommended minimum sample size for kernel density estimation (Seaman et al. 1999). We removed six individuals (2016: Birds 24, 29, 30, 76; 2017: Birds 36, 91) for which there were fewer than 30 wing-boom locations in a given year (range: 5-23 locations; Figure 4.A.1).

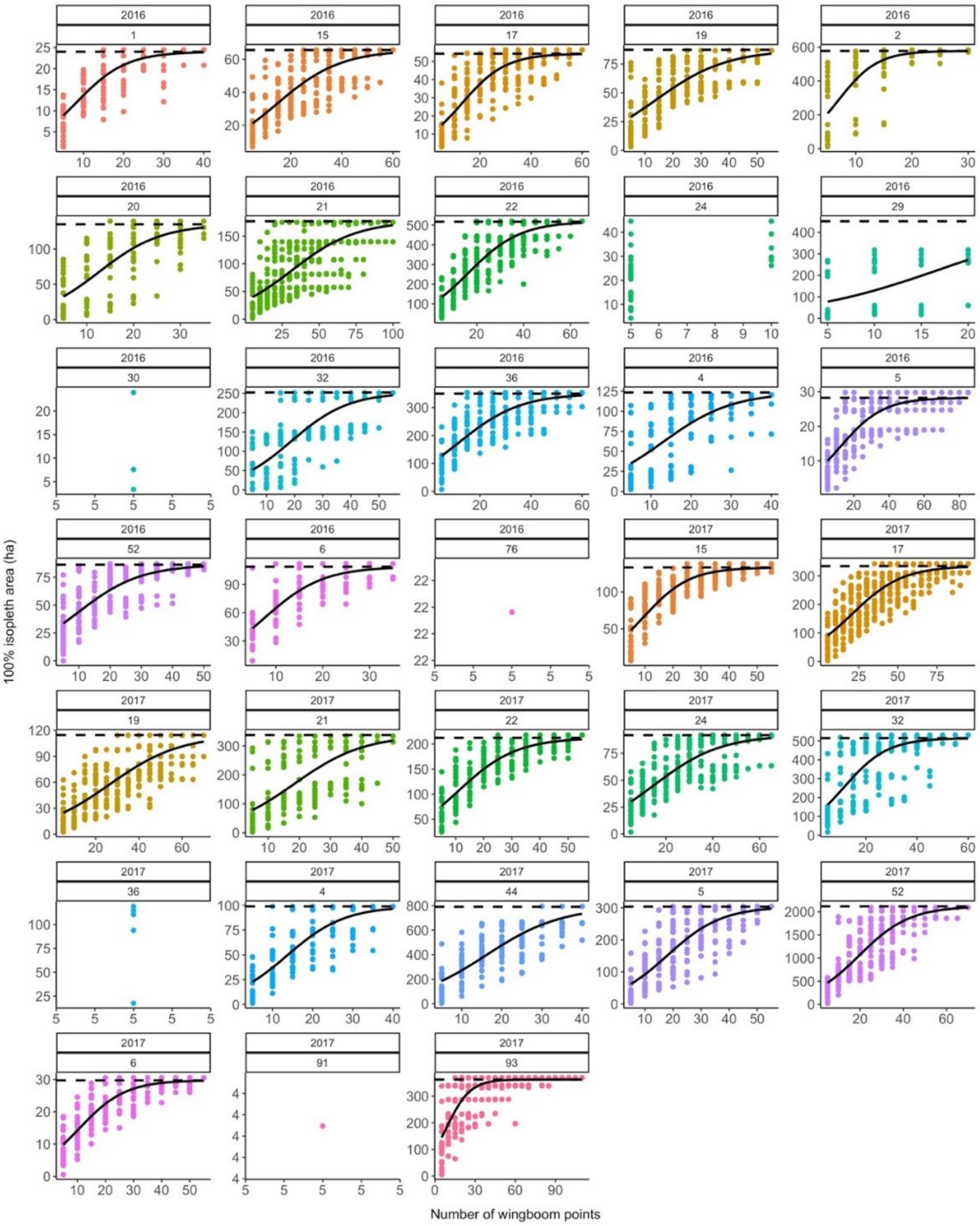


Figure 4.A.1. Growth curves of sample size and area of the 100% isopleth of the wing-boom use distribution of adult male common nighthawks from kernel density estimation. Samples were randomly selected from the available points for each individual in each year (2016, 2017), each sample size was bootstrapped 100 times, and a nonlinear least squares growth curve was fit to the results. Dashed lines represent the asymptote of 100% isopleth area for each individual.

Chapter 5. Movement range corresponds to scale of effect for single scale models but not for individual variables within models

Abstract

Movement is one of the proposed mechanisms for the scale at which a species responds most strongly to its environment, or the "scale of effect. Scarcity of empirical evidence for this hypothesis may be because studies determine scale of effect for individual environmental variables; however, seasonal movement is the complex product of reactions to multiple variables. We predicted scale of effect should correspond to movement range for the most predictive single scale habitat model ("overall scale of effect"), but not for individual predictors. We used passive acoustic monitoring and machine learning to model territorial and home range habitat for the common nighthawk. We modeled extents from 0.1 to 12.8 km to determine the overall scale of effect. We used the scale of effect for each predictor to build optimized multiscale models and evaluated their spatial predictive performance. The overall scale of effect was 0.2 km for territorial habitat and 1.6 or 6.4 km for home range habitat, which roughly equate to territory and home range size. The scale of effect for the strongest individual predictors did not correspond to overall scale of effect. Optimized, multiscale models offered no substantial improvements in predictive performance relative to overall scale of effect models. Our new perspective on scale of effect suggests that different mechanisms drive overall scale of effect and scale of effect of individual variables. Further research should revisit the relationship between movement and scale of effect in pursuit of a mechanistic framework for prediction.

Introduction

The effect of the surrounding landscape on species occurrence has long been understood to vary with scale (Wiens 1989; Levin 1992). In particular, the scale at which environmental variables are measured can influence the strength and even direction of the effect of those variables on species occurrence (Holland et al. 2004; Smith et al. 2011). Understanding the scale at which a species responds most strongly to the surrounding landscape for a particular response variable, or the 'scale of effect' (sensu Jackson and Fahrig 2012), is thus important for understanding habitat requirements, testing ecological predictions, and ultimately, guiding land use and wildlife management actions (Thornton & Jr 2013; Miguet et al. 2016).

While the importance of scale of effect is well understood, the mechanisms that drive this phenomenon have yet to be generalized. One of the most intuitive and thus common hypotheses is that scale of effect is driven by movement traits of a species because mobility determines the scale at which that species interacts with the surrounding landscape (Miguet et al. 2016). More mobile species should have larger scales of effect because they interact with environmental variables at a larger spatial extent. For small species like songbirds that conduct daily movements within a small breeding territory, scale of effect is predicted to be driven by dispersal movements (Tittler 2008). In contrast, scale of effect is predicted to be driven by seasonal home range movements for highly mobile species that travel within a larger home range for various resource needs (e.g., foraging, roosting, extra-pair copulation; Tittler 2008). A variety of predictions can be derived from the potential relationship between movement and scale of effect, including larger scales of effect for larger-bodied species, flying (i.e., more mobile) species, species at higher trophic levels, and migratory species (Miguet et al. 2016).

Despite the logical theoretical link between movement and scale of effect, the current consensus is that supporting evidence is limited (Miguet et al. 2016). There is simulation support for a link between dispersal range and scale of effect (Jackson & Fahrig 2012; Ricci et al. 2013); however, the linkage is difficult to evaluate empirically, and the one existing study with data for 22 bird species is not supportive (Tittler 2008). No evidence of a link between home range size and scale of effect has been found across nine studies (eight bird, one mammal; reviewed by Jackson and Fahrig 2015). This lack of support may be because scale of effect is rarely measured across a wide enough range of scales; many studies report scales of effect equal to the smallest or largest scale evaluated (Jackson and Fahrig 2015). Alternatively, there may be other mechanisms driving scale of effect including population density, reproductive rate, predator avoidance, and or competition, to name a few (Miguet et al. 2016).

We suggest that the lack of empirical support for a link between home range size and scale of effect for highly mobile species is because existing studies have focused on optimizing scale of effect for individual environmental variables; however, seasonal movement within a home range is the complex product of reactions to multiple environmental variables. The scale of effect of a single environmental variable is more likely to be driven by movements to fulfill specific resource requirements (Miguet et al. 2016), which typically do not encompass the entire home range. For example, the scale of effect for open water may be quite small relative to the home range for a species that travels short distances daily to drink during the course of larger and longer-term movements within the home range. In contrast, the scale of effect for the combination of environmental variables that a species is influenced by is more likely to correspond to home range size because the home range is the sum of all movement conducted to seek out the individual resource requirements those environmental variables reflect.

We predict that seasonal movement scale should correspond to the scale at which a single scale habitat model explains the most variation in species occurrence data (hereafter “overall scale of effect”). Comparing overall scales of effect between behaviours with different movement extents of habitat use would provide insight into the link between movement and scale of effect; however, this type of comparison has yet to be conducted (Miguet et al. 2016). In other words, is the overall scale of effect for a territorial habitat model smaller than the overall scale of effect for a home range habitat model? We used this behavioural comparison approach to test our hypothesis for a long-distance migratory bird, the common nighthawk (*Chordeiles minor*). The acoustic behaviour of the common nighthawk presents an opportunity to test whether overall scale of effect is linked to movement because this species has two behaviours with different movement extents that can be distinguished by acoustic signals. It uses a mechanical wing-boom display to defend a small territory (~10 ha; Chapter 4) for mating and nesting but vocalizes across a much larger home range for foraging and roosting (> 40 km^2; *unpublished data*). Common nighthawk habitat use can thus be studied across large spatial scales using passive acoustic monitoring.

We explicitly tested the hypothesis that movement range corresponds to scale of effect of single scale habitat models, but not for individual predictors within the models. To test our hypothesis, we built a dataset of common nighthawk detections by using signal recognition software to process a large dataset of acoustic recordings from Canada’s boreal forest. We used a standard multiscale approach to determine scale of effect, where we summarized our environmental variables of interest at multiple, nested buffers surrounding each survey location (i.e., extents; hereafter “scale”; summarized by Miguet et al. 2016). We included all environmental variables that we thought could affect habitat use in a single scale model and defined overall scale of effect as the single scale model that explained the most variation in the data. We also partitioned deviance for each individual variable in each model and determine the scale of effect for each variable. We predicted that the overall scale of effect for territorial habitat use would be smaller than that for home range habitat use, but that this relationship would not necessarily hold for individual environmental variables. Finally, we tested whether we could improve the predictive performance of our model by building a final multiscale model that contained the scale of effect for each individual environmental variable. We create spatial predictions for territorial and home range habitat use from the final models to facilitate management of this species in Canada's boreal forest.

Methods

Study Area

We defined our study area as an approximately 75,000 km^2 area in north-eastern Alberta (Figure 5.1). We selected the study area based on available sampling locations and extent of available environmental data (Supplementary Materials). The study area was primarily within the central mixedwood natural subregion (Natural Regions Committee 2006). The landscape was characterized by a mosaic of upland forest types and lowland peat bogs and fens. At the time of study, the region was subject to frequent and widespread disturbance, including active forestry, oil and gas development, and wildfire.

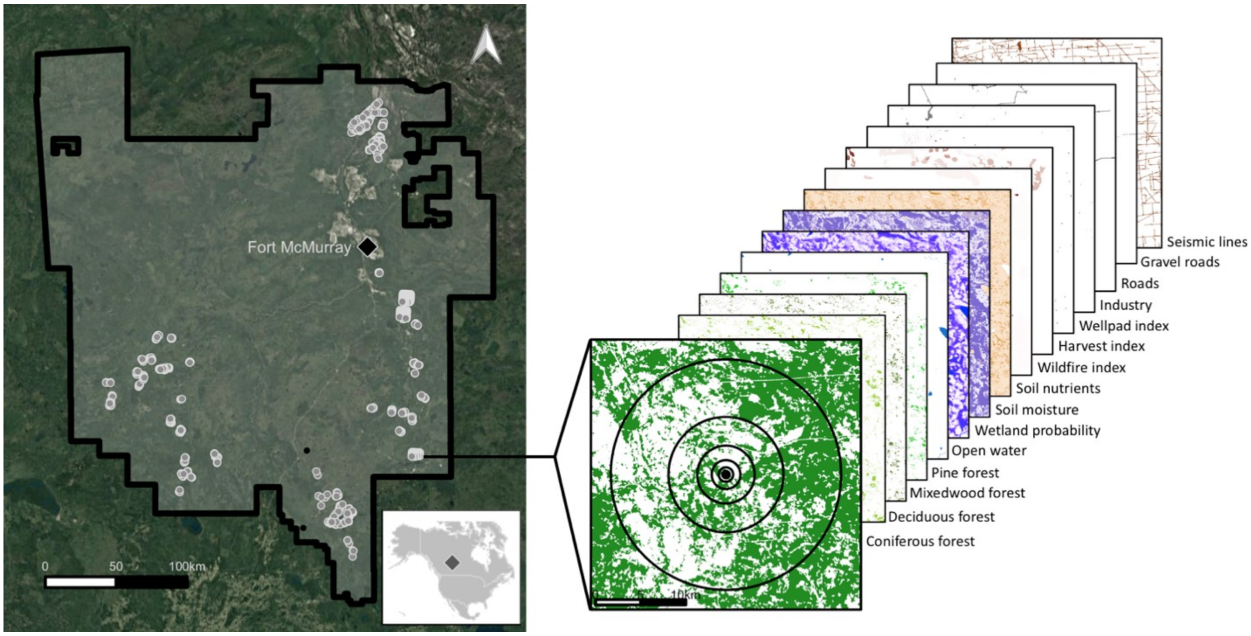


Figure 5.1. Study area, recording stations (dots), environmental variables, and extent of buffers (0.2 – 12.8 km) used to study scale of effect of common nighthawk territorial and home range habitat use.

Acoustic Recording Collection

We extracted common nighthawk detections from an archived bioacoustic database (https://www.wildtrax.ca/) of recordings collected with autonomous recording units (ARUs; SM2, SM2+, and SM3; Wildlife Acoustics Inc.). We selected 827 recording stations (Figure 5.1) that were sampled by ARUs following a standardized protocol (Lankau 2015) between June 1 and July 31, 2015. The 827 stations were part of six different bioacoustic projects with varying recording schedules, but all recordings were 10 minutes long. All recordings were made with two omnidirectional microphones at a sampling rate of 44.1 kHz with a 16-bit depth.

Acoustic Recording Processing

We used Song Scope (Wildlife Acoustics 2011) software to extract common nighthawk detections from the ARU recordings at each station (6,216 recordings in total). We created a recognizer by training Song Scope with 138 unmasked clips of common nighthawk calls recorded at close proximity in north-eastern Alberta and south-central British Columbia. We used the recognizer to scan the recordings using a minimum score threshold of 70 and a minimum quality threshold of 30, which we have previously shown optimizes the trade-off between false positive and false negative detections, while limiting the effective detection radius of the recognizer to approximately 150 m (Chapter 2, Chapter 3). The first author visually validated all potential common nighthawk detections to confirm whether they were true or false positive detections. During validation, the first author also noted all detections that were accompanied by a mechanical wing-boom display. Common nighthawks vocalize every time they perform a wing-boom display (*unpublished data*).

Environmental Data Processing

We extracted 15 environmental variables for each scale (Figure 5.1, Supplementary Materials). The environmental variables chosen were general dominant vegetation classes known to drive boreal bird community (Cumming et al. 2013), soil property variables that may influence common nighthawk ground nest site selection, and disturbance variables that may influence common nighthawk habitat use (Chapter 6). First, we resampled all data sources to 30 m raster layers to standardize the grain of our variables and facilitate further processing. We then reclassified each of the raster layers to extract the classes of interest. For three of the disturbance variables (fire, harvest, oil well site), we weighted each cell from 0 to 1 by dividing by time since disturbance. For the two soil property variables, we converted the classes to ordinal integers and standardized them as 0 to 1. The result was 15 raster layers of identical grain (30 m) and extent (Figure 5.1) and cell values ranging from 0 to 1.

Next, we calculated the mean value for each environmental variable within eight spatial scales for each station. We chose scales ranging from the smallest known territory radius (100 m; Chapter 4) to the largest known home range radius (~12 km) for common nighthawks (*unpublished data*). We decided to select scales of doubling value to avoid strong correlation between subsequent scales. The selected scales were thus 0.1, 0.2, 0.4, 0.8, 1.6, 3.2, 6.4, and 12.8 km (Figure 5.1). We extracted the mean value for each of our 15 environmental variables by using a moving window to create a new raster for each variable at each scale (120 raster layers in total), with the value of each cell calculated as the mean value within the specified scale. We then extracted the cell value for each recording station from each of the 120 raster layers. We used this raster-based approach instead of using buffers around each recording station to calculate the mean so that we could use the raster to create spatial predictions of habitat use later. All environmental data processing was done in ArcGIS version 10.5 (raster conversion & resampling only; Esri Inc 2016) and in R version 3.5.2 using the packages raster, rgdal, and rgeos (Hijmans 2020, Bivand and Rundel 2020, Bivand et al. 2021).

Statistical Analysis

Modelling Approach: We used boosted regression trees to build habitat use models for territorial and home range habitat use. Boosted regression trees (BRTs) are a machine learning approach that maximizes the predictive performance of the model by combining regression trees with boosting (Elith et al. 2008). The regression trees fit the explanatory variables to the response variable by recursively splitting the data into homogenous rectangular groups and fitting a constant for each explanatory variable to each group (De'Ath & Fabricius 2000). The boosting improves the predictive performance of the regression trees by iteratively fitting regression trees in a forward stagewise process to reduce deviance (Elith et al. 2008). We choose

BRTs to model habitat use because they can accommodate a large number of potentially correlated predictors and are robust to overprediction, which is necessary when building an overall habitat model to maximize predictive capacity. Further, boosted regression trees can model non-linear reactions and incorporate complex interactions between predictors, which could contribute substantially to model performance, and thus to correctly identifying the overall scale of effect.

Spatial Thinning: The sampling effort of the acoustic recordings used was strongly clustered in several areas, which can cause habitat model bias towards those clustered areas (Ploton et al. 2020; Robinson et al. 2017). We therefore used a grid sampling approach to spatially thin the dataset before habitat modelling. We divided the study area into 1 km cells, randomly picked one recording station from each cell, and retained data from the remaining stations for validation of spatial predictions (see 'Spatial Predictions' below).

Survey Effort: We standardized survey effort at each recording station by sampling a set number of recording minutes from each station from times and days when common nighthawk availability for detection was maximized. We used survival analysis in the package survival (Therneau and Grambsch 2020) to estimate the probability of common nighthawk territorial activity. In other words, for any given recording, what is the probability that a nighthawk is calling, if it is present at the study site? We fit a parametric survival regression model with time to first detection in each of our 6,216 recordings as the response variable (Sòlymos et al. 2018, 2013). For recordings in which a common nighthawk was not detected, we treated non-detections as censored events with a time of detection of 10 minutes (i.e., recording length). We included day of year, a quadratic effect of day of year, the sin of time of day, and the cos of time of day as predictor variables. We included time of day as trigonometric functions of time of day to allow for circularity. We built a global model with all predictor variables and compared it to all potential simpler models using AICc (Supplementary Materials). We selected the global model as the best fitting model. The probability of availability (p_i) for a 10-minutes recording is given as:

$$p_i = 1 - e^{-10a_i}$$

where a_i is the event rate per minute.

We then restricted the available data for each station to only the recordings that were collected at a day of year and time of day with a 99% probability of detection (Figure 5.2). From the remaining available recordings, we randomly selected eight (the minimum number available) for each recording station and summarized whether a common nighthawk was detected at that station for each detection type (wing-boom, call). The result was a binomial dataset of detected (1) or not detected (0) at each station (n=382) during 80 minutes of recording during peak detection probability time. We used this summary approach for each recording station instead of treating each recording as a visit with an occupancy (MacKenzie et al. 2002) or mixed modelling approach because our chosen model type, boosted regression trees, does not accommodate either approach. This standardized approach also allowed us to incorporate variation in ARU sampling at each recording station when bootstrapping (see ‘Bootstrapping’ below).

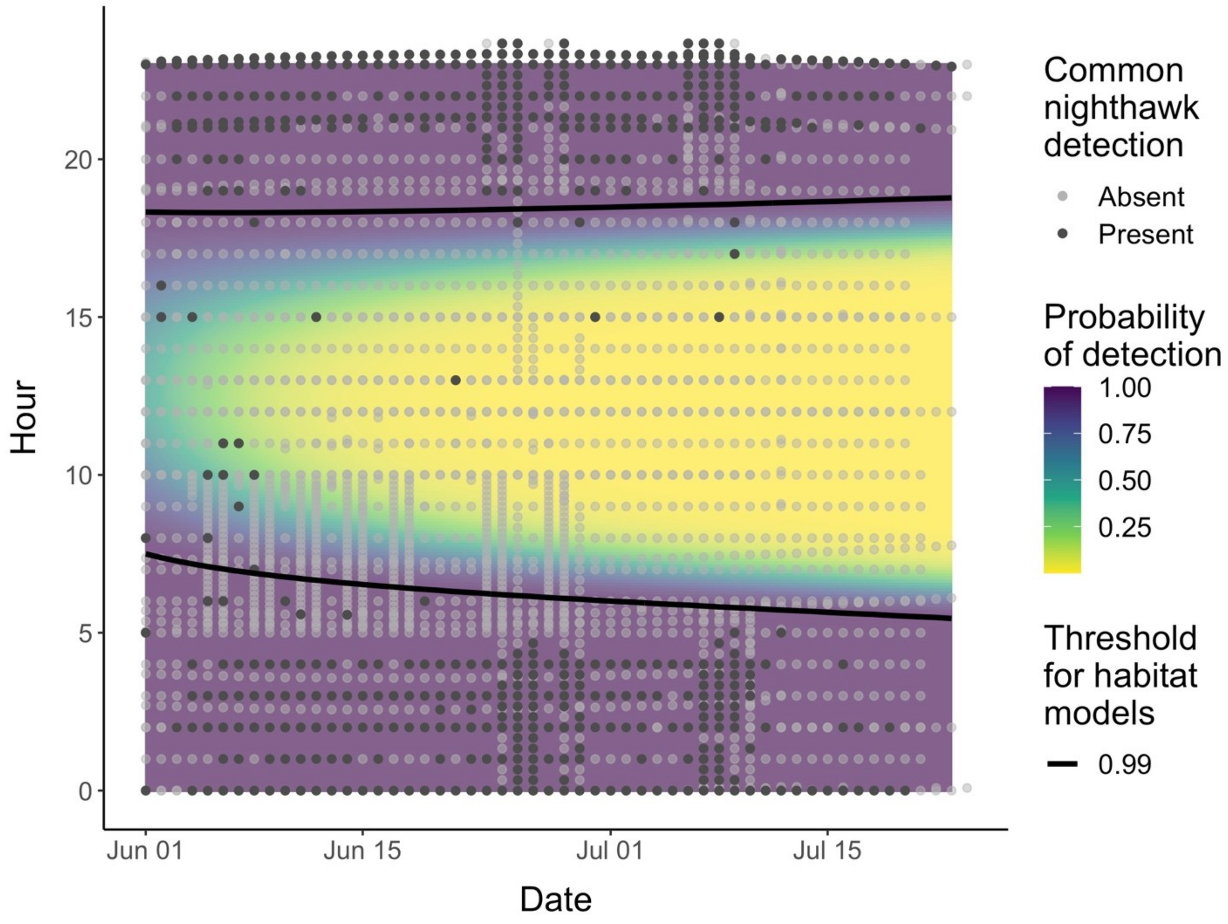


Figure 5.2. Probability of availability for detection of common nighthawks across day of year and hour of day.

Habitat Models: We modelled territorial and home range habitat use at each scale by using 10-fold cross-validation to fit a binomial BRT with the package dismo (Hijmans et al. 2020). We used detection or non-detection of the common nighthawk's territorial wing-boom signal at each recording station as the response variable for the territorial model and detection or non-detection of the more general call as the response variable for the home range model (Chapter 4). We selected a learning rate of 0.001 to ensure at least 1,000 trees (Elith et al. 2008), a tree complexity of 3 to allow up to third-order interactions between environmental variables, and a bag fraction of 0.75 for all models. The number of trees was determined by the lowest residual deviance across the ten cross-validation folds. We also used 10-fold cross validation to assess the predictive performance of each BRT.

Bootstrapping: We used bootstrapping to estimate variation in predictive performance and proportion of deviance explained for each type of habitat use at each scale. We bootstrapped the entire spatial thinning, survey effort standardization, and model fitting process 100 times, thereby incorporating variation in the station and recordings selected for each iteration. The final result was 16 sets (2 types of habitat use x 8 spatial scales) of 100 fitted boosted regression trees, each with 10-fold cross validation.

Scale of Effect: We used two performance parameters from the 10-fold cross validation to determine overall scale of effect: the area under the curve of the receiver operating characteristic (ROC AUC; Fielding and Bell 1997), and the percent test deviance explained. We determined overall scale of effect for each type of habitat use as the scale with the highest mean value across the 100 bootstraps. To determine scale of effect for each individual environmental variable, we calculated the percent test deviance explained by each variable (i.e., percent test deviance explained by the overall model multiplied by the relative influence of the variable) for each bootstrap and selected the scale with the highest mean percent test deviance explained across the 100 bootstraps.

Spatial Predictions: We used the scale of effect for each environmental variable to build a final multiscale model and make spatial predictions for each of the two types of habitat use. We used the same spatial thinning, survey effort standardization, and model fitting approach, and bootstrapped the whole process 100 times. For each bootstrap, we fit the model to the appropriate raster layers (resampled to 100 m resolution for efficiency) to create spatial predictions of common nighthawk territory and habitat use in northeastern Alberta. We then evaluated the performance of each prediction by comparing it to the data from the recording stations that were withheld during spatial thinning. We created this test dataset for each bootstrap by randomly selecting one test station for each remaining 1 km grid and using the same survey effort standardization method used to compile the training data. We used the dismo package (Hijmans et al. 2020) to calculate three metrics of predictive performance for each spatial prediction: (1) the ROC AUC, (2) the maximum correct classification rate of the model across a range of probability thresholds from 0 to 1, and (3) the maximum Cohen's kappa coefficient across a range of probability thresholds from 0 to 1 (Cohen 1960). We repeated the same prediction and evaluation process for the single scale models at the overall scale of effect to compare the performance of our multiscale and single scale models.

Covariate Effects: We used generalized additive models (GAMs) in the mgcv package (Wood 2011) to summarize the partial predictions of each environmental variable at each scale across the 100 bootstrapped models. We chose GAMs because boosted regression trees model nonlinear responses (i.e., partial predictions) and GAMs can model complex, nonlinear patterns by averaging multiple regressions with varying coefficients. For each environmental variable, we fit a Gaussian GAM with a penalized thin plate regression spline to the bootstrapped marginal effects predictions estimates with the value of that environmental variable as the predictor variable. We used generalized cross validation to determine the optimal number of knots (Table 5.1).

Results

Detections

There were 65,358 ten-minute recordings available from across our 827 recording stations. The detectability analysis indicated that common nighthawks in our study area have a detection probability of 0.99 after approximately 7:30 PM and before 5:00 AM, with some seasonal variation due to an interaction with day of year (Figure 5.2). Of those 65,358 recordings, common nighthawk home range behaviour (vocalization) was detected in 3,810 recordings and territorial behaviour (wing-boom) was detected in 1,174 recordings. Common nighthawk territorial and home range behaviour was detected at 17.3% and 44.4%, respectively, of the 827 recording stations. After spatial thinning to 382 recording stations and survey effort standardization to 80 recording minutes per station, the mean number of detections of common nighthawk territorial and home range behaviour was 8.6% and 21.6%, respectively.

Overall Scale of Effect

The scale of effect of territorial habitat use was smaller than the scale of effect for home range habitat use (Figure 5.3). Both the ROC AUC and the percent deviance explained indicated that 0.2 km was the overall scale of effect for territorial habitat use, with 0.1 km a close second. At scales larger than 0.2 km, both performance metrics dropped substantially and continued to decline with increasing scales. For home range habitat use, the two metrics supported different overall scales of effect. The ROC AUC indicated that 6.4 km was the overall scale of effect, while the percent deviance explained indicated 1.6 km was the overall scale of effect; however, the difference between performance of models built with 1.6, and 6.4 km scales was insignificant for both metrics. Model performance was lowest at the smallest scales for common nighthawk home range use. There were strong (mean of at least 6.3 across the 100 bootstraps) two-way interactions in all boosted regression tree models that likely contributed to the overall scale of effect. Interactions were stronger for the territorial movement scale than the home range movement scale and were also generally stronger at smaller scales.

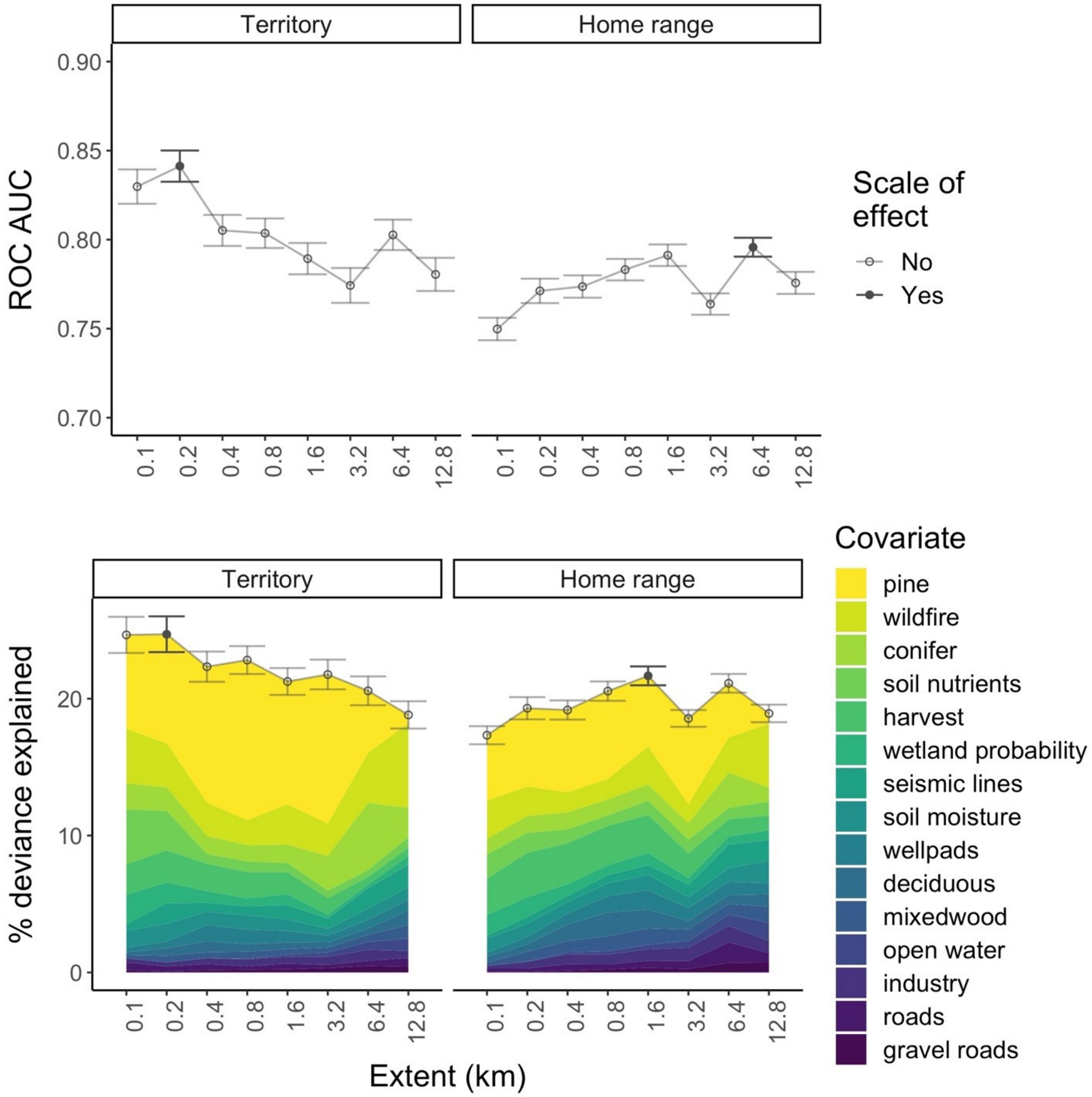


Figure 5.3. Model performance (top) of species distribution models for territorial and home range habitat use of common nighthawks. Error bars show the 95% confidence interval across 100 bootstraps and bold indicates the selected scale of effect for that model and/or variable.

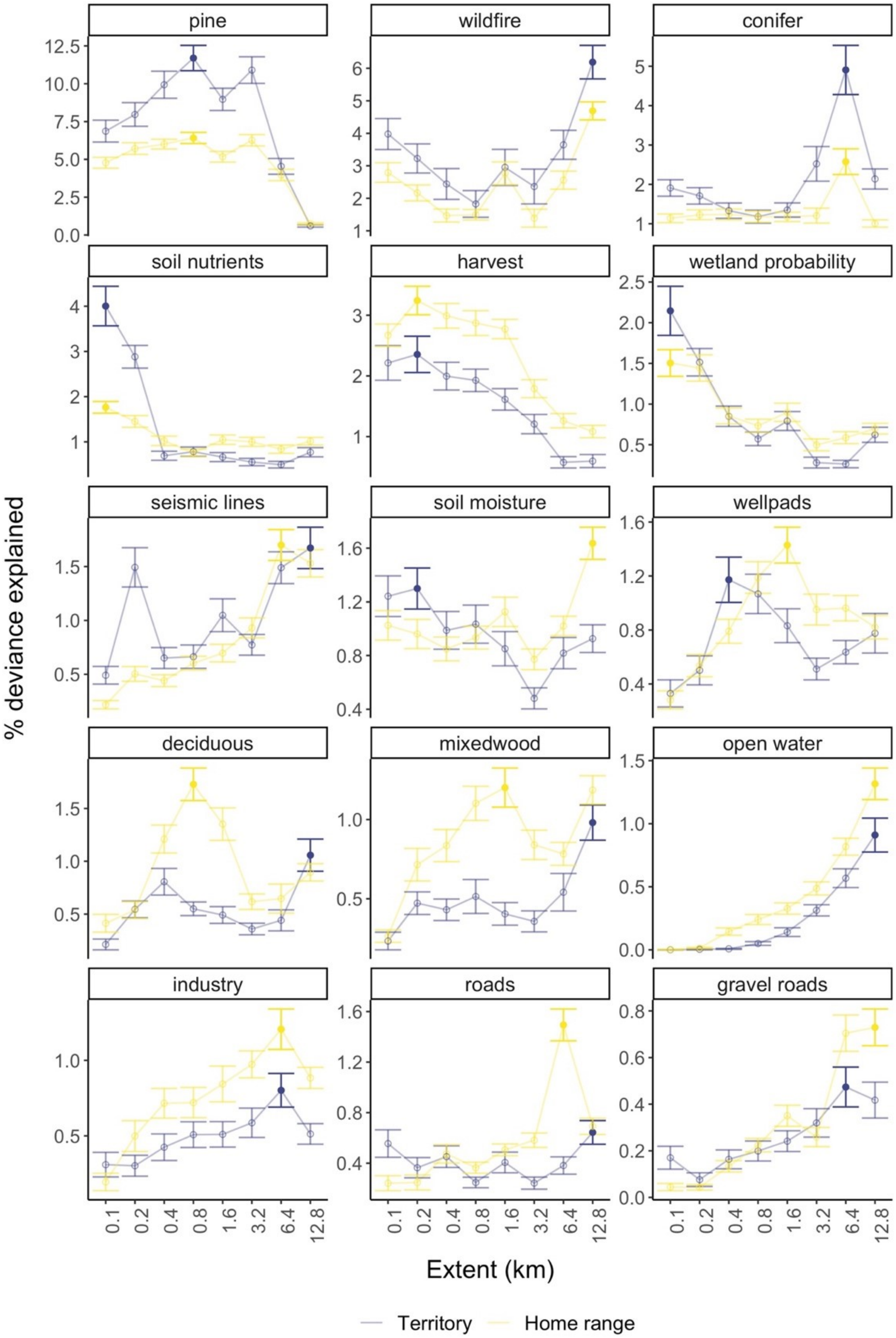


Figure 5.4. Percent deviance explained by individual environmental variables in species distribution models for two territorial and home range habitat use of common nighthawks. Error bars show the 95% confidence interval across 100 bootstraps and bold indicates the selected scale of effect for that model and/or variable.

Scale of Effect of Individual Environmental Variables

The scale of effect for individual environmental variables varied widely, with little relationship between the scale of effect for the most important variables and the overall scale of effect (Figure 5.4). Interestingly, the scale of effect for each environmental variable was similar between territorial and home range habitat use, with the exception of soil moisture, well site, mixedwood forest, and deciduous forest. The difference in overall scale of effect was instead driven by variation in proportion of deviance explained between the two types of habitat use. The scale of effect for territorial habitat use was driven primarily by a strong scale of effect for harvest index and soil nutrient level, both of which were among the top five most predictive environmental variables. The environmental variables driving the scale of effect for home range habitat use were less clear.

Spatial Predictions

The models for both types of habitat use predicted similar regions of high probability of habitat use, particularly the area in the northeastern section of the study area that burned in 2011 and is characterized by jack pine (*Pinus banksia*) forest with sandy soils (Figure 5.5). Overall, home range habitat use probability was higher than territorial habitat use probability across the study area, but also had a higher standard deviation across bootstraps. Using the scale of effect for each individual environmental variable in a multiscale model did not substantially improve model performance over single scale models at the overall scale of effect. The cross validation from model fitting indicated that the multiscale model was an improvement for home range habitat use; however, evaluation of the spatial predictions using the withheld data suggested this was not the case (Figure 5.5). Alternatively, the multiscale model provided slight improvement

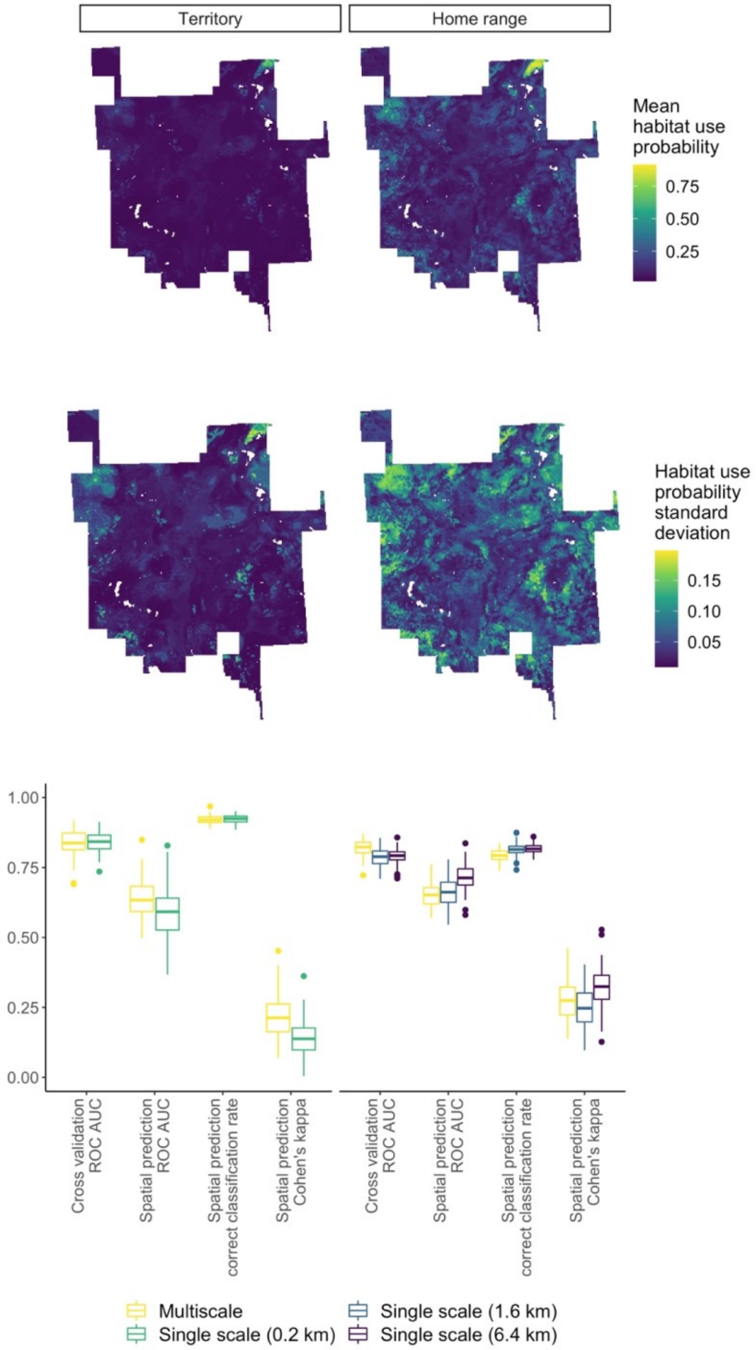


Figure 5.5. Mean and variance of 100 spatial predictions for territorial and home range habitat use of common nighthawks and predictive performance of those spatial predictions. Predictive models are multiscale models that include the specific scale of effect for each individual environmental variable included in the model. Predictive performance was evaluated with cross validation during model fitting ("cross validation) and after spatial prediction with a withheld dataset ("spatial prediction"). Lower and upper box boundaries are 25th and 75th percentiles; lower and upper whiskers are 1.5 times the interquartile range; individual points are those outside 1.5 times the interquartile range.

in model performance of territorial habitat use for the spatial predictions (mean 0.04), but not during cross validation. Comparison of ROC AUC from cross-validation and spatial prediction suggested cross-validation evaluation was overly optimistic; spatial prediction performance evaluation suggested model performance was “fair” (0.6-0.7; Duan et al. 2014).

Covariate Effects

Overall, the effect of the selected covariates was similar between territorial and home range habitat use when the scale of effect was also the same (Figure 5.6). When the scale of effect differed, the shape of the marginal effect generally also differed (e.g., soil moisture, roads). The proportion of pine forest was the strongest predictor for both territorial and home range habitat use (Table 5.1). The probability of habitat use was relatively low for both movement scales until approximately 30% cover, and then probability increased steadily to 0.3 (territory) and 0.5 (home range) probability of use at 100% cover (Figure 5.5). Wildfire index, harvest index, and the proportion of conifer forest were the next three strongest predictors, although their relative influence varied between the movement scales. All three variables had a positive influence on common nighthawk habitat use, but with varying thresholds for the values at which that effect took place. There was a particularly strong interaction between the proportion of pine forest and forest harvest (mean 13.3 across all bootstraps of all single scale models), particularly in smaller scale models.

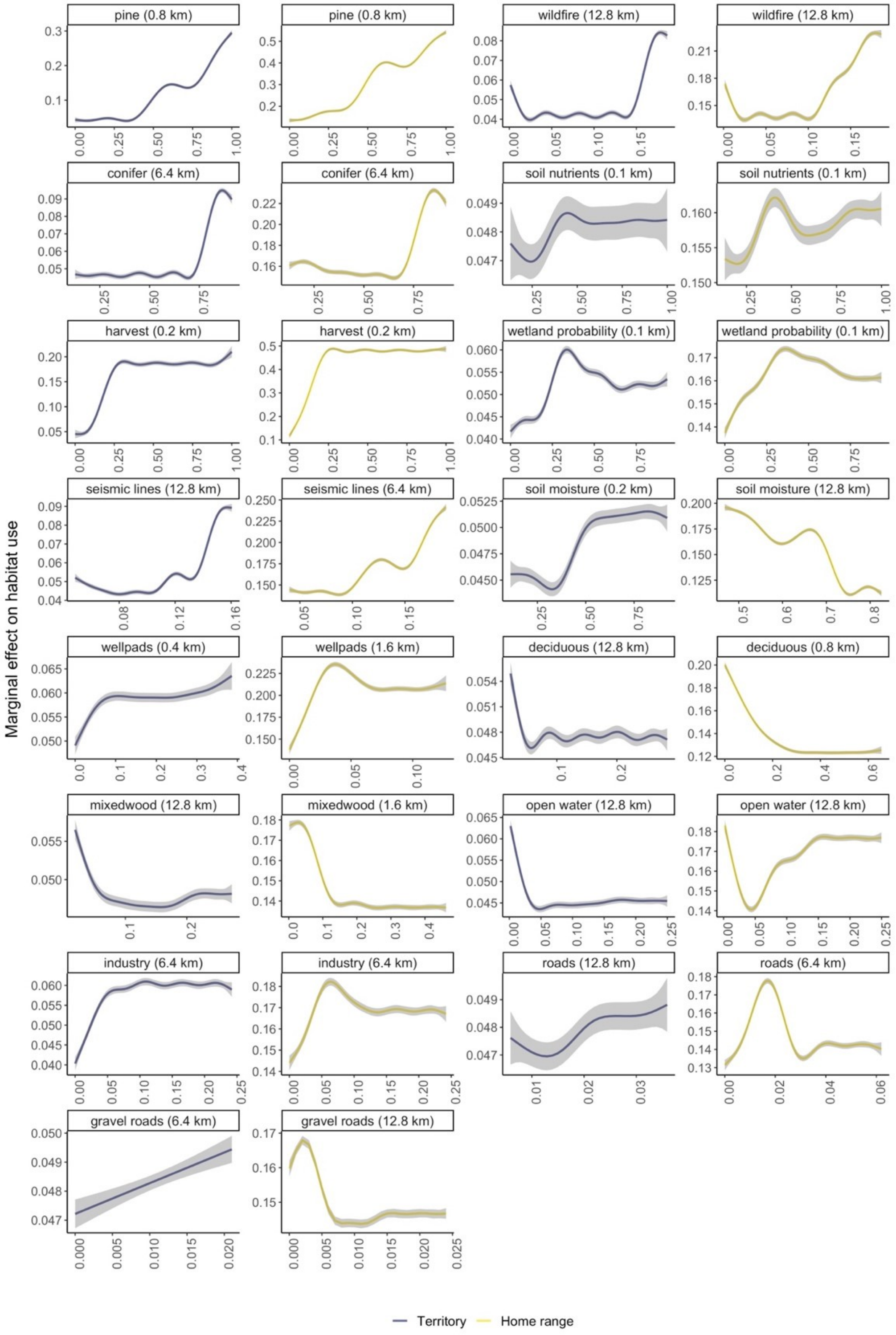


Figure 5.6. Marginal effect for the scale (extent) of effect of 15 environmental variables on territorial and home range habitat use of common nighthawks. Plotted effects are the mean (line) and 95% confidence interval (shaded ribbon) of a generalized additive model (GAM) fit to 100 bootstrapped species distribution models.

The GAMs used to fit the marginal effect plots varied in degrees of freedom from one to 9, and the amount of variation explained (R^2) by the GAM also varied from <0.01 to 0.69 (Table 5.1). There was an overall positive relationship between the relative influence of the covariates included in the boosted regression tree models and the R^2 value of the GAM fit to the 100 bootstraps of each model, suggesting that the stronger predictors had more consistent marginal effects on habitat use. The exception was soil nutrients, which explained 5.43% of the variation in the model, but the effect of the relationship was highly variable across bootstraps.

Table 5.1. Scale (extent) of effect and relative influence at that scale for 15 environmental variables on territorial and home range habitat use of common nighthawks. Standard deviation (SD) of relative influence was calculated across 100 bootstrapped species distribution models. Generalized additive models (GAMs) were used to smooth the marginal effects of each environmental variable on habitat use.

Habitat use type	Covariate	Scale of effect (km)	Relative influence (SD)	GAM df	GAM R^2
territory	pine	0.8	32.43 (13.03)	9	0.53
territory	wildfire	12.8	10.52 (7.59)	9	0.29
territory	conifer	6.4	9.93 (7.99)	9	0.3
territory	harvest	0.2	9.61 (4.88)	9	0.27
territory	soil nutrients	0.1	5.43 (3.1)	6	<0.01
territory	seismic lines	12.8	5.4 (3.97)	9	0.34
territory	soil moisture	0.2	5.21 (3.41)	7	0.05
territory	wetland probability	0.1	5.13 (3.92)	9	0.1
territory	industry	6.4	3.31 (2.34)	8	0.09
territory	open water	12.8	3.11 (1.86)	8	0.07
territory	deciduous	12.8	2.9 (2.69)	9	0.01
territory	well sites	0.4	2.47 (2.1)	6	0.02
territory	mixedwood	12.8	2.23 (1.67)	7	0.03
territory	roads	12.8	1.24 (1.02)	4	<0.01
territory	gravel roads	6.4	1.08 (0.88)	1	<0.01
home range	pine	0.8	23.15 (6.16)	9	0.69
home range	harvest	0.2	12.17 (3.37)	9	0.56
home range	wildfire	12.8	7.53 (4.35)	9	0.35
home range	conifer	6.4	6.92 (4.29)	9	0.36
home range	soil moisture	12.8	6.38 (2.4)	9	0.54

Habitat use type	Covariate	Scale of effect (km)	Relative influence (SD)	GAM df	GAM R^2
home range	deciduous	0.8	6.36 (2.45)	6	0.38
home range	oil well sites	1.6	5.53 (2.88)	8	0.15
home range	roads	6.4	5.08 (2.41)	9	0.18
home range	seismic lines	6.4	5.03 (2.48)	9	0.45
home range	open water	12.8	4.59 (1.85)	9	0.17
home range	industry	6.4	4.59 (2.03)	8	0.06
home range	mixedwood	1.6	3.42 (2.03)	9	0.3
home range	soil nutrients	0.1	3.34 (1.78)	8	0.01
home range	wetland probability	0.1	3.33 (1.66)	9	0.13
home range	gravel roads	12.8	2.58 (1.32)	9	0.11

Discussion

We defined a new type of scale of effect, "overall scale of effect" (i.e., scale of effect of an entire habitat model), and compared the overall scale of effect of two different types of habitat use for a single species, a line of inquiry first suggested by Miguet et al. (2016). We built territorial and home range habitat models for the common nighthawk to explore whether overall scale of effect corresponds to movement extent. We found support for our prediction: the overall scale of effect for common nighthawk home range habitat use was larger than the overall scale of effect for territorial habitat use. In contrast, the scale of effect for individual environmental variables varied widely, with little relationship between the scale of effect for the most important variables and the overall scale of effect. Using the scale of effect for each environmental variable in a final model did not improve spatial predictive performance relative to the single scale model for the overall scale of effect.

Support for a relationship between movement and scale of effect is sparse (reviewed by Jackson and Fahrig 2015), despite the obvious mechanistic link, and we suggest the lack of evidence may be in part because previous studies have not examined overall scale of effect. Our results support theoretical predictions that the scale of effect of environmental variables on species occurrence or abundance is driven by local movements (Jackson & Fahrig 2012; Miguet et al. 2016). The machine learning approach we used here may be particularly valuable in linking overall scale of effect to movement because not only does it allow for inclusion of many potential predictors, but it also models complex interactions that could be driven by local movement (Elith et al. 2008). For the common nighthawk, the overall scale of effect for territorial habitat use (200 m) corresponded almost perfectly to the mean radius size of territories in the study area (186 m; Chapter 4). Relating the overall scale of effect to home range movement is less straightforward. Tracking data are unavailable to estimate home range size for this species; however, the home range size available for a single individual within the study area (40 km^2; 3.6 km radius; *unpublished data*) suggests that the overall scales of effect (1.6 km, 6.4 km) we found correspond approximately to home range size. The other reason comparing overall scale of effect to home range size is difficult is because we found support for two overall scales of effect (1.6 km and 6.4 km). Martin and Fahrig (2012) suggest multiple scales of effect indicate a particular variable is important for more than one life stage. Although there is evidence that dispersal movement can explain scale of effect (Jackson & Fahrig 2012; Ricci et al. 2013), this is an unlikely explanation for multiple home range scales of effect because anecdotal evidence suggests common nighthawk dispersal range is two orders of magnitude larger than the overall home range scale of effect (several hundred kms; *unpublished data*). Predictions by Tittler (2008) also suggest that scale of effect is not related to dispersal for species that travel within large home ranges. Instead, we suggest the signal for scale effect for the home range is less clear perhaps because home range use combines many different drivers of movement (e.g., foraging, mating, roosting) and thus there is more noise in home range detection data. This conclusion is supported by a lower proportion of variation explained in the home range models than the territorial detection models.

Our prediction that individual environmental predictors would have varying scales of effect unrelated to the overall scale of effect was also supported. Variation in scale of effect across environmental predictors is a known phenomenon (Martin 2018; Galán-Acedo et al. 2018), but how to interpret this variation? If local movement also determines the scale of effect of individual variables, then the variables that share the overall scale of effect should be those that influence that movement behaviour most strongly. Our results are not consistent with this; for example, the variables that shared a scale of effect with territorial habitat use (forest harvest and soil moisture) were the 4th and 8th most predictive of territorial habitat use. Interpretation of scale of effect for individual environmental variables may be less straightforward and influenced by factors other than movement. Options include resource availability, predator movement, competition, and density dependence (Miguet et al. 2016). In fact, the similarity in scale of effect and marginal effects between territorial and home range habitat use for the top explanatory variables (pine, wildfire, conifer, soil nutrients, harvest) suggest that the scale of effect of these variables is explained by factors other than movement. For example, the scale of effect for pine forest, strongest explanatory variable, was 0.8 km for both types of habitat use, suggesting perhaps an indirect effect of conspecific attraction to neighbouring common nighthawks. The scale of effect for the next strongest explanatory variable, fire, was the largest scale measured (12.8 km) for both habitat use types and suggests we did not measure large enough scales to adequately detect the scale of effect for this variable (Jackson & Fahrig 2015). Fire scale of effect may perhaps be driven by dispersal because common nighthawks are a post-disturbance specialist in the boreal forest (Chapter 6), and dispersal may be critical to their population dynamics, as it is for Black-backed Woodpeckers (*Picoides articus*; Tingley et al. 2018). In contrast, environmental variables that differ in scale of effect between territorial and home range habitat use may be those that are specifically important for that behaviour. For example, soil moisture had opposite scales of effect for territorial and home range habitat use (0.2 vs 12.8 km). Concordant with common nighthawk ecology, soil moisture also had opposing marginal effects. Nesting territories require low soil moisture at small scales because nighthawks lay their eggs directly on the ground, while high soil moisture likely affects home range habitat at large scales use because nighthawks forage over wetlands for aerial insect prey (Brigham et al. 2011).

Due to the influence of context-dependent factors like density dependence and resource availability, scale of effect is often population-specific (McGarigal et al. 2016; Miguet et al. 2016; Galán-Acedo et al. 2019). For example, scales of effect are predicted and empirically shown to be lower in fragmented landscapes than contiguous landscapes because lower connectivity in fragmented landscapes leads due to disrupted dispersal (Galán-Acedo et al. 2019; Miguet et al. 2016). We suggest overall scale of effect is likely also population-specific, as movement ranges are also context-dependent. We reiterate the recommendations of others that caution should be taken when attempting to generalize the overall scales of effect from our study to other regions (Galán-Acedo et al. 2019; Miguet et al. 2016). In terms of scales of effect for individual environmental variables, the scale of effect for open wetland availability on common nighthawk home range habitat in Ontario, Canada (1.5 km) (Farrell et al. 2019) was much smaller from that reported here for open water (12.8 km; potentially larger) (Jackson & Fahrig 2015). This difference may indicate a greater importance of open water wetlands for common nighthawk foraging in Ontario relative to northern Alberta. Alternatively, the difference in scales of effect may simply be an artifact of varying quantitative definitions of wetland availability. In fact, methodological differences between studies may be one of the reasons for region-specific conclusions about scale of effect.

Ultimately, the purpose of determining and understanding scale of effect is to inform ecological analyses and inferences. McGarigal et al. (2016) recommend determining scale of effect for each environmental variable prior to building a final model to produce a “truly optimized multi-scale, multi-variable model” and comparing this model to the best single scale model. Similar to Martin and Fahrig (2012), we found minimal to no improvement in predictive performance from an optimized multiscale model. Other comparisons of multiscale and single scale model performance found a slight improvement of approximately 0.05 increase in ROC AUC (refs from Martin & Fahrig); however, variability in performance across bootstraps shown here and elsewhere (Martin and Fahrig 2012) suggests such improvements should be taken with a grain of salt. Martin and Fahrig (2012) provide two explanations for the similar or superior performance of single scale models over multiscale models. First, that similarity between single scale and multiscale models can be driven by same scale of effect of particular predictors. Second, if strong predictors have different scales of effect from the overall scale of effect, single scale models can still be good predictors if those predictors still have a strong effect at the overall scale of effect. Our study supports this second explanation, as the scale of effect of the most important predictors differed from the overall scale of effect. We further suggest that differences in the interactions between single and multiscale models may be partially responsible for the failure of multiscale models to perform better in our study.

The ideas and analysis here represent preliminary evidence that contributes to an ongoing conversation about the mechanistic underpinnings of scale of effect. If movement range is indeed a determinant of scale of effect but has gone undetected due to a focus on individual environmental variables, then our idea of overall scale of effect may be the first step towards “a framework for predicting scales of effect a priori” (Miguet et al. 2016). We used a novel approach of comparing habitat use of different behaviours of a single species to test this prediction. Next steps should explore a link between movement range and overall scale of effect for other populations or species, and then test for a relationship between overall scale of effect and home range size, as has been done for scale of effect of single environmental variables (Tittler 2008; Fisher et al. 2011; Desrochers et al. 2010). We note, however, that these predictions are likely specific to occurrence, occupancy, or abundance as the response variable. The scale of effect of other response variables such as species richness, reproductive success, or genetic diversity are likely driven by other mechanisms (Miguet et al. 2016; Moraga et al. 2019).

Regardless, our study suggests that the interpretation of scale of effect may be more nuanced than previously thought. Researchers and practitioners should consider differentiating between overall and individual variable scale of effect. Martin and Fahrig (2012) suggested that single scale models at the overall scale of effect could be used to inform overarching management actions for simplicity’s sake. We show however, that the scale of effect of the most important environmental variables does not necessarily correspond to the overall scale of effect; using a single scale of effect for management may still lead to suboptimal wildlife options (Thornton & Jr 2013; Holland et al. 2004; Smith et al. 2011). For example, the proportion of pine forest has the strongest positive effect on common nighthawk territorial habitat use but using the overall scale of effect or territory radius of 200 m to manage forest harvest of this species at risk may lead to less suitable nesting habitat. Furthermore, basing the scale of measurements for individual environmental predictors on estimates of movement ranges is inappropriate because our study suggests that the scale of effect of individual predictors is more likely attributed to other factors. Finally, our study adds to the large and continually growing body of literature that supports the importance of multiscale habitat studies as a critical step in informing wildlife management (Mayor et al. 2009; McGarigal et al. 2016; Martin 2018).

Supplementary Materials

Geospatial data sources.

Table 5.A.1 Source geospatial datasets used to extract environmental variables for common nighthawk habitat modelling.

Dataset name	Description	Proprietor	Original data format (cell size)	Variables extracted (calculations or reclassification)	Available from (access date)	Available reference or metadata
Historical Wildfire Perimeters	Polygon features representing wildfire perimeters dating back to 1931. The datasets from 1998 on should include most of the fires 12 hectares (29.7 acres) or larger that occurred within the legislated Forest Protection Area. The datasets prior to these years may include some fires smaller than 200 hectares (494.2 acres), but the emphasis on the data collection for these years was to provide Class E fires (wildfire larger than 200 hectares or 494.2 acres).	Alberta Agriculture and Forestry	Vector	Wildfire (weighted by time since fire)	https://wildfire.alberta.ca/resources/historical-data/spatial-wildfire-data.aspx (March 5, 2019)	https://wildfire.alberta.ca/resources/historical-data/default.aspx

Dataset name	Description	Proprietor	Original data format (cell size)	Variables extracted (calculations or reclassification)	Available from (access date)	Available reference or metadata
Human Footprint Inventory 2016	Twenty-one human footprint categories (based on more than 115 anthropogenic disturbance types) collated into a single integrated dataset by applying a specific order of precedence. The dataset was developed in part by a collaborative effort of data accumulation between the Government of Alberta and the ABMI, together forming the Alberta Human Footprint Monitoring Program (AHFMP) to better enhance and update human footprint data in Alberta.	Alberta Biodiversity Monitoring Institute	Vector	Harvest (class 17; weighted by time since harvest) Wellsite (classes 9 and 16; weighted by time since deactivation) Roads (class 3) Gravel roads (class 3, code 305 and 306) Seismic lines (class 20) Industrial (classes 1-16)	https://www.abmi.ca/home/data-analytics/da-top/da-product-overview/Human-Footprint-Products/HF-inventory.html (September 9, 2019)	https://ftp-public.abmi.ca/GISData/HumanFootprint/2016/HFI2016_Metadata.pdf
Alberta Vegetation Inventory	This photo-based digital inventory dataset is a compilation of the original blocks of Crown Alberta Vegetation Inventory data collected using aerial photography dating between 1987 and 2014. A variety of standards was used to collect the original blocks of data. It was developed to identify the type, extent and conditions of vegetation in the province.	Alberta Agriculture and Forestry	Vector	Pine (SP1=Pj)	https://geodiscover.alberta.ca/geoportal/rest/metadata/item/100b275712b442acbda4a0358d8a4951/html (Feb 12, 2013)	https://www.alberta.ca/vegetation-inventory-standards.aspx

Dataset name	Description	Proprietor	Original data format (cell size)	Variables extracted (calculations or reclassification)	Available from (access date)	Available reference or metadata
Boreal Wetland Probability	The Boreal Wetland probability dataset was generated using a machine learning framework in R statistical software. The machine learning model was based off of a Digital Elevation Model, optical satellite data, and Synthetic Aperture Radar data. The model was trained with photo-plots that were derived from high resolution 3D image interpretation and give detailed attribution of land cover information.	Alberta Biodiversity Monitoring Institute	Raster (10 m)	Wetland probability	https://abmi.ca/home/data-analytics/da-top/da-product-overview/Advanced-Landcover-Prediction-and-Habitat-Assessment—ALPHA—Products/Boreal-Wetland-Probability-Data.html (September 10, 2019)	Hird et al. 2017
2010 Land Cover of Canada	The Canada Centre for Remote Sensing created the 2010 Land Cover of Canada map as the Canadian contribution to the 2010 Land Cover Map of North America under the North American Land Change Monitoring System. The dataset was produced using observations from Thematic mapper ™ and Enhanced Thematic Mapper (ETM+) Landsat sensors.	Natural Resources Canada	Raster (30 m)	Deciduous (class 5) Coniferous (class 1 and 2) Mixedwood (class 6) Open water (class 18)	https://open.ccanad.ca/data/en/dataset/c668b87f-e85f-4842-b0e1-a8f79ebf1133 (September 9, 2019)	https://geoappext.nrcan.gc.ca/arcgis/rest/services/FGP/LandCover_EN/MapServer/0

Dataset name	Description	Proprietor	Original data format (cell size)	Variables extracted (calculations or reclassification)	Available from (access date)	Available reference or metadata
Alberta Derived Ecosite Phase	The province of Alberta is divided into 21 natural subregions, each of which contain distinct plant communities, or 'ecosites'. The Derived Ecosite Phase (DEP) dataset was developed by using the Alberta Vegetation Inventory and several LiDAR-derived datasets to classify areas into ecosites. Each ecosite is characterized by a unique combination of soil nutrient and moisture conditions.	Alberta Agriculture and Forestry	Vector	Soil moisture (moisture regime classes 2-9 rescaled from 0 to 1) Soil nutrients (nutrient regime classes A-E rescaled from 0 to 1)	https://open.alberta.ca/dataset/derived-ecosite-phase-version-2 (September 9, 2019)	https://open.alberta.ca/dataset/49cd2d77-4153-41bd-b0fc-059627b0a7a9/resource/64616ee4-8c5c-4e70-ab7c-d2ef002ca79a/download/af-derived-ecosite-phase-version-2-0-2020.pdf

Survival model comparison results.

Table 5.A.2. AICc ranking of parametric survival regression models for estimation of common nighthawk availability for detection in acoustic recordings. Day represents day of year. Sin and cos represent trigonometric functions of time of day and were used to allow for circularity. All models with quadratic terms also included a linear term of the same variable. Bold indicates the model selected to estimate availability for detection in acoustic recordings.

Model	df	logLik	AIC$_c$	ΔAIC$_c$	AIC$_c$ w
~day^2*sin^2 + day^2*cos^2	**11**	**34226.4**	**-68430.9**	**0.0**	**1.0**
~day*sin^2 + day*cos^2	10	34095.7	-68171.4	259.5	0.0
~day^2 + sin^2 + cos^2	7	33963.7	-67913.5	517.4	0.0
~day + sin^2 + cos^2	6	33848.4	-67684.7	746.2	0.0
~sin^2 + cos^2	5	33810.6	-67611.3	819.6	0.0
~day^2	3	29897.2	-59788.3	8642.5	0.0
~day	1	29563.2	-59124.4	9306.5	0.0
~1	2	29563.7	-59123.4	9307.4	0.0

Chapter 6. Specialist or generalist? It depends. Context-dependent habitat relationships provide insight into forest disturbance effects for a boreal bird species

Abstract

Understanding how disturbance affects species is a critical component of management in the boreal forest, particularly for disturbance specialist species that often help initiate the regeneration process. Emulation of natural disturbance rarely incorporates the disturbance response of those specialist species, which may be potentially sensitive to disturbance characteristics because they are attracted to resources provided by the disturbance itself. Identifying disturbance specialist species and the reasons they occur in post-disturbance habitat is thus a priority for maintaining biodiversity in managed forests. We used a novel approach of comparing two types of habitat use to fill this information gap for a potential boreal forest disturbance specialist, the common nighthawk (*Chordeiles minor*). The common nighthawk is of particular interest to forest managers because this wide-ranging species is declining and listed as of conservation concern in multiple jurisdictions. We used passive acoustic monitoring to survey for nighthawks at 400 locations in the boreal forest of northeastern Alberta, Canada where land use pressure for resource extraction is high. We used an occupancy modelling framework to compare two competing hypotheses for why nighthawks use post-disturbance habitat: nesting or foraging resources. The nesting resource hypothesis was supported by all three lines of evidence that we examined. First, time since disturbance negatively affected territorial habitat use, but not extraterritorial habitat use, confirming that this species is only a disturbance specialist for the territorial component of its home range. Second, territorial habitat use differed between disturbance types, with a higher probability of habitat use for post-harvest and abandoned well

sites, which are more likely to have lower amounts of residual vegetation. For our third line of evidence, open pine forest mitigated the effects of time since disturbance for territorial habitat use for all three disturbance types, but especially for postfire areas. We discuss each line of evidence and provide industry-specific recommendations for incorporating the common nighthawk into forest management plans. Our research emphasizes that understanding the context-dependence of species responses to disturbance provides insights that facilitate effective forest management. We advocate that understanding the behavioural context of habitat use can provide this insight and is often more likely to be aligned with the operational scale of forest management.

Introduction

The ecology and management of the boreal forest are governed by disturbance. The frequency, intensity, size, season, and spatial distribution of recurrent fires control vegetation structure and species assemblage, nutrient cycling, energy flow and the resulting wildlife community (Bonan & Shugart 1989; McLauchlan et al. 2020). The distribution of fire characteristics on the landscape creates a "shifting mosaic" of heterogeneous patches of forest that change dynamically over time (Bormann and Likens, 1979). More recently, this mosaic includes patches of anthropogenic disturbance. Most of the southern extent of the North American boreal forest is used for forestry (Pasher et al. 2014; Pickell et al. 2015; Timoney 2003). In the northwest, oil and natural gas extraction is increasingly prominent, and the resultant seismic exploration lines, pipelines, well sites, and open pit mines are novel disturbances relative to the natural boreal forest disturbance regime (Pickell et al. 2015).

One approach to mitigating the effects of these anthropogenic disturbances is the "emulation of natural disturbance", which mimics the frequency, size, and residual organic matter of the long-term spatiotemporal patterns of natural disturbance through actions like variable retention and prescribed burning (Hunter 1993; Kuuluvainen & Grenfell 2012; Long 2009). To successfully maintain biodiversity on the boreal landscape, emulation of natural disturbance must provide the same balance of resources for wildlife as the natural disturbance regime. Understanding how wildlife responds to disturbance is thus important for informing forest management; however, disturbance ecology is primarily founded in the responses of disturbance tolerant or avoidant species and largely ignores the resource requirements of disturbance specialist species. Disturbance specialists may be more sensitive to disturbance type because they are attracted to resources provided by the disturbance itself. For example, infestations of bark- and wood-boring beetles and the cavity-nesting opportunities provided by wildfire disturbance are the resources that attract Black-backed and other postfire specialist woodpeckers (*Picoides articus*; Hobson & Schieck 1999; Schieck & Song 2006; Hutto 2002). Meta-analysis revealed that western boreal bird communities differ between postburn and postharvest areas and only begin to converge between 11 and 30 years post-disturbance (Schieck & Song 2006). Incorporating disturbance specialists into plans for emulating natural disturbance is important because they facilitate succession by initiating seeding and accelerating nutrient cycling, among others (Lindenmayer et al. 2019).

Comparing habitat use of disturbance specialist species between disturbance types can thus provide insight into their ecology because those varying effects are ultimately driven by differences in the resources provided. Forest harvest may be planned to emulate wildfire through variable retention that mimics postfire residual vegetation structure; however, emulating the ecological processes of nutrient cycling such as postfire beetles that provide foraging resources is more difficult (Kimmins 2008, McLean et al. 2015). Disturbances linked to energy infrastructure like oil well sites, seismic exploration lines, and transmission lines are even less likely to emulate the ecological processes of fire because they are more dissimilar to the natural disturbance regime than forest harvest (Pickell et al. 2013). Post-disturbance specialist species that occupy postfire areas for foraging opportunities should thus be sensitive to disturbance type, while those that are attracted to the open vegetation structure will be less sensitive.

Comparing disturbance effects on habitat use for different behaviours can thus also provide insight into the resources that disturbed areas provide. Although songbirds typically acquire all their resources from within their territory boundaries, many other bird species and some mammal species hold territories for reproduction but undertake extraterritorial movements to obtain other resources like food (Whitaker & Warkentin 2010). Understanding the behavioural context of disturbance effects will help facilitate successful emulation of natural disturbance because it provides insight into the relevant spatial and temporal scale of disturbance effects.

In the boreal forest, the common nighthawk (*Chordeiles minor*) is thought to be a disturbance specialist because it primarily inhabits postfire and postharvest areas and is generally absent from older forest stands (Hagar et al. 2004; Foley 2018; Sidler 2017). Common nighthawks are crepuscular birds that lay their eggs directly on the ground within small, exclusive territories (~10 ha; Chapter 4) and conduct extraterritorial movements to constant-flight forage for aerial insect prey within large, overlapping home ranges (~40 km^2; *unpublished data*). Territorial and extraterritorial habitat use are differentiated by an aerial wing-boom display around the nest location, where the male produces a ‘vroom’ sound by flexing his primary feathers downward at the bottom of a steep aerial dive (Miller 1925, Chapter 4). The degree to which nighthawks use recently disturbed areas for foraging resources relative to nesting resources is unknown. The absence of vegetation and abundance of postfire wood-boring beetles likely provide foraging opportunities and common nighthawk diet in the boreal forest is dominated by wood-boring beetles (Knight et al. 2018), suggesting that the resource provided by disturbed areas in the boreal forest is aerial insect prey. In contrast, the scarcity of vegetation likely provides suitable nesting substrate and in other parts of its breeding range, however, the common nighthawk inhabits open vegetation types such as grasslands, sand spits, and pine forests (Brigham et al. 2011), suggesting that the resource provided by disturbed areas in the boreal forest is scarcity of vegetation. Furthermore, common nighthawks in northwestern Ontario's boreal forest, occupied postfire, postharvest, and wetland areas equally, further suggesting no reliance on disturbed areas for foraging; however, this study only examined home range habitat use (Farrell et al. 2017). The common nighthawk is a species of conservation concern across its range due to population declines with habitat loss via fire suppression as a potential mechanism for those declines (Environment Canada 2016a); therefore, a species-specific understanding of its disturbance response and the appropriate scale of that response is critical for incorporation into forest management plans.

Table 6.1. Three sets of predictions for each of two competing hypotheses about common nighthawk use of post-disturbance areas in the boreal forest.

	Nesting Resource Hypothesis: post-disturbance areas provide nesting habitat	Foraging Resource Hypothesis: post-disturbance areas provide foraging habitat
1. Time since disturbance	Time since disturbance only affects territorial habitat use because succession affects availability of bare ground.	Time since disturbance affects both territorial and extraterritorial habitat use because it affects open vegetation structure for foraging and aerial insect availability (Morissette et al., 2002; Schowalter et al., 1981).
2. Disturbance type	Postfire habitats have lower initial probability of territorial habitat use due to greater amounts of coarse woody debris and retained vegetation on the ground (McRae et al., 2001, Bognounou et al. 2021).	Postfire habitats have higher initial probability of territorial and extraterritorial habitat use due to an abundance of wood-boring beetle prey and other aerial insects associated with retained vegetation (Deans et al., 2005).
3. Vegetation	Pine forest mitigates negative effects of time since disturbance on territorial habitat use because the sandy soils remain relatively free of ground cover.	Wetland probability mitigates negative effects of time since disturbance on both types of habitat use because it provides abundant aerial insects. Pine forest does not mitigate negative effects of time since disturbance because the sandy soils drive low aerial insect productivity.

Our goal was to inform boreal forest management for the common nighthawk by confirming that this species is a disturbance specialist and understanding how it uses post-disturbance areas. We explored differences in common nighthawk habitat use between natural and anthropogenic disturbance types to derive three lines of evidence for the competing hypotheses that common nighthawks use post-disturbance areas for nesting or foraging (Table 6.1). We used a large, archived bioacoustic data set to study common nighthawk habitat use in multiple disturbance types in Canada's boreal forest. We selected recordings from sites across a range of times since disturbance for three disturbance types: wildfire, forest harvest, and petroleum well site (hereafter "fire", "harvest", and "well site"). We determined common nighthawk detection or non-detection in each recording with automated processing and used those detections in an occupancy framework to study probability of habitat use. We built two sets of analyses to compare between different types of habitat use by using the common nighthawk wing-boom signal to differentiate between territorial and extraterritorial habitat use. For each set, we conducted our analysis in three stages. First, we determined which detectability covariates to include in our occupancy models. Second, we tested for effects of two vegetation covariates, pine forest, which provides open area with minimal ground vegetation for nesting, or wetland probability, which provides abundant aerial insects for foraging. This second step was a prerequisite to evaluating our hypotheses about whether vegetation can mitigate disturbance effects (Table 6.1). Third, we fit a model that tested our hypotheses by testing for effects of disturbance type, time since disturbance, and vegetation covariates.

Methods

Study Area

We conducted our study in north-eastern Alberta where the boreal forest is subject to intensive resource extraction including active forestry and multiple types of oil and gas exploration and extraction (Figure 6.1). The study area was primarily within the Central Mixedwood Natural Subregion (Natural Regions Committee 2006) where the natural landscape is characterized by a mosaic of upland forest types and lowland peat bogs and fens. Our study area is an approximately 75,000 km2 area based on available sampling locations and extent of available environmental data (Supplementary Materials).

Site Selection

We selected sites from available survey locations in the WildTrax database, which is an online archive of audio recordings from autonomous recording units (ARUs) deployed for various research and monitoring projects across the study area and beyond (https://www.wildtrax.ca/home). Some sites had ARUs deployed in multiple years, providing multiple years of time since disturbance, so we used ARU deployments (i.e., combinations of site and deployment year) as our unit for random selection. We used inventories of each disturbance type (Supplementary Materials) to randomly select 400 ARU deployments with a disturbance (fire, harvest, and well site) within 200 m. We used 200 m because we have previously determined this is the scale of effect for common nighthawk nesting (*unpublished data*) and it is also the approximate effective detection radius of our acoustic processing (Chapter 3; see below). During site selection, we prioritized disturbance type and age class combinations based on the strata with the fewest available sites so that we sampled as evenly as possible across a post-disturbance temporal gradient (Figure 6.1). We used the most recent event to determine time since disturbance if there were overlapping disturbances of the same type at a potential site. We only used sites for which there was one disturbance type within the 200 m radius. Across our 400 sites, we sampled 33 individual fires, 159 individual harvest blocks, and 95 individual well pads.

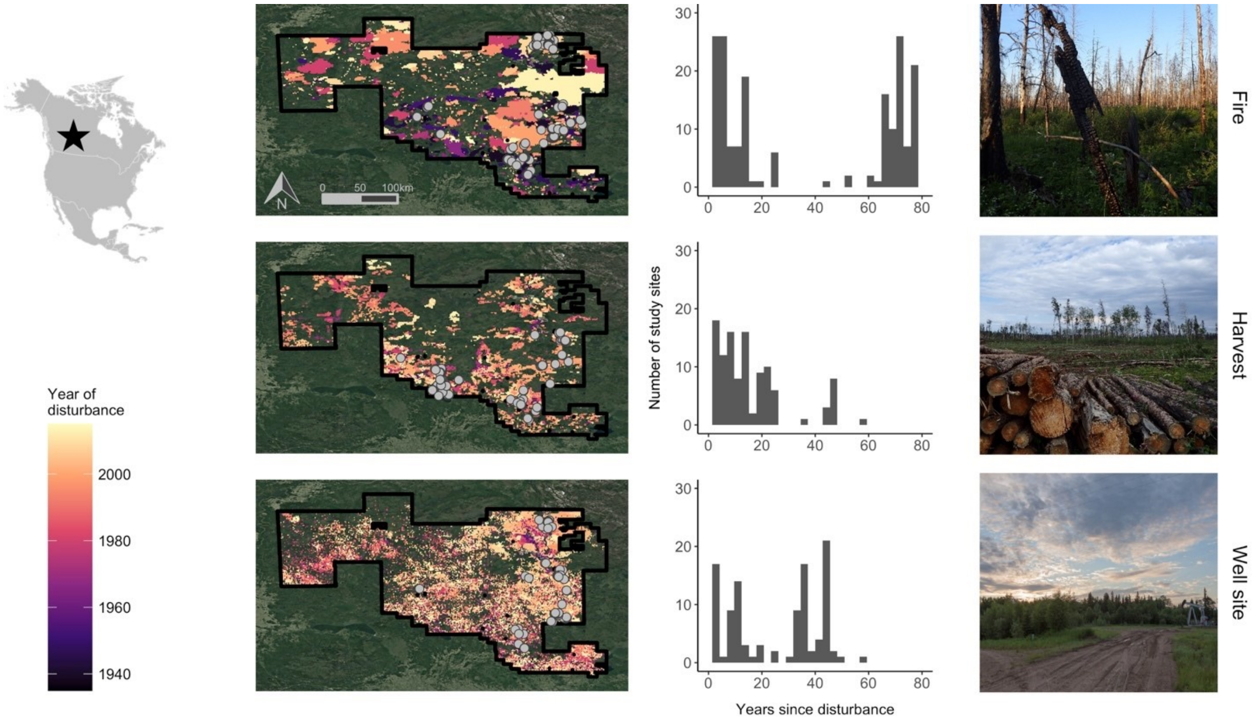


Figure 6.1. Variation in disturbance ages sampled with acoustic recorders (dots) at three disturbance types for common nighthawk habitat use in the boreal forest of northeastern Alberta, Canada.

Acoustic Data Selection

For each ARU deployment, we randomly selected up to 20 three-minute recordings (mean=8.4, SD=3.3, range=1-20). We used only recordings collected between June 1 and July 31 when common nighthawks are on their breeding territories in the study area (Knight et al. 2021a; Ng et al. 2018) and within a half hour before sunset to 1.5 hours after sunset, when the probability of detecting common nighthawks in the study area is maximized (Knight et al. 2021b, Hannah et al. In Review). All recordings were collected by ARUs with two omnidirectional microphones (ARUs; SM2, SM2+, and SM3; Wildlife Acoustics Inc.) at a sampling rate of 44.1 kHz with a 16-bit depth and following a standardized deployment protocol (Lankau 2015).

Acoustic Data Processing

We used Song Scope software (Wildlife Acoustics Inc.) to extract common nighthawk detections from all ARU recordings at each station. We used the vocalization for acoustic data processing because it is easier for a computer to detect and classify and because the wing-boom signal is always accompanied by a call. We trained a recognizer with 50 audio clips of common nighthawk vocalizations that were recorded at close-range (mean = 11.7 m; Chapter 3), as determined by a human observer (Yip et al. 2020). We used the recognizer to scan the recordings using a minimum score threshold of 60 and a minimum quality threshold of 20, which we have previously shown optimizes the trade-off between false positive and false negative detections, while limiting the effective detection radius of the recognizer to approximately 200 m (Chapter 2, Chapter 3). The first author visually validated all potential common nighthawk detections identified by Song Scope to confirm they were true positive detections. During validation, the first author also noted all detections that were accompanied by a mechanical wing-boom display.

Occupancy Covariates

For each disturbance type, we quantified the time since disturbance relative to the time the recording was collected. For well sites, time since disturbance was calculated as 1 year for active wells and the time since well abandonment (i.e., year when well is sealed, taken out of service, and the site is no longer maintained) for inactive wells. We did not use initial clearing date to calculate time since disturbance because well sites are maintained as vegetation free while active (i.e., continued disturbance). We did not quantify total disturbance size because there was a near perfect confound between disturbance size and disturbance type (mean ± standard deviation: well site 0.7 ± 0.4 ha, harvest 59.3 ± 92.5 ha, fire 64,197 ± 155,149 ha).

Where multiple disturbances of the same type overlapped, we assigned that area to the most recent disturbance. We did not include intensity of disturbance because the majority of fires in this region of the boreal forest are classified as burn class 5 (>94% tree kill; Supplementary Materials) and there is little to no residual vegetation remaining after forest harvest (< 2% residuals; Alpac 2015) and well site clearance (0%).

We also quantified two vegetation covariates within the 200 m radius that could mitigate the effects of time since disturbance on habitat use (Table 6.1). We defined pine forest as any areas that had jack pine (*Pinus banksia*) as the primary tree species and quantified the proportion of the 200 m radius that was pine forest at each site. Wetland area fluctuates with annual variation in precipitation, so we instead calculated the mean wetland probability within the 200 m radius for each site from the boreal wetland probability data set, which describes the probability of wetland habitat at 10 m resolution for Alberta's boreal forest (Supplementary Materials).

Detectability Covariates

We calculated the day of year and time relative to sunset for each recording to account for any phenological or circadian differences in availability for detection. We also quantified the signal-to-noise ratio (StN) and power spectrum density (PSD) of two frequency bands (0.6-1.2 kHz, 4.4-5.6 kHz) in each recording to quantify potential weather effects on perceptibility via sound masking or degradation. We chose these signal bands because they have been shown to be effective at classifying heavy rainfall, but also because they correspond to the two acoustic signals of the common nighthawk, the call and the mechanical wing-boom signal. We used the hardRain package, which uses these two measures to classify heavy rainfall (Metcalf et al., 2020). To account for potential changes in rainfall intensity during a recording, we quantified

StN and PSD for each thirty second interval and calculated the mean for each metric in each recording.

Statistical Analysis

We used an occupancy modeling framework to estimate the probability of common nighthawk habitat use using the unmarked package (Fiske 2011) in R version 4.0.3 (R Core Team 2020). We interpreted our results as probability of seasonal habitat use instead of occupancy because common nighthawks have large home ranges (~40 km^2 in the study area; *unpublished data*) and thus do not satisfy the closure assumption that requires they are present within the effective detection radius during the time of sampling (MacKenzie et al., 2002). We used each recording as a visit to build two sets of habitat models: 1) territorial models that used detection or non-detection of the territorial wing-boom signal as the response variable, and 2) extraterritorial models that used detection or non-detection of the more general peent call as the response variable. For the extraterritorial models, we used only detections at sites where there were no wing-booms detected because we were interested in how disturbance affects habitat use outside of the breeding territory (i.e., foraging, travelling, and roosting). For each set, we conducted our analysis in four stages. Prior to each analysis, we screened our predictor variables and removed any with a variance inflation factor (VIF) greater than 5 or covariation greater than 0.7.

First, we tested for effects of day of year, time relative to sunset, PSD and StN on common nighthawk detectability to determine which variables to include in subsequent analyses. For each model set (territorial, extraterritorial) we built a global model with all detectability covariates and compared it to all potential simpler models with AICc. We inspected all variables with generalized additive models (GAMs) and included StN values as second-order polynomials in the occupancy model based on these results. We excluded PSD for the higher frequency band (4.4-5.6 kHz) because it was highly correlated with PSD of the lower frequency band (corr = 0.88) and had the highest VIF value of the two PSD measurements (4.60 vs 4.57). We did not include any occupancy covariates in these models. We used small-sample Akaike's information criterion (AICc; Akaike 1987) to compare models and select the most parsimonious within ΔAICc < 2. For each model set, we included the covariates of the best-fitting model as detectability covariates in all subsequent analyses.

Second, we tested for effects of our vegetation covariates, proportion of pine forest and mean wetland probability, on common nighthawk habitat use to confirm their use in subsequent analyses. To prevent disproportionate influence of heavily sampled areas, we divided our study area into a 1 km grid and randomly selected one site for each grid cell from all available sites (Ploton et al., 2020; Robinson et al., 2017). For each model set, we built a global model that included the amount of pine forest and mean wetland probability and the interaction between the two. We included mean wetland probability as a second-order polynomial because common nighthawk probability of occurrence has previously been shown to be highest at intermediate values of wetland probability (Chapter 5). We bootstrapped this sampling and model fitting process 100 times and calculated the mean AICc value for each model across bootstraps. We selected the best fitting model for each model set as the most parsimonious model with mean ΔAICc < 2. For each model set, we added the vegetation covariates of the best fitting model as occupancy covariates in subsequent analyses.

Third, we investigated whether disturbance type affects common nighthawk habitat use, and if so, whether there is an interaction with time since disturbance. We again spatially thinned our dataset for this analysis by selecting up to one site for each 1 km grid cell for each disturbance type. For sites where there was more than one disturbance, we used the minimum time since disturbance as the response variable (0.85 correlation with mean time since disturbance). We also included an interaction between time since disturbance and the relevant vegetation covariates to assess our predictions about vegetation type mitigating negative effects of time since disturbance (Table 6.1). For each model set, we built a global model with both interactions and all additive effects and compared it to all simpler models that included the relevant vegetation covariate for that model set (pine for territorial models, wetland for extraterritorial models). We bootstrapped this sampling and model fitting process 1000 times. We selected the best fitting model for each model set as the most parsimonious model with mean ΔAICc < 2.

Results

We processed 4,523 3-minute ARU recordings from 326 sites and 400 ARU deployments (i.e., combinations of site and deployment year). The recognizer reported a possible 25,129 common nighthawk detections, 16,351 of which were true positive common nighthawk calls. 672 of those true positives were accompanied by a territorial wing-boom display. The territorial detections occurred at 44 of the sites, 116 of the ARU deployments and 460 of the ARU recordings, resulting in a predicted probability of detection of 0.20 (0.16-0.25 95% CI) when all detectability covariates were held at their mean. The remaining extraterritorial detections occurred at 50 of the sites, 56 of the ARU deployments and 497 of the ARU recordings, predicted probability of detection of 0.18 (0.15-0.22 95% CI) when all detectability covariates were held at their mean.

Detectability

The selected model for both territorial and extraterritorial habitat use included only the signal to noise ratio (StN) of the higher frequency band (4.4-5.6 kHz; Table 6.2, Figure 6.2). Time relative to sunset and day of year both had positive effects on territorial and extraterritorial habitat use, but were not included in the selected models, likely because we selected recordings for times and dates with high detectability.

Table 6.2. Occupancy model selection results for detectability of common nighthawk territorial and extraterritorial habitat use. Comparison included global model and all potential simpler models with time relative to sunset (time), day of year (day), power spectrum density (PSD) and signal to noise ratio (StN) of two frequency bands. Only the top 5 models are shown, sorted in descending order by mean model weight (Wt). Bold indicates the most parsimonious model within ΔAICc < 2 that was selected to include the covariates of in subsequent analyses. Psi ~ 1 for all models.

Model	K	AICc	ΔAICc	Wt
		Territorial habitat use		
$p \sim \mathrm{PSD1} + \mathrm{StN2}^2 + \mathrm{time}$	6	875.67	0.00	0.18
$p \sim \mathrm{StN2}^2 + \mathrm{time}$	5	876.05	0.38	0.15
$p \sim \mathrm{StN2}^2$	4	876.12	0.46	0.14
$p \sim \mathrm{PSD1} + \mathrm{StN2}^2$	5	876.31	0.64	0.13
$p \sim \mathrm{StN2}^2 + \mathrm{time} + \mathrm{day}$	6	877.57	1.90	0.07
		Extraterritorial habitat use		
$p \sim \mathrm{StN2}^2$	4	996.25	0.00	0.56
$p \sim \mathrm{StN1}^2 + \mathrm{StN2}^2$	6	996.59	4.03	0.07
$p \sim \mathrm{StN1}^2 + \mathrm{StN2}^2 + \mathrm{day}$	7	998.06	5.50	0.04
$p \sim \mathrm{StN2}^2 + \mathrm{time}$	5	998.13	5.57	0.03
$p \sim \mathrm{StN2}^2 + \mathrm{day}$	5	998.13	5.58	0.03

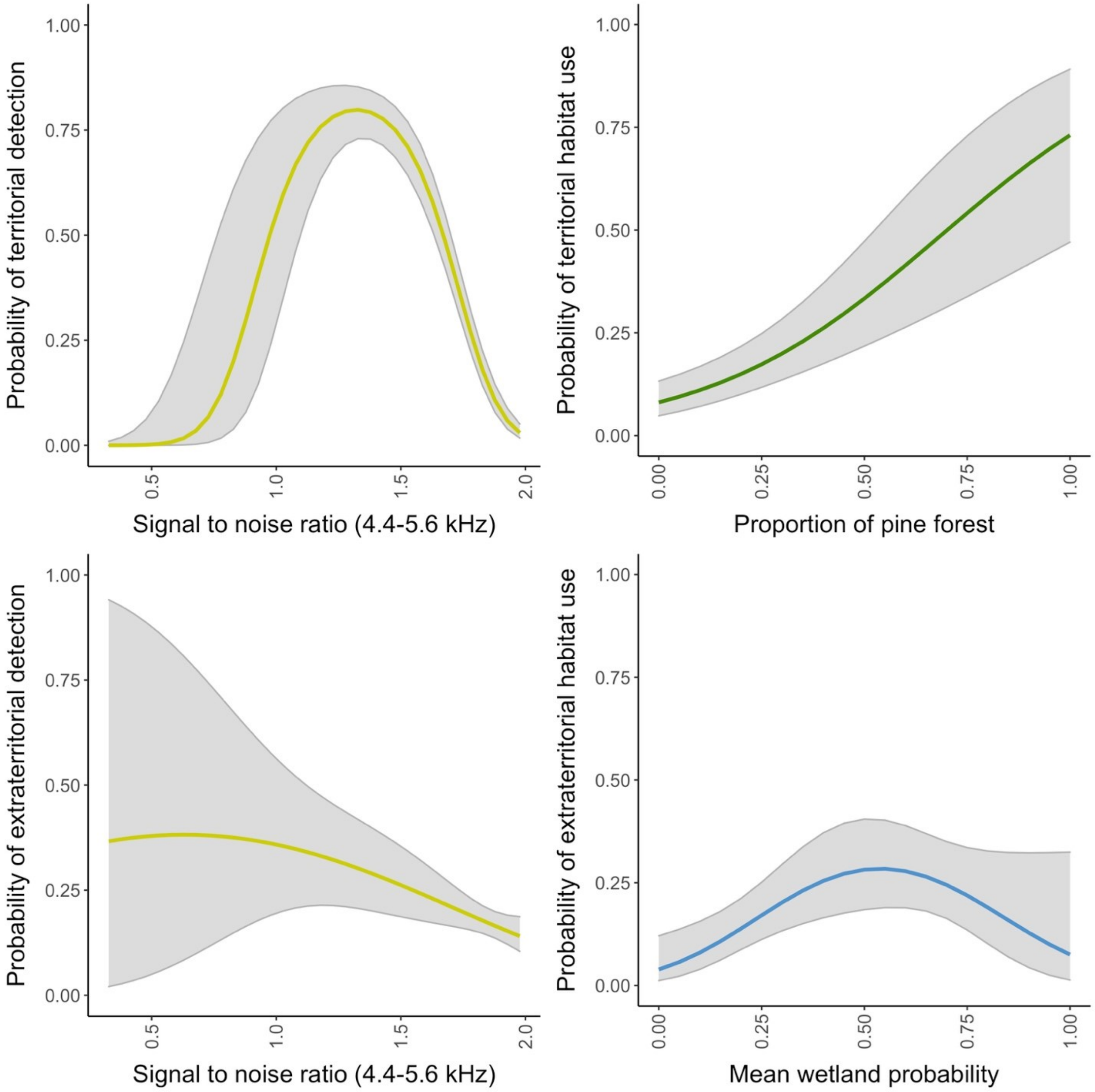


Figure 6.2. Predicted effects of detectability covariates on probability of detection (left) and of vegetation covariates on probability of habitat use (right) for territorial and extraterritorial common nighthawks. Lines and shaded areas represent predicted mean and 95% confidence intervals from occupancy models. Vegetation covariate predictions are the mean of 100 bootstraps used to spatially thin the available dataset.

Vegetation Effects

The selected model for territorial habitat use included only the proportion of pine forest within 200 m and the selected model for extraterritorial habitat use included only a quadratic effect of mean wetland probability within 200 m (Table 6.3, Figure 6.2). Across the 100 bootstraps, there was a mean 0.10 increase in probability of detection for every 10% increase in proportion of pine forest within 200 m. The highest probability of extraterritorial habitat use was at approximately 50% wetland probability.

Table 6.3. Occupancy model selection results for effects of vegetation on common nighthawk territorial and extraterritorial habitat use. Models included the proportion of pine within 200 m (pine) and a quadratic effect of the mean wetland probability within 200 m (wetland2; also includes the linear term). Data were spatially thinned to a 1 km grid and the analysis was bootstrapped 100 times. Results are the mean and standard deviation (SD) of AICc, ΔAICc, and model weight (Wt) across the 100 bootstraps. Bold indicates the most parsimonious model within mean ΔAICc < 2 that was selected to include the covariates of in subsequent analyses.

Model	K	AICc	ΔAICc	Wt
		Territorial habitat use		
psi ~ pine + wetland^2	7	494.97 (SD=48.42)	0.61 (SD=0.9)	0.43 (SD=0.17)
psi ~ pine	5	495.67 (SD=48.61)	1.3 (SD=1.36)	0.32 (SD=0.15)
psi ~ pine*wetland^2	9	496.1 (SD=48.55)	1.74 (SD=1.28)	0.26 (SD=0.15)
psi ~ wetland^2	6	513.8 (SD=52.61)	19.44 (SD=6.06)	0.00 (SD=0.00)
psi ~ 1	4	522.08 (SD=53.86)	27.72 (SD=6.69)	0.00 (SD=0.00)
		Extraterritorial habitat use		
psi ~ wetland^2	6	698.8 (SD=30.5)	0.00 (SD=0.00)	0.65 (SD=0.03)
psi ~ pine + wetland^2	7	700.69 (SD=30.47)	1.89 (SD=0.23)	0.25 (SD=0.02)
psi ~ pine*wetland^2	9	704.17 (SD=30.45)	5.38 (SD=0.63)	0.05 (SD=0.01)
psi ~ 1	4	705.23 (SD=30.6)	6.43 (SD=1.61)	0.03 (SD=0.02)
psi ~ pine	5	706.67 (SD=30.53)	7.87 (SD=1.56)	0.02 (SD=0.01)

Disturbance Effects

The selected model for territorial habitat use was the global model including time since disturbance and an interaction with both disturbance type and proportion of pine forest (Table

6.4). The probability of territorial habitat use declined with time since disturbance for all disturbance types but was mitigated by the proportion of pine forest present within 200 m. When there was a high proportion of pine forest within 200 m, postharvest and post-disturbance well sites had an almost perfect mean predicted probability of habitat use, while postfire areas had a slightly lower initial probability (0.75; Figure 6.3). In contrast, the predicted probability of habitat use at postfire sites with minimal pine forest was between 0.01 and 0.50 across the tree disturbance types. Probability of habitat use dropped steeply at postharvest sites, reaching 0 at approximately 20 years postharvest for sites with minimal pine forest and approximately 40 years for sites with high proportions of pine forest. Mean probability of habitat use dropped less steeply at post-disturbance well sites, and even more slowly at postfire sites with a high proportion of pine forest; however, the 95% confidence intervals were quite wide for all estimates. Interestingly, the probability of habitat use at postfire sites with minimal pine increased with time since disturbance, reaching approximately 0.21 at 77 years post-disturbance. We investigated the landscape context of the six sites with wing-boom detections that were driving this relationship and found that all of them were in lowland areas interspersed with islands of sparse pine forest and that all had areas of pine forest just outside our 200 m buffer radius.

There was minimal evidence that common nighthawk extraterritorial habitat use outside the territory is influenced by disturbance; the selected model for extraterritorial habitat use included only a quadratic effect of mean wetland probability (Table 6.4).

Table 6.4. Occupancy model selection results for effects of disturbance type (type) and time since disturbance (time) on common nighthawk territorial and extraterritorial habitat use. Territorial models also included the proportion of pine within 200 m (pine) and extraterritorial models also included the mean wetland probability within 200 m (wetland). Data were spatially thinned to a 1 km grid and the analysis was bootstrapped 100 times. Results are the mean and standard deviation (SD) of AICc, ΔAICc, and model weight (Wt) across the 100 bootstraps. Bold indicates the most parsimonious model within mean ΔAICc < 2 that was selected to include the covariates of in subsequent analyses.

Model	K	AICc	ΔAICc	Wt
Territorial habitat use				
psi ~ time*type + time*pine	11	421.02 (SD=21.13)	0.22 (SD=0.51)	0.52 (SD=0.23)
psi ~ time*type + pine	10	423.09 (SD=21.16)	2.29 (SD=2.48)	0.23 (SD=0.14)
psi ~ pine	5	424.39 (SD=20.51)	3.59 (SD=2.4)	0.12 (SD=0.09)
psi ~ time + pine	6	426.35 (SD=20.5)	5.55 (SD=2.4)	0.05 (SD=0.03)
psi ~ time*pine	7	427.51 (SD=20.31)	6.71 (SD=2.25)	0.03 (SD=0.03)
Extraterritorial habitat use				
psi ~ wetland^2	6	597.12 (SD=33.36)	0.03 (SD=0.14)	0.43 (SD=0.09)
psi ~ time + wetland^2	7	598.68 (SD=33.46)	1.58 (SD=0.47)	0.20 (SD=0.04)
psi ~ type + wetland^2	8	599.28 (SD=33.36)	2.19 (SD=1.06)	0.15 (SD=0.06)
psi ~ time*wetland^2	9	599.82 (SD=33.52)	2.73 (SD=1.17)	0.12 (SD=0.06)
psi ~ time + type + wetland^2	9	601.27 (SD=33.41)	4.17 (SD=1.06)	0.06 (SD=0.02)

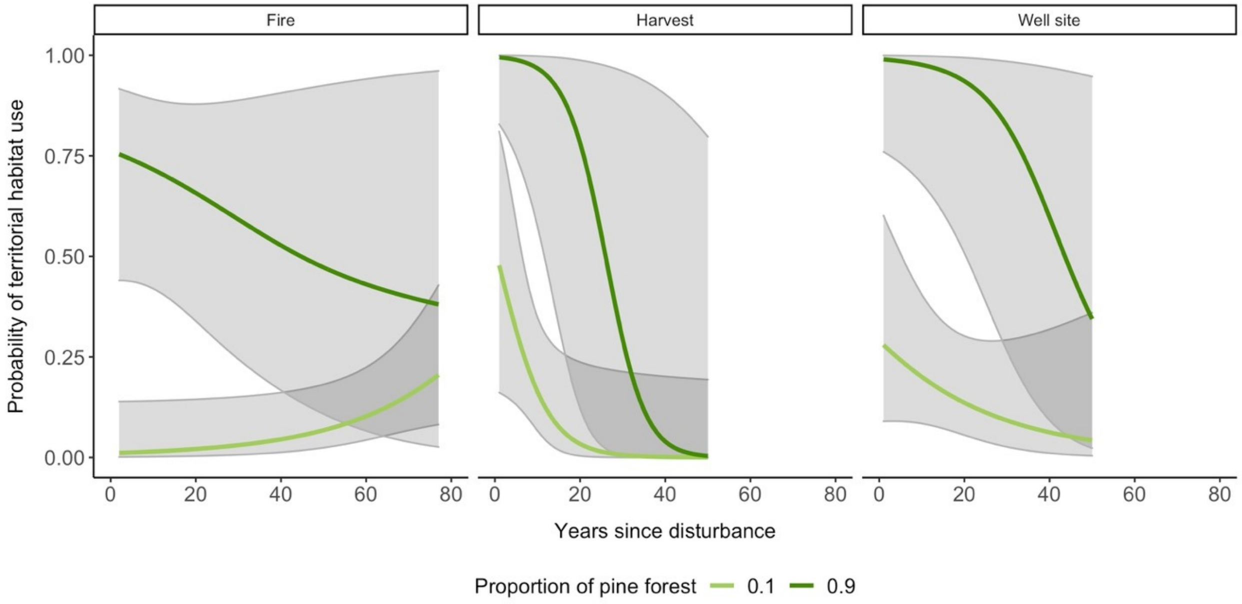


Figure 6.3. Predicted effects of disturbance and proportion of pine forest on probability of territorial habitat use (right) for common nighthawks. Lines and shaded areas represent predicted mean and 95% confidence intervals from occupancy models. Predictions are the mean of 100 bootstraps used to spatially thin the available dataset.

Discussion

To inform boreal forest management, we modelled different types of habitat use to test two competing hypotheses for why common nighthawks use post-disturbance areas in the boreal forest: nesting or foraging resources. We used occupancy models of territorial and extraterritorial detections of common nighthawks from passive acoustic monitoring to test three lines of evidence for our two hypotheses (Table 6.1). Across all three lines of evidence, we found support for the nesting resource hypothesis. For our first line of evidence, time since disturbance negatively affected territorial habitat use, but not extraterritorial habitat use, confirming that this species is only a disturbance specialist for the territorial component of its home range. For our second line of evidence, territorial habitat use differed between disturbance types, with higher probability of habitat use for harvest and well site disturbances immediately after disturbance. For our third line of evidence, pine forest mitigated the effects of time since disturbance for territorial habitat use for all three disturbance types, but especially for postfire areas. Extraterritorial habitat use was affected only by mean wetland probability. Although common nighthawks almost certainly also use post-disturbance areas for foraging resources as well, we conclude that the relative importance of nesting resources is the primary driver of post-disturbance habitat use. We discuss each line of evidence below and the management implications of our results.

Time Since Disturbance

Successional stage affects the abundance and distribution of many boreal bird species (Hobson and Schieck, 1999; Schieck and Song, 2006). Whether the effects of succession are because birds are adapted to disturbance itself or to the resultant vegetation structure is unclear (Bunnell 1995, Thompson 2008). Regardless, we confirmed that the highest probability of common nighthawk territorial habitat use is immediately after disturbance. We conclude this effect of time since disturbance is due to vegetation succession that renders sites unsuitable for laying eggs on bare ground for nesting, as opposed to change in the foraging resources because we did not find any disturbance effects for extraterritorial habitat use and common nighthawks do not restrict their foraging to the territory (Brigham et al., 2011). Hagar et al. (2004) concluded common nighthawks only occupied thinned forest stands and not intact forest due to increased nesting opportunities and foraging opportunities provided by a more open canopy. Farrell et al. (2017) previously found no difference between harvest, fire, and wetland sites; however, they also examined habitat use at the home range scale and thus were unable to disentangle the effects of wetlands on extraterritorial habitat use and disturbance on territorial habitat use.

Although our time since disturbance results confirm common nighthawks track disturbance across the landscape, the mechanism through which this occurs is less clear. Do common nighthawks find and colonize new disturbances during natal or adult dispersal? In other words, does the breeding habitat at post-disturbance sites become unsuitable as succession proceeds and breeding adults relocate elsewhere, or does the population at a given site go extinct as the breeding adults die off? In landscapes with relatively high disturbance rates, like the boreal forest, specialists are expected to have high dispersal abilities (Büchi & Vuilleumier 2016; Johst et al. 2002) and the extreme mobility of the common nighthawk corroborates this dispersal ability. Territorial fidelity is high in this species (Gross 1940, Dexter 1961, Ng 2018, Chapter 4) pointing towards natal dispersal as the mechanism driving metapopulation dynamics. Within a population, the abundance of disturbance specialists that rely on natal dispersal to find suitable habitat is expected to track the same trajectory as the survival curve for that species (Southwood 1977); however, the lifespan of the common nighthawk is unclear, so we are unable to evaluate this hypothesis (Brigham et al., 2011). Future research to inform common nighthawk management should use mark-recapture techniques to disentangle the relative roles of survivorship and natal and adult dispersal on the metapopulation dynamics of common nighthawks (Driscoll et al. 2010).

Disturbance Type

Differences between disturbance types in the rate at which probability of habitat use declines with time since disturbance suggests adult dispersal is at least partially responsible for the decline in territorial habitat use. Succession rate likely differs between disturbance types, which results in adults dispersing to new breeding territories at different rates. Succession at well sites is slower than at forest harvest sites (Osko 2001), and we found the rate of decline of common nighthawk habitat use reflected this difference in succession. Succession at postfire sites is also slower that at forest harvest sites because fire kills the understory layer, while harvest either simply tramples or in some cases retains it (Seedre et al. 2011), and we also found the rate of decline of habitat use reflected this difference. Differences in rate of decline of habitat use between disturbance types may also be driven by the vegetation community, which regenerates at different rates. Harvested areas are typically in mixedwood or trembling aspen (*Populus tremuloides*) stands and occur during winter (Chen & Popadiouk 2002), which regenerate fastest, while fires and oil well sites can occur in any vegetation type and fires can also be a range of severities.

We also found differences in the initial probability of territorial habitat use between disturbance types. As per our predictions for the nesting resource hypothesis, we suggest differences in initial probability are likely driven by differing effects on the amount of residual vegetation immediately following disturbance (Bognounou et al. 2021). If foraging resources were the driver of post-disturbance habitat use, then we would expect to see the highest initial probability of habitat use in postfire areas because they a) have higher availability of wood-boring beetle prey in the first decade after disturbance, which common nighthawks are known to consume (Knight et al. 2018), and b) postfire areas likely have higher aerial insect availability due to higher amounts of residual vegetation (Deans et al. 2005; Morissette et al. 2002). Instead, we saw the lowest initial probability of habitat use at postfire areas, likely due to those higher amounts of downed woody debris and residual vegetation (McRae et al. 2001) that reduce nesting substrate availability. We note, however, that there is substantial variability in downed and residual vegetation across postfire and postharvest sites (Lindenmayer et al. 2019), which may be in part responsible for the large confidence intervals around our estimates. This is in direct contrast to many other disturbance-specialist species, which are attracted to post-disturbance sites specifically for the wood-boring beetle prey and cavity nesting opportunities (Saab et al. 2007; Tingley et al. 2020).

There are also potential confounds that could contribute to differences between disturbance type. Disturbance size is directly confounded with disturbance type in the boreal forest, and so we were unable to examine its effects here. As per island biogeography theory, fire size is expected to be positively related to occupancy and abundance (Turner et al. 1997). Black-backed woodpeckers show the opposite relationship because they are limited by dispersal distance (Tingley et al. 2018). Given that common nighthawks are unlikely to be dispersal limited, we would expect postfire areas to have the highest probability of habitat use because they are the largest disturbance type; however, this is opposite of our results.

Vegetation

Pine forest is the least productive stand type in the boreal forest because it typically grows in sandy soil with low organic content. The positive effect of proportion of pine forest on common nighthawk territorial habitat use was stronger than any of the other covariates we examined and pine forest mitigated the effects of time since disturbance by increasing the initial probability of habitat use and decreasing the rate of decline of habitat use. This effect of pine forest provides strong support for the nest resource hypothesis because if common nighthawks were attracted to disturbance for the foraging resources, we would expect them to avoid pine forest due to the relatively low availability of aerial insects. In fact, this propensity for pine forest suggests that common nighthawks are not necessarily disturbance specialists, but rather “bare ground specialists”, at least on territory. This label jibes with common nighthawk nesting habitat in other biomes, including coastal sand dunes and beaches, rocky outcrops, sagebrush and grassland habitat, and flat gravel rooftops (Brigham et al., 2011). Across biomes, common nighthawk select nest sites with high amounts of bare ground (Allen & Peters 2012; Lohnes 2010).

Outside of the territory, mean wetland probability was the only covariate that affected habitat use. Areas of open water are important for breeding common nighthawks in southern landscapes because large flocks of individuals are frequently seen foraging for emergent insects over lakes and rivers in southern ecosystems (Brigham et al., 2011) and presence-absence modelling in southern Saskatchewan found that common nighthawks use home ranges closer to areas of open water (Ng, 2009). Although the distribution of aerial insect availability in the boreal forest is more homogenous than in grasslands, wetland areas may still provide richer foraging resources. Probability of extraterritorial use was highest at intermediate levels of mean wetland probability, suggesting that common nighthawks are using wetland edges for foraging, as opposed to over the middle of large wetland complexes. Maximal use at wetland edges may be because the vegetation zonation from emergent to submergent community supports higher species richness and abundance of aquatic emergent insects (Stagliano et al. 1998, Bush and Wissinger 2016).

Management Implications

The conclusion that the common nighthawk is a disturbance specialist or “bare ground specialist” bodes well for the persistence of this species on the boreal landscape. In the northwestern boreal, where anthropogenic disturbance from forestry and oil and gas extraction is the dominant disturbance (Pickell et al. 2014), we found no detrimental differences in common nighthawk habitat use between natural and anthropogenic disturbance types. We also suggest increases in the frequency and severity of fires due to climate change (de Groot et al. 2013) will likely benefit nighthawks due to the resultant greater availability of bare ground for nesting. We caution that although we found no negative differences in disturbance effects of anthropogenic disturbance types on common nighthawk habitat use, that there may still be differences in vital rates (e.g., nest success, juvenile survivorship) that reflect differences in habitat quality between disturbance types (Johnson 2007). For example, postharvest areas can be ecological traps for Olive-sided Flycatchers (*Contopus cooperi*) because they preferentially select postharvest areas but have lower reproductive success than in postfire areas (Robertson & Hutto 2007). We also note that our study only examined disturbance effects at the local scale, but the spatial context of disturbance on the landscape (i.e., disturbance regime) can also affect habitat use (Turner 2010); thus, it remains unknown whether fire, forestry, and oil and gas disturbance regimes have differing effects on common nighthawk habitat use. The disturbance regime affects the process of colonization for metapopulation-structured species (Amarasekare & Possingham 2001) and the oil and gas disturbance regime is a completely novel regime on the boreal landscape (Pickell et al. 2013); therefore, future research should use a cumulative effects framework to examine if the novel disturbance regime of the northwestern boreal forests drives common nighthawk metapopulation dynamics.

Although we found no negative differences in common nighthawk habitat use between disturbance types, care must be taken in disturbed areas to avoid undue harm to disturbance specialist species that are already using that habitat. Given that common nighthawks use recently disturbed areas as ground-nesting habitat, post-disturbance activities such as chain-dragging, furrow trenching, and replanting cutblocks, clearing and maintenance of wells and lines at well sites, retention logging of burned areas, and vehicle access to any disturbed area all have the potential to disturb nesting adults and/or destroy their nests. These activities should be done outside the breeding season whenever possible to avoid disturbing nesting common nighthawks, particularly in Canada where this species is protected under the *Migratory Bird Conservation Act SC 1994* and the *Species at Risk Act SC 2002*. We note that the common nighthawk breeding season extends later than most songbirds in the boreal forest, with the potential for eggs from the beginning of June through mid-August and the potential for flightless young possible through late August (*unpublished data*). If activities cannot be scheduled outside this breeding season, pre-activity surveys should be conducted to evaluate the potential for nesting common nighthawks. We showed here that common nighthawks only use disturbed areas for territorial purposes, and so not all common nighthawk detections should be of concern. Surveys should focus on the presence of the wing-boom signal, which is an indication of the nest location (Chapter 4). These recommendations should be applied to other managed disturbance types including transmission lines and open-pit mines.

Conclusions

We conclude by suggesting that understanding the behavioural context of habitat use is important for effective management. The operational scale of forest management (e.g., stand size) is more likely to be aligned with the scale of habitat components than that of the home range for highly mobile species like the common nighthawk. Examining different types of habitat use not only provided insight into our contrasting nest and food resource hypotheses, but also revealed that the 'disturbance specialist' label is context-dependent: Common nighthawks are not disturbance specialists outside of the territory. Behaviour-specific habitat studies have similarly revealed attributes of function-specific habitat components like encampment behaviour of elephants (Roever et al. 2013) and prey capture locations of burrowing owls (*Atene cunicularia*; Marsh et al. 2014) that would otherwise have been masked with home range analyses. Behavioural context of habitat use is particularly important for highly-mobile species because they often use spatially distinct habitats to fulfill their life-history needs as opposed to multitasking within a single geographic area (Frans et al. 2017; Law & Dickman 1998; Roever et al. 2013). Furthermore, ignoring habitat function can lead to erroneous conclusions about habitat attributes because function-specific habitats can have opposing attributes that nullify, change the strength of, or distort the signal of habitat use (Roever et al. 2013). We advocate that understanding the behavioural context of habitat analyses can help inform management of other species with large home ranges.

Supplementary Materials

Methods for quantification of occupancy covariates from remotely sensed data.

Disturbance covariate data sources: We used a different data source to quantify each disturbance type (fire, harvest, wells). Historical wildfire perimeters were sourced from the Alberta Ministry of Agriculture and Forestry (2018; Table 6.A.1). Harvest polygons was sourced from the Alberta Biodiversity Monitoring Institute (ABMI) wall-to-wall human footprint inventory (2016; Table 6.A.1) using the feature type "HARVEST-AREA". Well polygons were also sourced from the ABMI wall-to-wall human footprint inventory using both active wells (WELL-BIT, WELL-GAS, WELL-CASED, WELL-OTHER, WELL-OIL feature types) and abandoned wells (WELL-ABAND feature type). We assigned time since disturbance as 1 year for active wells and used the year of disturbance field for abandoned wells, which identifies the year that well was last drilled. We then converted all three disturbance type layers to raster to prioritize the most recent disturbance where multiple disturbances of the same type overlapped. We resampled each raster layer to 30 m resolution to standardize it across datasets. We then converted the raster layers back to polygons for further processing with buffers (see *Covariate quantification* below).

Vegetation covariate data sources: We also sourced each vegetation covariate (pine, wetland) from separate data sources. We used the Alberta vegetation inventory (AVI) from the Alberta Government (2013; Table 6.A.1) to quantify pine forest habitat by identifying all polygons for which the primary dominant tree type was jack pine (*Pinus banksia*). We used the

ABMI boreal wetland probability (Hird et al. 2017; Table 6.A.1) dataset to quantify wetland habitat. We also resampled both vegetation layers to 30 m cell size to standardize resolution.

Covariate quantification: We quantified all three of our occupancy covariates (time since disturbance, proportion of pine forest, mean wetland probability) for a 200m radius buffer around each site where an ARU was deployed. For time since disturbance, we determined whether or not there was a disturbance (i.e., 1 or 0) within the 200 m buffer for each of the three disturbance types by clipping each disturbance polygon layer with a layer of 200 m buffers for each site. For pine forest, we quantified the proportion of the 200 m buffer that was pine forest by clipping the pine polygon layer with the 200 m buffer layer. For wetland probability, we calculated the mean wetland probability within the 200 m buffer.

Table 6.A.1. Source geospatial datasets used to extract environmental variables for common nighthawk habitat modelling.

Dataset name	Description	Proprietor	Original data format (cell size)	Variables extracted	Available from (access date)	Available reference or metadata
Historical Wildfire Perimeters	Polygon features representing wildfire perimeters dating back to 1931. The datasets from 1998 on should include most of the fires 12 hectares (29.7 acres) or larger that occurred within the legislated Forest Protection Area. The datasets prior to these years may include some fires smaller than 200 hectares (494.2 acres), but the emphasis on the data collection for these years was to provide Class E fires (wildfire larger than 200 hectares or 494.2 acres).	Alberta Agriculture and Forestry	Vector	Time since fire	https://wildfire.alberta.ca/resources/historical-data/spatial-wildfire-data.aspx (March 5, 2019)	https://wildfire.alberta.ca/resources/historical-data/default.aspx
Human Footprint Inventory	Twenty one human footprint categories (based on more than 115 anthropogenic disturbance types) collated into a single integrated dataset by applying a specific order of precedence. The dataset was developed in part by a collaborative effort of data accumulation between the Government of Alberta and the ABMI, together forming the Alberta Human Footprint Monitoring Program (AHFMP) to better enhance and update human footprint data in Alberta.	Alberta Biodiversity Monitoring Institute	Vector	Time since harvest, time since well site disturbance	https://www.abmi.ca/home/data-analytics/da-top/da-product-overview/Human-Footprint-Products/HF-inventory.html (September 9, 2019)	https://ftp-public.abmi.ca/GISData/HumanFootprint/2016/HFI2016_Metadata.pdf

Dataset name	Description	Proprietor	Original data format (cell size)	Variables extracted	Available from (access date)	Available reference or metadata
Alberta Vegetation Inventory	This photo-based digital inventory dataset is a compilation of the original blocks of Crown Alberta Vegetation Inventory data collected using aerial photography dating between 1987 and 2014. A variety of standards was used to collect the original blocks of data. It was developed to identify the type, extent and conditions of vegetation in the province.	Alberta Agriculture and Forestry	Vector	Proportion of pine	https://geodiscover.alberta.ca/geoportal/rest/metadata/item/100b275712b442acbda4a0358d8a4951/html (Feb 12, 2013)	https://www.alberta.ca/vegetation-inventory-standards.aspx
Boreal Wetland Probability	The Boreal Wetland probability dataset was generated using a machine learning framework in R statistical software. The machine learning model was based off of a Digital Elevation Model, optical satellite data, and Synthetic Aperture Radar data. The model was trained with photo-plots that were derived from high resolution 3D image interpretation and give detailed attribution of land cover information.	Alberta Biodiversity Monitoring Institute	Raster (10 m)	Mean wetland probability	https://abmi.ca/home/data-analytics/da-top/da-product-overview/Advanced-Landcover-Prediction-and-Habitat-Assessment--ALPHA--Products/Boreal-Wetland-Probability-Data.html (September 10, 2019)	Hird et al. 2017

Chapter 7. Deriving density estimates from single-point recordings processed with automated recognition

Abstract

Density and population estimation is at the core of wildlife management, particularly at the regional, national, or species range scale. Passive acoustic monitoring and automated recognition have the potential to greatly facilitate population estimation via efficient collection of multi-visit data; however, existing approaches to estimate density from data produced by recognizers are limited and not particularly well-suited for application at large geographic scales. We developed a five-step density estimation approach for recognizer data that is based on the theoretical cascade of processes that occur from habitat suitability through to species detection. First, we used time to detection from the occurrence dataset produced by recognizers to estimate availability for detection. Second, we used conditional likelihood zero-inflated models to estimate occupancy of suitable sites from a dataset of abundance at a subset of our acoustic recordings. Third, we use the results from the previous two steps as inputs for a modified occupancy model that estimates habitat suitability and the probability of species presence during survey (i.e., closure). Fourth, we used a known distance dataset to estimate the effective survey area of our recognizer. Fifth, we combined the parameters from the previous approaches to estimate and predict density and territory size across our study sites and areas. We used our approach to estimate density of territorial male common nighthawks in four study areas in Canada's western boreal forest: two in recently burned areas ("wildfire") and two in forests of mid to late seral stages ("multi-seral"). We used two pieces of information to validate our approach. First, density was much higher in wildfire areas (0.038 and 0.030 males/ha) than multi-seral areas (0.000 and 0.005 males/ha), which is congruent with the “disturbance-specialist” habitat preferences of territorial common nighthawks. Second, our territory size estimates for one of the wildfire areas (9.7 ha/male) was nearly identical to estimates from VHF-tagged birds for the same area (10.2 ha/male). Our novel approach to density estimation is flexible, can be applied at large scales, incorporates habitat effects, and is suitable for mobile species that do not necessarily satisfy the closure assumption of multi-visit modelling. We are confident it will be an excellent tool in the toolbox of density estimation, particularly for species assessment and recovery, which occur at large geographic scales.

Introduction

Across the globe, species density and population size estimates are a fundamental component of wildlife management and conservation. Population size estimates are used in addition to trend estimates as criteria for assessing species conservation status across agencies and jurisdictions, usually following criteria set out by the International Union for the Conservation of Nature (IUCN 2012). Population size is also often used as part of the recovery objectives for species of conservation concern (National Recovery Working Group 2005). Finally, spatial variation in population density needs to be quantified for prioritizing landscapes with higher conservation values (e.g., the abundance center hypothesis) (Brown 1984). For example, information on spatial variation in density and abundance helps inform Critical Habitat Identification for species listed as Threatened in Canada (Government of Canada 2016). Density estimation as a means of estimating population size and providing a more nuanced understanding of habitat use (vs. occupancy models) is thus an important tool for management of bird populations.

A variety of density estimation approaches are available due to the importance of this metric for wildlife management; however, existing methods that provide reliable estimates are time and/or cost intensive and thus limit their application at the national or range-wide scales that are often used for wildlife management. For example, distance sampling requires highly trained observers to reliably estimate distance of vocalizing animals (Buckland et al. 2005, 2015), repeated count approaches require multiple visits by trained observers (MacKenzie et al. 2002, 2006), and mark-recapture requires intensive trapping and handling effort of wild animals (Pollock et al. 1990). In contrast, methods like the Partners in Flight approach that use existing community science datasets like the North American Breeding Bird Survey (BBS) are inexpensive to estimate population size at national and range-wide scales (Will et al. 2020) but are known to be imprecise in various situations (Thogmartin et al. 2006; Sòlymos et al. 2020; Thogmartin 2010). Furthermore, these wide-ranging approaches are ineffective for species that are rare, irruptive, colonial or semi-colonial, and vocalize rarely or at different times of day and thus are not well monitored by traditional dawn point count surveys like the BBS (Matsuoka et al. 2014; Rosenberg et al. 2017).

Passive acoustic monitoring (PAM) offers an alternative approach to data collection for wildlife monitoring and research at large geographic scales. The number of studies that use PAM for all purposes, including density estimation (Pérez-Granados & Traba 2021), has increased over the last two decades (Shonfield & Bayne 2017; Gibb et al. 2018; Sugai et al. 2018). Autonomous recording units (ARUs) are cost and time effective because they allow for multiple visits, are better for detecting rare species, and can be scheduled to record those species that are not well surveyed by dawn recordings.

One of the drawbacks to PAM is the time and skill required to process the acoustic recordings into detections of species (Shonfield & Bayne 2017; Gibb et al. 2018; Sugai et al. 2018). In response, a variety of computer algorithms have been developed to automatically process acoustic recordings (Priyadarshani et al. 2018; Stowell et al. 2018). Acoustic indices quantify and summarize recordings into single metrics of the amount of sound in each one-minute interval (i.e., the soundscape), which can be used to quantify disturbance and biodiversity (Lawrence et al. 2019). Several authors have attempted to use acoustic indices to estimate density, with limited success (Arneill et al. 2020; Orben et al. 2019). The alternative is automated recognition, where the computer scans acoustic recordings and assigns classification probabilities to detected sounds for the species (or multiple species) it has been trained to classify (Gibb et al. 2018; Priyadarshani et al. 2018). The hurdle with automated recognition is that it produces occurrence rather than abundance data because the algorithm is trained at the species level and individuals are indistinguishable to the computer (Priyadarshani et al. 2018).

Existing methods for density estimation from recognizer data therefore rely on supplementary information sources. The first type of model uses vocal activity rates as a proxy for abundance in a distance sampling framework (Sebastián-González et al. 2018; Pérez-Granados & Traba 2021). Unfortunately, the variation in vocal activity rate across time, space, and ecological conditions limits the use of these models to the context under which the vocal activity rate data was collected and they are unlikely to be reliable if implemented at broad geographic scales (Pérez-Granados et al. 2019; 2021). The other type of density model for automated recognition data uses paired human point counts for abundance information in a multi-visit n-mixture framework (Doser et al. 2021). These models are thus not full passive acoustic monitoring approaches; they still require data collection by skilled observers.

Furthermore, they assume population closure between visits (i.e., recordings) to compute density, however, species with large home ranges are sure to move in and out of the survey area between visits (Rota et al. 2009). A third group of density model exists, which uses arrays of ARUs (Marques et al. 2012; Stevenson et al. 2015); however, the resources required to deploy arrays of ARUs as opposed to single point records also renders this approach costly at large geographic scales.

Our goal was to develop a density estimation approach that could be applied to data from single point ARUs processed with recognizers and implemented at large geographic scales. Our model is based on a theoretical cascade of processes that determine survey outcomes, starting with site suitability and ending with detection (Figure 7.1). The model uses four pieces of information from three data sources. 1) The time to first detection in each recording from occurrence data was used to estimate probability of common nighthawk activity. 2) The detection and non-detection information across multiple recordings (i.e., visits) from the occurrence data was used to estimate probability of availability, 3) the abundance data from recordings with known occurrence was used to estimate mean abundance for each study area, and 4) the known distance data was used to estimate perceptibility. The model itself is an adaptation of single season single species occupancy modelling (MacKenzie et al. 2002) that uses inputs from survival modelling and zero-inflated modelling to account for a) multiple individuals at a location (i.e., abundance), and b) deviations from the closure assumption. We chose this approach because passive acoustic monitoring, particularly when data is processed by recognizers, is well-suited for hierarchical occupancy-style modelling because it creates multi-visit datasets.

We estimated density of the common nighthawk (*Chordeiles minor*) at four study areas in the boreal forest as a case study for our method. Reliable density estimates do not exist for this species because it is nocturnal and thus poorly surveyed by existing point count programs (Knight et al. 2021b) and is a highly mobile aerial insectivorous species that occupies large home ranges (at least 40 km^2; *unpublished data).* Generating meaningful population size estimates for the common nighthawk is a conservation priority because it is listed as Threatened under the Canadian Species at Risk Act (Environment Canada 2016a) and management for recovery and future assessment requires setting population objectives. We collected ARU recordings in Canada's western boreal forest and derived detections of territorial males using automated recognition. We used the mechanical wing-boom display of the common nighthawk to identify territorial males within recordings (Chapter 4). We estimated density of territorial male common nighthawks for four study areas: two that were recently burned by wildfire and two with a mosaic of mid and late seral stages ("multi-seral"). We used these study areas because the common nighthawk is a 'disturbance specialist' in the boreal forest Chapter 6; Sidler 2017; Farrell et al. 2017, 2019; Foley 2018), and thus there are reliable predictions about higher density in burned areas that we used to rationalize which study areas would be expected to have higher density. Territory size estimates are also available for one of our burned study areas (Chapter 4), which we also used to validate our estimates.

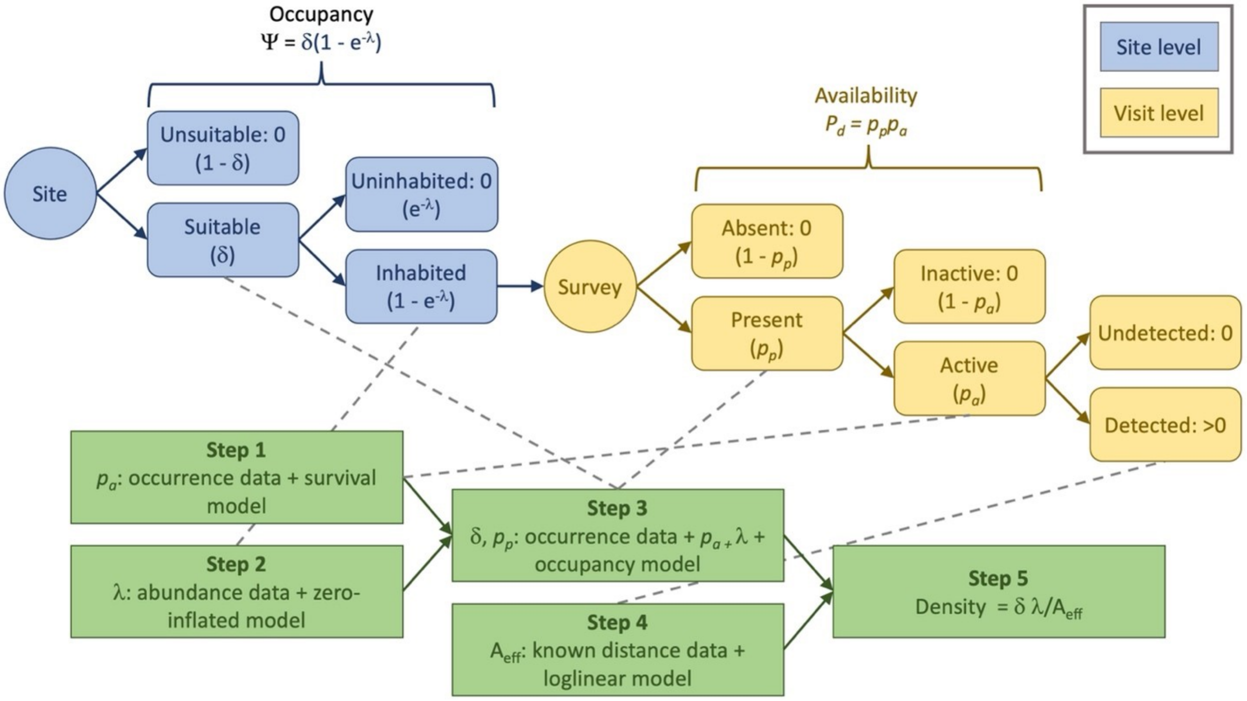


Figure 7.1. Overview of modelling framework for estimating density from passive acoustic monitoring data. The top diagram represents the theoretical cascade of processes that occur between survey site selection and species detection. The bottom diagram represents the five steps of the modelling process. The dashed lines indicate connections between the process and the step of the modelling process that estimates the probability of that process.

Methods

Study Area

We selected four study areas in Canada's western boreal forest for common nighthawk density estimation: two in large, recent (< 10 years) wildfires and two in areas with a mix of mid and late seral stages (multi-seral; Figure 7.2). One of each study area type was in the southern portion of the western boreal forest in northeastern Alberta, Canada where land use pressure for resource extraction is high, and one of each type was in the relatively undisturbed northern portion in the Northwest Territories, Canada. We chose these two regions to compare density estimates at different latitudes and different levels of anthropogenic disturbance. The vegetation (pre-fire for the wildfire sites) was a typical mix of coniferous and deciduous upland forests and lowland peatlands, with varying amounts of jack pine (*Pinus banksia*), which has previously been associated with common nighthawk habitat use (Chapter 5, 6). We chose the perimeter for the wildfire study areas as the footprint of the fire that our sampling locations spanned. We chose the perimeter for the north multi-seral area as the perimeter of the Edéhzhíe Protected Area that contained our sampling locations. We determined the perimeter for the south multi-seral area by placing a 5 km buffer around the sampling locations for that region. Known mean territory size of territorial male common nighthawks was 10.2 ha (SD = 11.7 ha; Chapter 4) in the south wildfire study area.

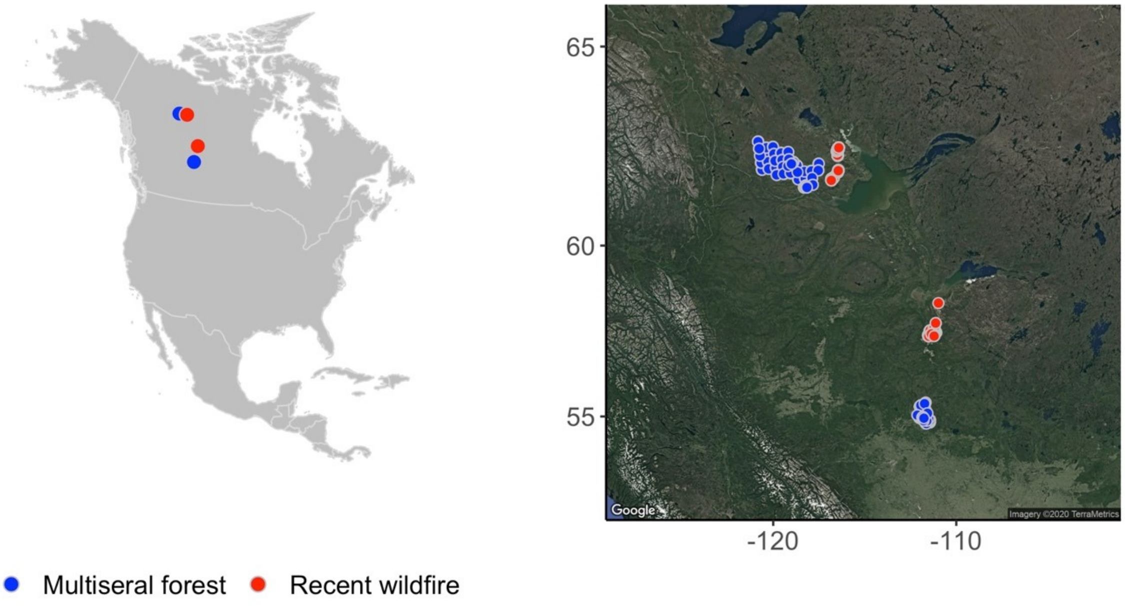


Figure 7.2. Study areas (left) and individual sampling locations (right) for density estimation of common nighthawks from passive acoustic monitoring data. Two study areas were in recent (< 10 years) wildfires and two were in multi-seral landscapes.

Acoustic Recording Collection

We randomly selected sampling locations for each of our four study areas from an archived dataset of acoustic recordings collected between 2012 – 2016 (https://www.wildtrax.ca/home). Within each study area, we randomly selected at least 104 sampling locations, which was the minimum available at our smallest study area. We only selected locations that were at least 500 m apart to ensure common nighthawks were not double counted between locations, resulting in slightly different sample sizes at each study area (southern multi-seral: n=104, southern wildfire: n=105, northern multi-seral: n=134, northern wildfire: n=134). At each of those sampling locations, we randomly sampled 30 ten-minute recordings that were collected between 22:00 and 05:00 when common nighthawks are most vocally active, and between June 1 and July 30 when common nighthawks are present on the breeding grounds (Ng et al. 2018; Knight et al. 2021a). All recordings were collected by Wildlife Acoustics Inc. ARUs (SM2+, SM3, SM4; Maynard, Massachusetts, USA). All recordings were made with two omnidirectional microphones at a sampling rate of at least 44.1 kHz with a 16-bit depth.

Acoustic Recording Processing - Occurrence

We used Song Scope (Wildlife Acoustics 2011) software to extract common nighthawk detections from all ARU recordings at each station. Song Scope software has previously been shown to perform well for building occurrence datasets for the common nighthawk calls (Chapter 2). We trained our recognizer with 138 unmasked clips of common nighthawk calls recorded at close proximity (< 50 m). Training recognizers with calls recorded at close proximity is important to ensure the processed data meets two important assumptions of density estimation (Chapter 3). First, that the probability of detection is approximately one at zero metres, and second, that it decays with increasing distance following a half-normal distance curve (Buckland et al. 2015; Sòlymos et al. 2013).

We used the recognizer to scan the recordings using a minimum score threshold of 65 and a minimum quality threshold of 30, which we have previously shown optimizes the tradeoff between false positive and false negative detections, while ensuring that the probability of detecting an individual at 0 m is near 1 (Chapter 2, 3). All potential detections of common nighthawks reported by the recognizer were visually and/or aurally validated to confirm whether they were true or false positive detections. During validation, observers also noted all detections that were accompanied by a mechanical wing-boom display that indicates territorial habitat use. Common nighthawks vocalize when they perform a wing-boom display (*unpublished data*).

Acoustic Recording Processing - Abundance

Using the processed occurrence results, we randomly selected at least 14 (minimum number available for a study area; south multi-seral: 14, south wildfire: 42, north multi-seral: 20, north wildfire: 19) recordings with confirmed nighthawk presence from each study area to determine abundance of territorial nighthawks in each recording. Individual common nighthawks are vocally distinct (Armstrong 1965) and can be separated by experts through aural and visual interpretation of the spectrogram based on call duration, band width, region of maximum amplitude, and time between vocalizations (Figure 7.3). The first author reviewed all vocalizations with wing-booms in each selected recording and assigned an individual identification to each.

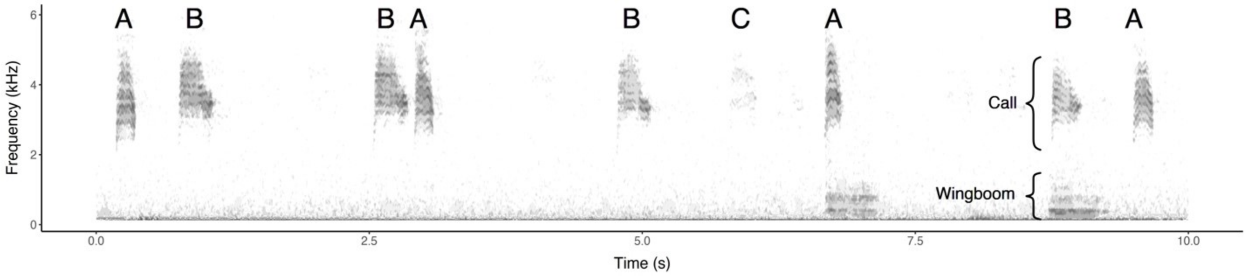


Figure 7.3. Spectrogram of multiple individual common nighthawks vocalizing and performing aerial wing-boom displays. Individuals are labelled A, B, and C and were distinguished visually based on call duration, band width, region of maximum amplitude, and time between vocalizations. Recorded on June 7, 2016 north of Ft. MacKay, Alberta, Canada.

Acoustic Recording Collection & Processing – Known Distance

We collected recordings of common nighthawk vocalizations with known distances so that we could estimate the effective survey area of our recognizer, allowing us to convert our abundance estimates into density estimates. We collected 495 clips of common nighthawk vocalizations with known distance by attracting territorial males with conspecific broadcast calls to the beginning of an transect of ARUs (detailed methods in Chapter 3; Yip et al 2020). We scanned those vocalizations with our recognizer using the same score and quality thresholds we used to produce the occurrence dataset (60, 35, respectively); EDR decreases with increasing score threshold due to the trade-off between true and false positive detections (Chapter 3). The first author visually reviewed the recognizer results to remove false positives. The result was a binomial dataset of 495 detections and non-detections for vocalizations of known distance.

Environmental Covariate Data Collection

We determined two environmental covariates for each study area that were shown to affect common nighthawk habitat use (Chapter 5, 6). The first was the proportion of coniferous forest within 200 m of each survey location. We reclassified the North American Land Cover Classification (30 m cell size; Canadian Centre for Remote Sensing 2010) as coniferous (classes

1 and 2). We created a new raster layer with the value of each cell calculated as the mean percent coniferous cover within a 200 m buffer. We used this raster-based approach instead of using buffers around each recording station to calculate the mean so that we could use the raster to create spatial predictions of habitat use later. The second covariate was fire year at each site, for which we sourced vector layers for Alberta (Alberta Agriculture and Forestry 2020) and Northwest Territories (NWT Centre for Geomatics 2020) and converted to raster layers at the same grain and extent as the conifer layer. We extracted the point value of each raster for each survey location. For the fire layer, we then converted fire year to time since fire by subtracting fire year from recording year. Any locations with no historic fire information were assigned the highest time since fire in that region's dataset. We similarly converted the fire raster layer to time since fire for spatial prediction by subtracting fire year from the year the layer was produced.

Abundance Models

Model Overview: We estimated the density of common nighthawks from recognizer-processed data by adapting the traditional single-season occupancy model (MacKenzie et al. 2002) to include elements of the qPAD approach (Sòlymos et al. 2013) and zero-inflated conditional likelihood (Lambert 1992; Sòlymos et al. 2012). PAM, particularly when data is processed by recognizers, is well-suited for hierarchical occupancy-style modelling because it creates multi-visit datasets that can be used to estimate detectability. Our model is based on a theoretical cascade of processes that determine survey outcomes, starting with site suitability and ending with detection (Figure 7.1). As an adaptation of occupancy modelling, these processes can be split into two levels: the site level, and the visit level.

At the site level, we split the probability of occupancy into two separate parameters: suitability and inhabitancy (to differentiate from “occupancy”; Figure 7.1). A site can either be suitable or unsuitable habitat, which is determined by the environmental resources (i.e., predictors) available at that site. If a site is suitable, it can either be inhabited or uninhabited if the habitat is not saturated. Because we are interested in density, not inhabitancy, there can be 1, 2, 3, or more individuals at occupied sites. In a single season occupancy model, the effect of environmental predictors is considered when estimating the probability of occupancy; however, we have separated the effect of predictors out by adding this separate suitability parameter so that we can use zero-inflated conditional likelihood to estimate the probability of suitable yet uninhabited sites from our abundance data, which separates truly uninhabited sites (i.e., true zeros) from false negative detections.

At the visit level, we also split the probability of availability for detection into two components: territorial activity and territorial presence (Figure 7.1) (Marsh & Sinclair 1989). An occupied site can either have at least one bird that is present or absent at the time of survey. If a bird is present, they can either be active and available for detection or not. We separated out availability for detection into these two different components, presence and activity, because our study species (and many others) are often away from the study site somewhere else in their home range. Because we had two different sources of information for availability of detection, time to first detection and multiple visits, we were able to separate out presence from activity. And finally, if a bird is active on territory, it can be detected or not detected by the ARU and recognizer, and this probability of perception is a function of distance. Perceptibility is a critical component of density estimation because it determines the amount of area that individuals are detected within. We used our separate dataset of detections of known distance to estimate the effective detection radius of our surveys to convert our abundance estimates from the modelling process into density estimates.

In summary, our approach used four pieces of information from three different data sources (Figure 7.1). 1) The time to first detection in each recording from occurrence data was used to estimate probability of common nighthawk activity. 2) The detection and non-detection information across multiple recordings (i.e., visits) from the occurrence data was used to estimate probability of availability, 3) the abundance data from recordings with known occurrence was used to estimate mean abundance for each study area, and 4) the known distance data was used to estimate perceptibility.

Step 1. Territorial activity: First, we used survival analysis to estimate the probability of common nighthawk territorial activity. In other words, for any given recording, what is the probability that a nighthawk is active and performing wing-boom displays, if it is present at the study site? We fit a parametric survival regression model with time to first wing-boom detection in each of our 14,310 recordings as the response variable. For recordings in which a common nighthawk was not detected, we treated non-detections as censored events with a time of detection of 10 minutes (i.e., recording length). As predictor variables, we included region to account for latitudinal differences in territorial activity (Hannah et al. *In review*), day of year and a quadratic effect of day of year to account for seasonal differences in activity, and the sin and cos of time of day to account for circadian patterns of activity. We included time of day as trigonometric functions of time of day to allow for circularity. The probability of territorial activity (p_{aj}) for a given recording is thus:

$$p_{aj} = 1 - e^{-10a_j}$$

where a_i is the event rate per minute. We built a global model with all predictor variables and compared it to all potential simpler models using AICc. We selected the most parsimonious model within ΔAICc = 2 and used it to predict the probability of territorial activity for each recording in our dataset.

Step 2. Inhabitancy: Next, we used a conditional likelihood model to estimate the proportion of uninhabited but suitable sites in each of our four study areas (Sòlymos et al. 2012). In other words, what is the proportion of true non-detections (i.e., true zeros) in our dataset, given the distribution of abundances observed at sites with detections? We used our dataset of abundance per recording in a zero-inflated model and to estimate the mean abundance of common nighthawks (λ_h), conditional on the abundances observed in recordings of known occurrence. We fit two sets of models: one with study area as a predictor and a null model. For each set, we fit a zero-inflated Poisson model (ZIP) and a zero-inflated negative binomial (ZINB) model and selected the model with the lowest AICc. Including ZINB models at this step is important because the Poisson distribution does not always reliably fit the abundance-occupancy relationship that this conditional likelihood step estimates (He et al. 2016). We selected the model with the lowest AICc (ZIP with treatment) and then used a sandwich estimator to estimate λ_h because our abundance dataset had multiple recordings sampled for some survey stations. We randomly sampled one recording for each survey site within the abundance dataset and fit our best fitting model to that sample. We repeated the process 300 times and calculated the mean, 2.5% and 97.5% quantiles of the λ_h estimates. We estimated the probability that a given site (N_i) is uninhabited for each study region as:

$$P(N_i = 0|\delta_h) = e^{-\lambda_h}$$

Step 3. Suitability and Territorial Presence: Next, we used the outputs from the previous two steps as inputs for our modified occupancy model. We made two modifications to the traditional single species single season occupancy model, one to each level of the model (detectability and occupancy levels) (MacKenzie et al. 2002).

First, we included the predicted probabilities of territorial activity for each recording as a statistical offset on the detectability side of the model, which essentially accounts for differences in detectability by converting count data to rates. Including this second source of availability information to the multi-visit information in the occupancy model allowed us to separate out the two components of availability in each recording (p_{dj}): activity (p_{aj}) and presence (p_{pj}; Figure 7.1):

$$p_{dj} = p_{aj} p_{pj}$$

The benefit of separating these two components of availability is that it removes the closure assumption of occupancy modelling, which assumes that if a site is occupied during one survey, that it is occupied during all surveys (MacKenzie et al. 2002; Rota et al. 2009). In other words, the site is "closed" to movement in and out of the survey radius between visits or recordings. Estimating the probability of presence also provides a quantitative estimate of the closure assumption for other applications.

Second, we redefined the probability of occupancy (Ψ_i) at a given site as the sum of suitability (δ_i) and occupied given suitable ($1-e^{-\lambda h}$; Figure 7.1):

$$\Psi_i = \delta_i\left(1 - e^{-\lambda_h}\right)$$

The covariates included in the model therefore are interpreted as those that predict habitat suitability, not occupancy.

We fit a global model with both environmental covariates (coniferous, fire) and an interaction between the two as predictors of suitability. We compared that global model to all simpler models and selected the model with the lowest AICc for density estimation and prediction.

Step 4. Effective survey area: We then used our third dataset of detections and non-detections for vocalizations of known distance to estimate the effective survey area of our recognizer (A_{eff}). We fit loglinear models to our binomial dataset with distance as the predictor, although we transformed distance as the negative squared distance so that it was a linear predictor. We did not include vegetation covariates in our loglinear models because common nighthawks vocalize and perform the wing-boom display above the canopy and thus should be relatively unaffected by differential attenuation of varying vegetation type. We fit two models: one with a log link and one with a complementary log-log link and selected the model with the lowest AICc value. Following the half-normal detection function (refs), we then calculated the effective detection radius (τ) of our recognizer as

$$\tau = \left(\frac{1}{\beta}\right)^{-0.5}$$

where β is the coefficients from the selected loglinear model and τ is the distance at which the number of individual birds detected outside τ is equal to the number of individual birds missed within τ. Finally, we converted τ to A_{eff} using the area of a circle:

$$A_{eff} = \pi\tau^2$$

Step 5. Density Estimation & Population Prediction: Finally, we put all the pieces together to estimate density at each study site.

$$Density_i = \frac{\delta_i \lambda_h}{A_{eff}}$$

We also converted our density estimates into mean territory size estimates using the occupancy estimates:

$$Territory_h = \frac{1}{Density_i * \Psi_i} = \frac{1}{Density_i * \delta_i(1 - e^{-\lambda_h})}$$

We used the density estimates for each combination of study area and habitat variables to create spatial predictions and population estimates for each study area. First, we aggregated our raster files by a factor of ten (30 m cells to 621 m cells) to reduce computation requirements. We used the selected occupancy model to predict δ_i for all possible combinations of covariates, calculated density for each combination in each study area, converted our per hectare density estimates to individuals per raster cell, and assigned the appropriate density estimate to each raster cell in each study area. We then summed the raster cell density estimates to create a population estimate for each study area.

Results

Step 1. Territorial Activity

The mean time to first wing-boom detection of the raw data was 4.54 minutes for recordings with detections and 8.91 minutes for all recordings, with non-detections assigned the maximum time of 10 minutes. The selected survival model for probability of territorial activity was the global model, which included a linear and quadratic term for day of year, the cos and sin of time of day, and latitudinal region (Supplementary Materials). The probability of territorial activity was highest just after midnight and during the month of June, with higher overall probability of territorial activity in the south study region (Figure 7.4). The mean predicted time to first wing-boom detection from the selected model was 4.29 minutes (SD = 6.00 minutes) and the mean probability of territorial activity in a 10-minute recording was 0.18 (SD = 0.21).

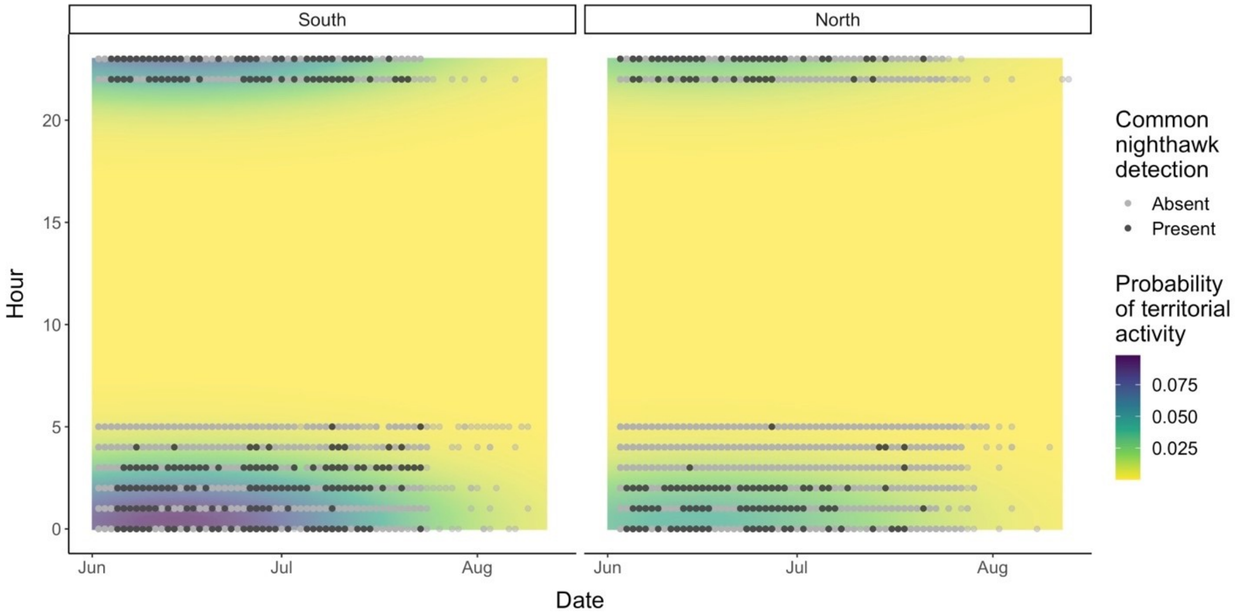


Figure 7.4. Probability of common nighthawk territorial activity during the breeding season at two latitudes of Canada's western boreal forest (south = northeastern Alberta, north = Northwest Territories). Probabilities were predicted from the coefficients of a parametric survival regression model.

Step 2. Inhabitancy

The number of nighthawks detected in a recording ranged from 1-4, with the maximum detected at the south wildfire site. The ZIP model with study area as a covariate was the selected model for prediction of mean abundance (ΔAICc = 2.23, Supplementary Materials). The wildfire areas had higher predicted mean abundance at suitable sites than the multi-seral sites, with the highest overall abundance at the south wildfire site (Table 7.1). The mean abundance (λ) of suitable sites across all four study areas was estimated as 0.25. The mean suitable and uninhabited probability was 0.89. The mean suitable and inhabited probability was 0.11.

Table 7.1. Suitability, inhabitancy, abundance, density, and territory size estimates for common in Canada's western boreal forest at four study areas (south = northeastern Alberta, north = Northwest Territories).

Study area	Mean habitat suitability	Proportion of suitable sites that were inhabited	Mean abundance at suitable sites (λ_h; 2.5 – 97.5% quantile)	Mean density (males/ha)	Mean territory size (ha)
South multi-seral	0.80	0.00	0.00 (0.00 – 0.01)	0.000	14.6
South wildfire	0.93	0.42	0.54 (0.32 – 0.81)	0.038	9.7
North multi-seral	0.80	0.08	0.08 (0.00 – 0.16)	0.005	14.0
North wildfire	1.00	0.32	0.39 (0.22 – 0.44)	0.030	9.7

Step 3. Suitability and Territorial Presence

We detected territorial common nighthawks at 159 of the 477 study sites: 6 in the south multi-seral study area, 50 in the south wildfire area, 19 in the north multi-seral area, and 84 in the north wildfire area. The selected modified occupancy model included a strong positive effect of the proportion of conifer within 200 m and a weak negative effect of years since the most recent wildfire (Table 7.2). Suitability across the study sites ranged from 0.31 to 1.00, with a mean suitability of 0.89. Mean suitability was higher in the wildfire study areas than the multi-seral areas (Table 7.1).

Table 7.2. AICc ranking of modified occupancy models for estimation of common nighthawk density. Conifer represents the proportion of coniferous forest in a 200 m radius and firetime represents the number of years since the most recent wildfire.

Model	df	logLik	AIC_c	ΔAIC_c	AIC_c w
~conifer*firetime	5	-2057.04	4124.20	0.00	0.49
~conifer + firetime	4	-2058.88	4125.85	1.66	0.21
~conifer	3	-2060.07	4126.19	2.00	0.18
~firetime	3	-2061.89	4127.81	3.61	0.08
~1	2	-2061.53	4129.11	4.91	0.04

We detected territorial common nighthawks in 853 of the 4,770 recordings at sites with territorial common nighthawk occupancy: 18 at the south multi-seral study area, 511 at the south wildfire study area, 57 at the north multi-seral study area, and 267 at the north wildfire study area. After accounting for the probability of territorial activity and the overall probability of availability for detection in these multi-visit data, the probability of territorial presence after accounting for imperfect probability of territorial activity was 0.95.

Step 4. Effective Survey Area

Three hundred of the 495 common nighthawk vocalizations of known detection distance were detected by the recognizer (Figure 7.5). The log model was the best-fitting model (ΔAICc = 116.44). The effective survey area of the recognizer was 13.3 ha.

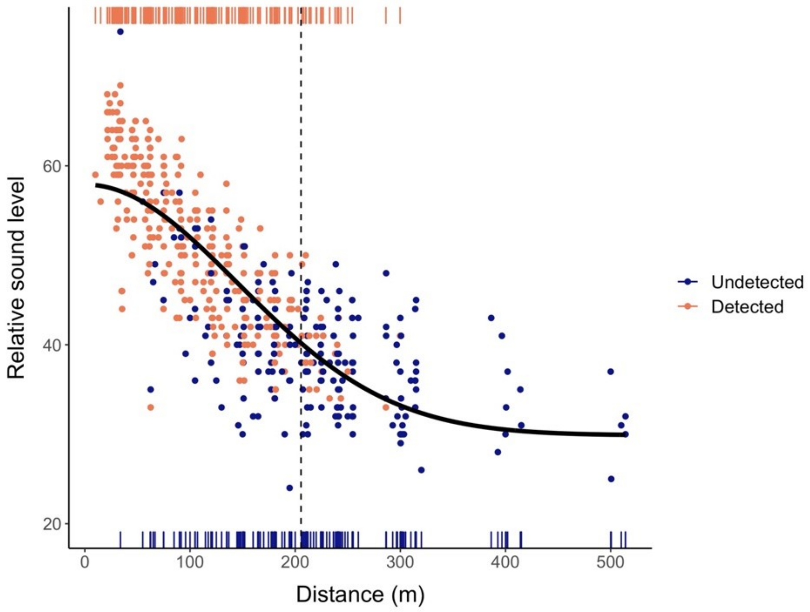
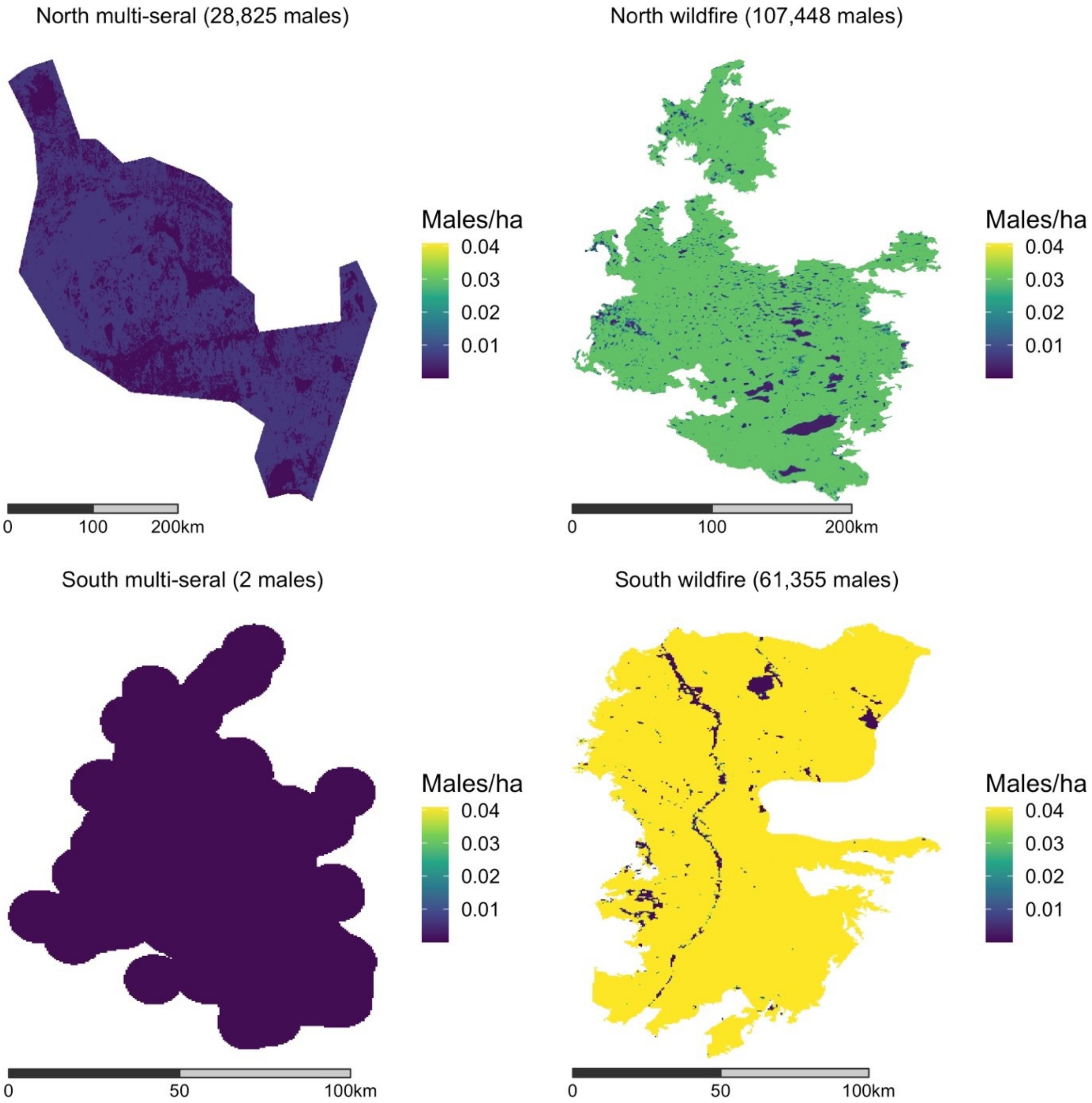


Figure 7.5. Relative sound level and distance from the recorder for common nighthawk vocalizations and whether those vocalizations were detected with automated recognition. The black line indicates the probability of detection (i.e., perceptibility), as determined with a loglinear function. Dashed line represents the effective detection radius; the distance at which the number of individual birds detected outside that radius is equal to the number of individual birds missed within. Rugs indicate detected (orange) and undetected (blue) data points.


North multi-seral (28,825 males)

North wildfire (107,448 males)

South multi-seral (2 males)

South wildfire (61,355 males)

Figure 7.6. Density estimates of territorial male common nighthawks at four study areas in Canada's western boreal forest. Two study areas were in recent (< 10 years) wildfires and two were in areas with multi-seral forest. One of each type was in northeastern Alberta (south) and one of each in the Northwest Territories (north).

Step 5. Density Estimation & Population Prediction

The mean estimated density across the 477 study sites was 0.018 territorial male nighthawks per ha, with much higher densities estimated for the wildfire areas (Table 7.1, Figure

7.6). The estimated territory size for the south wildfire area was very similar to existing territory size estimates from VHF-tagged birds from the same area (9.7 vs 10.2 ha: Chapter 4).

Discussion

We developed and implemented a new model for density estimation of animals from acoustic recordings collected by single point autonomous recording units (ARUs) and processed with automated recognition. Our method used three different data types to estimate density in five steps, including a modified occupancy model. We applied our method to estimate density of territorial male common nighthawks from ARU recordings collected in four different study areas: two recently (< 10 years) burned and two with forest of mid and late seral stages ("multi-seral). As predicted, we found much higher densities in the wildfire study areas (0.038 and 0.030 males per ha) than in the multi-seral areas (0.000 and 0.005 males per ha). Our territory size estimate for the south wildfire area was similar to existing territory size estimates from VHF-tagged birds in the same study area (9.7 vs 10.2 ha), suggesting our approach yields accurate density estimates.

The motivation for our density estimation approach was to develop a model that could be applied at broad geographic scales. To limit the site-specificity of our approach, we based our model on a theoretical cascade of processes from site suitability to detection and estimated parameters at each step in that cascade. Estimating each parameter separately allowed us to add covariates at each step of the model, therefore adjusting each parameter according to the conditions under which the data was collected. The ability to incorporate covariates is particularly important for availability of detection because vocal activity of animals is known to vary with time of day and season (York et al. 2014; Amrhein et al. 2002), moon phase (York et al. 2014), breeding status (Amrhein et al. 2002; Upham-Mills et al. 2020) and more. In contrast, vocal activity rate methods of density estimation incorporate availability by averaging it out across dates, times, and locations (Pérez-Granados & Traba 2021; Pérez-Granados et al. 2021). Taking an average as opposed to accounting for availability with covariates in the modelling process limits the applicability of the vocal activity estimates. Call rate methods of estimating density are thus unlikely to be applicable at broad geographic scales because exogenous and endogenous variation in call rate may render this method imprecise and only applicable to the study sites where vocal activity information was collected (Pérez-Granados & Traba 2021; Sebastián-González et al. 2018; Stevenson et al. 2015). Furthermore, if the ARU dataset is skewed in the sampling of any predictor of availability, the density estimates will also be skewed.

Other approaches for estimating density from single ARUs that accommodate covariates of availability have other assumptions that renders then unsuitable for many species. Doser et al.'s (2021) model uses a multi-visit n-mixture framework, which assumes population closure and thus is unsuitable for species with large home ranges. Our model does not assume closure, and in fact, provides an estimate of the probability of closure ("presence" in our model). The probability of closure between recordings for territorial common nighthawks was 0.95, which is expected because the mean territory size in the study area is slightly less than the effective survey area of our recognizer (9.7 vs 13.3 ha). Quantification of closure and subsequent effects on occupancy estimates suggest it is common and can bias estimates (Rota et al. 2009).

An additional advantage of including covariates at each step is that our density estimation approach can incorporate predictors of habitat suitability directly in the model. Separate parameter estimates for suitability and inhabitancy will prove particularly useful for applications like critical habitat modelling, which aim to provide habitat for the long-term recovery and conservation of species (Government of Canada 2016). The downside of incorporating covariates into the model is that it can bias predictions if the wrong covariates are used. For example, we found coniferous forest strongly affected common nighthawk habitat suitability; however, this is association is likely driven by a strong selection for pine forest specifically (Chapter 5, 6), and thus the density predictions for spruce-dominated forests are likely to be high. We also recommend that if the study area is large enough to contain multiple metapopulations or biozones, users should include region as a covariate in the model or model each region separately (including separate abundance datasets), with overlap, and blend to create a final model.

There are limitations and disadvantages to our model, the greatest of which remains the challenge of obtaining the abundance and detection distance datasets. We used a spectrogram analysis approach for obtaining our abundance dataset; however, spectrogram analysis does not work for all species and so any of the other established methods for obtaining abundance data could be used in lieu, including paired point counts (Doser et al. 2021) or vocal activity rate (Pérez-Granados et al. 2021). Regardless, we recommend that the abundance dataset used in the model is a paired subset of the ARU recordings used for the survival and occupancy models; otherwise, population differences in abundance can bias density estimates.

The detection distance dataset is a critical component of all density estimation approaches; without information on the perceptibility of the ARU surveys, abundance estimates cannot be converted to density. Perceptibility can either be incorporated via distance sampling or by estimating the effective survey area. Our approach requires estimation of the effective survey area, which can be done either with a dataset of reference recordings at known distances, as done here, or by comparing the abundance dataset to paired point recordings (Wilgenburg et al. 2017).

We echo the endorsement of Perez-Granados (2021) for the reference dataset approach because it can be used for other purposes, like estimating distance from relative sound level (Darras et al. 2018, Yip et al. 2020) and using that information to limit the radius of ARU surveys (Hedley et al. 2020). Sourcing the reference dataset from the same study sites and area as the density estimation is less important; however, for many species, it will be important to at least sample vegetation types that are representative of the main ARU dataset because effective survey area depends on vegetation type via differential amounts of sound attenuation (Yip 2020). Regardless of the approach used to incorporate perceptibility in density estimation, we reiterate the importance of doing so for the automatic recognition algorithm and classification threshold used to process the data per se, as both components can greatly affect the survey area, in addition to the recorder type (Chapter 3, Yip et al. 2017); using the survey radius of human listeners and especially human point counts is inappropriate and will underestimate density by overestimating the survey area.

We conclude that our model is another tool in the density estimation toolbox. Vocal activity rate approaches are relatively easy and quick to implement for small-scale studies (Sebastián-González et al. 2018), n-mixture model approaches that incorporate false positive detections are efficient and suitable for species that have small home ranges (Doser et al. 2021), and acoustic SECR approaches with ARU arrays are best for applications that require highly precise estimates (Stevenson et al. 2015). Our approach is well-suited for estimation at broad geographic scales or species with large home ranges that do not satisfy the closure assumption. Although we designed our density modelling approach with data from automated recognition in mind, our model is extendable to acoustic recordings processed by human listeners or even human point count data. Our model can also easily incorporate integration of multiple datatypes via the availability and perceptibility steps of the model. As the availability of acoustic recordings that cover large geographic extents increases (Sugai et al. 2018), so does the capacity to estimate population size more accurately. Our method may be preferable to existing population estimates for species that are poorly monitored by existing surveys (Rosenberg et al. 2017), that occupy remote regions (Sòlymos et al. 2020), or that have geographic variability in availability for detection (Hannah et al. in review). The common nighthawk checks all three of these boxes, and our results suggest that the existing population estimate is low (Partners in Flight 2020). For example, our density estimate for the wildfire study area in Alberta's boreal forest (~61,000 territorial males) is almost as large as the PIF estimate for the entire province (71,000 individuals), despite sizeable populations in many other areas of the province. We therefore suggest that population estimation for wildlife management move towards an integrative, species-specific "if the shoe fits" approach, as opposed to the current "one size fits all" approach, and our density estimation approach will be a valuable tool to help this shift towards more accurate estimates, especially for wide-ranging species.

Supplementary Materials

Model comparison results for survival and zero-inflated models.

Table 7.A.1. AICc ranking of parametric survival models for estimation of common nighthawk availability for detection. Day represents day of year. Sin and cos represent trigonometric functions of time of day and were used to allow for circularity. Region represents a categorical covariate for study area region (north, south) to account for differences in civil twilight between latitudes. All models with quadratic terms also included a linear term of the same variable. Models with quadratic terms also include the linear term of that covariate.

Model	df	logLik	AIC_c	ΔAIC_c	AIC_c w
~day^2 + sin + cos + region	6	-3679.23	7373.68	0	0.96
~day^2 + cos + region	5	-3683.93	7380.09	6.4	0.04
~day + sin + cos + region	5	-3686.33	7384.88	11.2	0
~day +cos + region	4	-3690.82	7391.06	17.38	0
~ sin + cos + region	4	-3704.17	7417.77	44.08	0
~ cos + region	3	-3708.27	7423.37	49.69	0
~day^2 + sin + cos	5	-3753.33	7518.88	145.19	0
~day + sin + cos	4	-3757.36	7524.15	150.47	0
~day^2 + cos	4	-3757.89	7525.21	151.53	0
~sin + cos	3	-3761.52	7529.86	156.18	0
~day + cos	3	-3761.8	7530.43	156.75	0
~cos	2	-3765.76	7535.91	162.23	0
~day^2 + region	4	-3995.59	8000.61	626.92	0
~day + region	3	-3997.58	8001.99	628.31	0
~region	2	-4014.96	8034.31	660.63	0
~1	1	-4187.17	8376.48	1002.79	0
~day	2	-4187.14	8378.69	1005	0
~day^2	3	-4187.1	8381.02	1007.34	0

Table 7.A.2. AICc ranking of zero-inflated models for conditional likelihood estimation of common nighthawk occupancy and abundance at suitable sites. Two distributions were compared, Poisson and negative binomial. For each distribution, a null model was compared to a model with study area as a categorical covariate.

Model	df	logLik	AIC_c	ΔAIC_c	AIC_c w
Poisson(λ),~ study area	4	-52.17	112.8	0.00	0.74
NegBinom(r, p), ~ study area	5	-52.17	115.0	2.23	0.24
Poisson(λ),~ 1	1	-59.86	124.3	8.98	0.01
NegBinom(r, p), ~ 1	2	-59.74	128.6	10.82	0.00

Chapter 8. Synthesis

Dissertation Overview

In this thesis, I demonstrated the importance of spatial and behavioural context for using passive acoustic monitoring (PAM) and automated recognition in ecology, with application to the common nighthawk (*Chordeiles minor*). In Chapters 2 and 3, I showed that the classification threshold (hereafter, "score threshold") of recognizers is a proxy for survey area and that the performance of a recognizer strongly depends on the threshold used. In Chapter 4-6, I showed that the common nighthawk's wing-boom display is an indication of breeding territoriality and nesting, and that this behavioural context can provide greater insight into common nighthawk habitat use and ecological principles. In Chapter 5, I used the behavioural context of the wing-boom display to understand common nighthawk habitat use in the western boreal forest and explore a new idea about the relationship between scale and movement. In Chapter 6, I again used the wing-boom display to test two competing hypotheses for why common nighthawks use recently disturbed areas in the boreal forest. In Chapter 7, I combined my results on spatial and behavioural context to a new method for density estimation from recognizer data and applied it to show common nighthawk populations are particularly dense in post-wildfire areas. Below I synthesize my research across my research chapters and the scientific literature for common nighthawk conservation and passive acoustic monitoring. I have included select references to media features that highlight my research throughout the following sections.

Common Nighthawk Conservation

Although the goal of my thesis was to contribute to the broad-scale use of PAM, the motivation was to contribute to knowledge and conservation of the common nighthawk

(*Chordeiles minor*). I have admired, studied, and championed this species for over a decade. Given that my thesis and associated work make up a large proportion of existing common nighthawk research to date, my perspective on its ecology and management are below, with reference to chapters of my thesis for evidence and recommendations.

Boreal Forest Conservation

The thesis components of my common nighthawk research were set in Canada's western boreal forest, which has previously been estimated to provide breeding habitat for a large portion of Canada's common nighthawk population (Haché et al. 2014). Three other studies have examined common nighthawk habitat associations, one in the western boreal forest (Sidler 2017) and two in the eastern boreal forest (Foley 2018; Farrell et al. 2017, 2019). All three support the conclusion of my research that common nighthawks are "disturbance specialists" in the boreal forest, with higher occupancy or abundance detected in earlier seral forests (media features: CBC News 2017, Struzik 2017). Chapter 6 of my thesis adds nuance to this label by specifying that common nighthawks are disturbance specialists for territorial habitat, but not for extraterritorial habitat use. This distinction will greatly facilitate conservation prioritization for this species in the boreal forest because it narrows down areas of conservation interest, as seen in the spatial predictions of Chapter 5.

The confirmation of common nighthawks as disturbance specialists in the boreal forest does present a challenge, however, because protection of post-disturbance areas where common nighthawks currently nest is not an effective long-term or even medium-term conservation strategy. Instead, regional land use planning in Canada's boreal forest should incorporate an emulation of natural disturbance approach that includes early seral habitats. Regenerating forests are the most poorly understood seral stage (Lindenmayer et al. 2019); incorporating the early seral component of the boreal forest's shifting mosaic into land use planning will also benefit other disturbance specialist species in the boreal forest like the black-backed woodpecker (*Picoides arcticus*) (Tingley et al. 2016, 2020, 2018), tree swallow (*Tachycineta bicolor*) (Hobson & Schieck 1999), and olive-sided flycatcher (*Contopus cooperi*) (Haché et al. 2014; Environment Canada 2016b). Alternatively, Chapters 5-7 of my research also show, that pine forest is an even better predictor of common nighthawk territorial habitat in the boreal forest. If long-term habitat protection for the common nighthawk is ever considered in the boreal forest, it should focus on open pine forest, particularly the sandy, post-glacial areas of the Athabasca Plain. Similarly, any designation of critical habitat based on a set of environmental conditions (Rosenfeld & Hatfield 2006) should include pine forest and sandy soils.

That being said, the fire regime in the boreal forest is shifting with climate change (de Groot et al. 2013). As fire frequency increases, so will the proportion of the boreal landscape that is suitable for common nighthawk nesting and other post-disturbance specialists. Which begs the question, do common nighthawks need habitat protection in the boreal forest? Perhaps not, although further research on effects of fire severity on common nighthawk nesting habitat suitability should be conducted prior to drawing any conclusions. Community-level research suggests common nighthawks are most abundant after low severity fires (Knaggs 2018) and fire severity is also expected to increase with climate change. Population forecasting would also be helpful to confirm whether habitat protection is needed, and the tools in my thesis, particularly Chapter 7, provide an analytical foundation to do so. Regardless of forecasting results, care should continue to be taken to avoid disturbance or damage to common nighthawk nests, as per the *Migratory Bird Conservation Act SC 1994*, and my research in Chapter 4 provides recommendations on how to do so.

Range-wide Conservation

Although my thesis was restricted to the boreal forest, synthesis of my results with research from the rest of the common nighthawk's breeding range provides further insight into its ecology. The limited literature available suggests that common nighthawks use separate areas for nesting, foraging, and roosting across the breeding range (Fisher et al. 2004; McGuire & Brigham 2017; Caccamise 1974; Armstrong 1965; Newberry et al. 2018); therefore, behavioural context is important for interpreting the habitat associations of this species in southern landscapes as well. I discuss nesting and foraging habitat separately below.

In non-boreal landscapes, wetlands are a key habitat component for aerial insectivore foraging. Tree swallows select for wetlands for foraging, particularly in agroecosystems (Michelson et al. 2018; Elgin et al. 2020) and aquatic insects improve nestling development and breeding success (Twining et al. 2018) due to added nutritional benefits of omega-3 fatty acids relative to terrestrial insects (Génier et al. 2021). Although the importance of wetlands for common nighthawk foraging has not been directly tested, the presence of nighthawks was positively related to distance to water in southern Saskatchewan (Ng 2009) and flocks of common nighthawks are known to forage over open water during the breeding season (McGuire et al. 2021). I predicted wetlands to be unrelated to common nighthawk habitat use in the boreal forest based on the dominance of terrestrial beetles we found in food boluses (Knight et al. 2018) and the proliferation of aerial insects in the boreal forest; however, the results from Chapter 6 suggests that wetlands may be important common nighthawk foraging habitat across the range. Farrell et al. (2017) found similar common nighthawk occupancy rates in wetlands and post-disturbance areas, which would be explained by the combined effects of foraging and nesting, respectively. Given that wetlands are emerging as a fundamental component of resource requirements for the common nighthawk and other aerial insectivorous species, I echo the call to strengthen wetland protection and management across Canada to conserve aerial insectivores and preserve valuable ecosystem services (Nebel et al. 2020), including in the boreal forest where wetland loss also occurs (Morissette et al. 2019).

Common nighthawk nesting habitat in non-boreal parts of the range is highly diverse; however, there are two characteristics that most nesting habitats have in common. The first is a paucity of ground cover, at least at the nest site level (Lohnes 2010; Allen & Peters 2012; Jennifer 2015). The second characteristic is the relative permanence of nesting habitats. In the boreal forest, post-fire nesting habitat availability shifts across the landscape over time. In contrast, non-boreal nesting habitats like gravel rooftops, rocky outcrops, sand bars and dunes (Brigham et al. 2011) are permanent fixtures on the landscape. The exception would be areas like thinned Oregon forests where common nighthawk is disturbance associated (Hagar et al. 2004). Even within grassland habitat, the areas with minimal ground cover that are most suitable for nesting are likely to be relatively spatially stable because they are driven by soil characteristics. Conservation of common nighthawk nesting habitat in non-boreal landscapes is therefore a much more important consideration for habitat protection. High densities of common nighthawk nest sites (e.g., 25 m between nests; Sutherland 1963) in some locations further emphasizes that nesting habitat is limited in some non-boreal populations. No non-boreal habitat studies have been conducted that model the wing-boom signal per se, however, and so I recommend nesting habitat research for southern populations is important to begin the process of identifying habitat for protection. The PAM and behavioural context tools from Chapters 2, 3, 4, and 7 of my thesis will greatly facilitate nesting habitat research. In the meantime, the community science dataset from the Canadian Nightjar Survey differentiates the wing-boom signal from the call and could be used immediately to pinpoint areas with particularly high densities of breeding males for protection (Knight et al. 2019a).

The difference in permanence of nesting sites between boreal and non-boreal populations suggests that there are likely regional differences in metapopulation dynamics for this species. In landscapes with relatively high disturbance rates like the boreal forest, specialist species are expected to have high dispersal abilities (Büchi & Vuilleumier 2016; Johst et al. 2002). In contrast, species that inhabit more stable environments are expected to be characterized by high competition and low dispersal (Büchi & Vuilleumier 2016). Common nighthawk dispersal in the boreal forest is virtually unstudied; however, one male captured and tagged for our migratory connectivity study relocated 125 km northeast between breeding seasons (Knight et al. 2021a). In non-boreal populations, gene flow between cities is low for rooftop nesting common nighthawks (Mays et al. 2019). Furthermore, there is a area of low predicted probability of occurrence of common nighthawks between the boreal and non-boreal portions of the common nighthawk breeding range, at least in western North America, that could be a barrier to gene flow (Haché et al. 2014, Knight et al. 2021a). Anecdotally these three pieces of evidence together suggest there may be genetic differences between boreal and non-boreal populations. A range-wide genoscape (Ruegg et al. 2014) of this species would provide insight into the dispersal ecology of the species and much-needed direction for species management.

As it stands, the common nighthawk is assessed and listed as a single species under Canada's *Species at Risk Act SC 2002*, despite its massive breeding range and obvious differences in ecology between boreal and non-boreal populations. Lumping both regions together under a single listing has led to an ineffective conservation process. For example, the threat assessment process during the recent species reassessment by the Committee on the Status of Endangered Wildlife in Canada (COSEWIC) concluded that threats to the common nighthawk were “high – low” due to differences between boreal and non-boreal populations. The result of the reassessment process was that the species was recommended for down-listing from Threatened to Special Concern, in part due to the abundance of common nighthawks detected in the boreal forest by the studies mentioned in this thesis (COSEWIC 2018). This decision may come at the cost of non-boreal populations, which have declined 68% since 1970 and continue to do so, albeit at a slower rate than previously (COSEWIC 2018). Unfortunately, COSEWIC has taken a “guilty until proven innocent” approach with this species and has refused consideration of splitting the common nighthawk into two designatable units (i.e., separate assessment and listing for boreal and non-boreal populations), citing “genetic and other consistency across the three subspecies” (COSEWIC 2017); however, there is no existing genetic research for boreal common nighthawks and subspecies designations are exclusively phenotypic. Research to determine whether there are genetic differences between boreal and non-boreal populations of common nighthawks should therefore be conducted as soon as possible before removing the species off Schedule 1 of the *Species at Risk Act SC 2002* further threatens non-boreal common nighthawk populations. I have initiated the process of collecting tissue samples for genetic analysis with a network of collaborators to facilitate this top priority.

Comparing and contrasting boreal and non-boreal populations could also provide some insight into mechanisms of population decline. Our recent work on migratory connectivity suggests that the drivers of differential common nighthawk population trends may be on the breeding grounds because populations from across North America mix during migration and on the wintering grounds (Knight et al. 2021b) (media features: The Globe and Mail 2018, American Bird Association 2020). As mentioned, there is a myriad of potential causes for common nighthawk population declines, even on the breeding grounds, but one of the major differences between boreal and non-boreal regions is the historical role of natural disturbance and the contemporary level of anthropogenic disturbance. Industrial agriculture has been pinpointed as a potential cause of aerial insectivore declines due to breeding habitat loss, wetland drainage, pesticide-caused reductions of insect abundance, and sub-lethal effects of neonicotinoid pesticides on birds themselves; however, empirical evidence of effects is mixed (Spiller & Dettmers 2019; Berzins 2020). Relating long-term monitoring data to land use change and pesticide use could provide some insight into population threats, although it will need to be done at a fine spatiotemporal grain because there is little spatial concordance in aerial insectivore population trends (Michel et al. 2015).

Unfortunately, quantitative comparison of boreal and non-boreal common nighthawk populations is impeded by limited long-term monitoring data on the boreal breeding grounds. Thanks to latitudinal differences in vocal activity (Hannah et al *In Review*) that are likely driven by differences in civil twilight (Sidler 2017), there are limited common nighthawk datapoints for the boreal forest (Haché et al. 2014). Even existing datasets for non-boreal regions like the North American Breeding Bird Survey are suboptimal for management because they are dawn surveys (Knight et al. 2021b). The recently formalized Canadian Nightjar Survey will improve long-term population monitoring for the common nighthawk and other nightjar species (Media feature: Global News 2021); however, additional ARU data collection and/or processing of existing ARU data would further facilitate habitat modelling and population estimation. The PAM tools in my thesis, both for data processing and analysis are transferrable across the species' breeding range to facilitate future research, although the behavioural context study in Chapter 4 should be repeated at a location where nest density is high (e.g., Sidney Island, British Columbia, Canada) to confirm territoriality in other populations (Brigham et al. 2011). Additional tools for common nighthawk management will likely become available as PAM continues to advance. Development of deep learning methods to differentiate common nighthawk individuals (Armstrong 1965) would be particularly useful to collect demographic data via acoustic mark-recapture. Current demographic datasets for the common nighthawk are few and far between because individuals are nearly impossible to recapture (Knight et al. 2021a) and nest monitoring is impeded by semi-precocial chicks (Kramer & Chalfoun 2012)

The caveat for much of the boreal versus non-boreal comparison above is that it is based on Western science. Common nighthawks have likely bred in the boreal forest since the last glacial maximum and being a conspicuous species, the Cree and Dene likely hold much wisdom about their ecology over time. Indigenous participation in conservation can lead to better outcomes via the inclusion of multiple perspectives (e.g., Schuster et al. 2019) and is a critical step towards reconciliation, decolonization, and upholding indigenous rights. If I regret anything about the last six years of my PhD research, it is not investing the time and effort to build collaborative relationships with the indigenous communities in the Fort MacKay and Fort Chipewyan areas of northeastern Alberta where much of my research was conducted. Above all else, I recommend future common nighthawk research involve indigenous people and perspectives.

Automated Recognition

The Value of Context

The failure of macroecology to find many broad rules for ecology is a nod to the importance of context in ecological research (Lawton 1999). The literature is rift with studies describing the context-dependency of ecological phenomena (e.g., Subalusky et al. 2019) and so it is our duty as ecologists to always be looking for opportunities to understand more about our data and how it affects our ecological inferences. When using PAM and recognizer technology, Chapter 2 of my thesis supports the conclusion of others that it is critical to think through objectives and analysis before deciding to use automated recognition (Priyadarshani et al. 2018). Practitioners should ask themselves whether the lack of context inherent in ARU data will 1) invalidate any assumptions of the planned statistical analysis, or 2) have the potential to provide misleading inferences. Below I discuss the value of considering the spatial and behavioural context of recognizer data.

Spatial Context

Together, Chapters 2 and 3 of my thesis showed that the ‘performance’ of recognizers varies with score threshold, the mechanism of which is spatial context. In Chapter 3, I showed that score threshold is a proxy for survey area. The implication of this work is that survey areas for recognizer-processed data are influenced by three factors: 1) the ARU model used (Yip et al. 2017b), 2) the recognizer algorithm and training data (Chapter 2, 3), and 3) the score threshold applied (Chapter 3). Recognizers will almost always have fewer detections than human listeners due to this survey area phenomenon. While maximizing recall (i.e., number of detections of the target species) is valuable because more detections provide greater statistical power for analysis, I suggest PAM researchers should move away from the dogma that human listeners are the “gold standard” and that false negatives relative to human observers are inherently bad (Priyadarshani et al. 2018; Stowell et al. 2018). Lower recall of recognizers relative to human observers is not necessarily because they are inferior to human observers, but that they are simply sampling a smaller area. Quantifying this survey area is important for density estimation (Pérez-Granados &

Traba 2021) and can also have other applications like limiting the survey radius to limit the inferences of a study to specific areas (Hedley et al. 2020).

Several methods are available for determining the survey area of PAM data. Chapter 3 of my thesis provides one approach for determining the effective survey area of a recognizer and Chapter 7 shows how that area can be used to convert abundance estimates to density. Alternatively, the relative sound level (RSL) of sounds in ARU data can also be automatically derived during the automated recognition process, used to predict the distance of vocalizing animals, and analyzed in a distance sampling framework for density estimation (Sebastián-González et al. 2018, Yip et al. 2020). A third method involves training human listeners to estimate distance, although this is a more time-intensive and subjective process because it is not automated (Darras et al. 2018). A fourth and final method involves paired ARU and human point count datasets and calculating an offset for known effective survey areas of human points counts via the relative number of detections between the two datasets (Van Wilgenburg et al. 2017).

Central to all these approaches for providing spatial context to recognizer data, save the paired dataset approach, is a dataset of sounds of the focal species recorded at varying known distances. I cannot emphasize enough how valuable a known distance dataset is for ecological applications of recognizer data. The known distance dataset I collected was at the foundation of this thesis as well as an additional paper that used machine learning to automatically weed out false positives from unvalidated recognizer data (Knight et al. 2020). I echo the call of Pérez-Granados and Traba (2021) for building sound libraries of known distance recordings to support future density estimation studies. There are three known methods for developing known distance datasets. The first and most precise method uses triangulation from an array of ARUs; however, this method is laborious to set up and post-process to obtain individual locations. The second method is the one used in my thesis, which uses broadcast calls to attract an individual to a transect of recorders and a human surveyor to estimate bird location relative to that transect. The third method, presented by Sebastián-González et al. (2018), involves manually recording any vocalizing individuals that are visually detected while simultaneously measuring distance with a laser rangefinder. A similar variant, presented by Darras et al. (2018), uses an ARU paired with a surveyor who measures distance to any visual detections of vocalizing individuals during a point count survey. The pros and cons of these methods are discussed at length in Yip et al. (2020).

Behavioural Context

Chapters 4-7 showed that understanding the behavioural context of species detections can deepen ecological inferences and open doors to new lines of inquiry. Behaviour-specific habitat modelling can be particularly important for highly mobile species because they are more likely to use spatially distinct habitats to fulfill their life history needs (Roever et al. 2013; Law & Dickman 1998; Frans et al. 2017) and because they spend substantial amounts of time moving through suboptimal areas where they can be detected (Marsh et al. 2014). Despite these known pitfalls, studies rarely compare presence-absence models to function-specific models (Roever et al. 2013; Frans et al. 2017; Brambilla & Saporetti 2014). For example, the home range habitat spatial predictions from Chapter 5 showed moderate suitability across the landscape, while the territorial habitat predictions were much better for pinpointing areas of high suitability.

Movement is another type of context that is lacking from recognizer data, but that could potentially be derived using the principles in Chapter 3 and Yip et al. (2020). If RSL and/or score are proportional to distance, then the change in either measurement between vocalizations should be a reasonable proxy for movement rate. Relative signal strength between stereo microphones could be used to rule out changes in RSL/score caused by changes in directionality of vocalization. Furthermore, changes in RSL/score between vocalizations could be used a parameter in a hidden Markov model (HMM) to predict behavioural context. HMMs are commonly used to identify behavioural patterns from animal tracking data in movement modelling (Joo et al. 2019), and could easily be extended to PAM.

Looking Forward: Multispecies Recognition

The field of automated recognition is transitioning from single species to multispecies recognizers as deep learning methods emerge that can be trained on an infinite number of classes (Stowell et al. 2018, Kahl et al. 2021). Two annual birdsong classification competitions are driving the development of bigger and better algorithms (Stowell et al. 2018; Kahl et al. 2020). Neural networks are so powerful that highly accurate multispecies classifiers can be built with small training datasets and pretrained networks (Stowell et al. 2018; Knight et al. 2019) (media feature: CBC Radio 2020). There are several large multispecies neural net recognizers currently available to the public. For example, BirdNET from the Cornell Lab of Ornithology, which is trained to recognize 984 species from North American and Europe, is even available as a smartphone app (Kahl et al. 2021).

As multispecies classifiers become mainstream and accessible, it continues to be important to think about the context of the species detection datasets they produce. The spatial context principles I demonstrated in this thesis may not necessarily apply to multispecies recognizers because classification probability differs between binary and categorical classifiers (Sokolova & Lapalme 2009). Where single species recognizers report a single probability from zero to one that a particular sound is the target species, multispecies recognizers report a probability from zero to one for each species they have been trained on, and those probabilities sum to one. Multispecies classifiers therefore do not typically use score thresholds. Instead, a particular species is reported as detected if it has the highest classification probability for that sound. Research should be conducted to determine whether classification probability remains a proxy for score in multispecies classification; it is possible that the probabilities of the other species in the classifier interfere with this relationship.

As ecology becomes more technologically-driven, there is a risk that ecologists will lose touch with the context of data they use (Hebblewhite & Haydon 2010). Although PAM and automated recognition have great potential to contribute to ecology and conservation, there is no replacement for time spent in the field observing the species or system of focus. My thesis and the questions herein were inspired by several thousand hours spent observing, monitoring, catching, tracking, and appreciating common nighthawks across North America. Given the demonstrated importance of context in my thesis, I strongly encourage all ecologists considering the use of PAM to ensure they keep one foot in the field because it's hard to understand context without first-hand experience.

A Call for Interdisciplinary Collaboration

To conclude, I come back to the value of collaboration. My thesis exposed a disciplinary gap in PAM between the computing scientists that design recognizer algorithms and the ecologists that apply them (Thessen 2016). I believe one of the drivers of this disciplinary gap is that success is measured by a recognizer's "accuracy" and thus algorithm designers are motivated solely by this one-dimensional goal. I recommend birdsong classification competitions consider a broader suite of purpose-driven challenges and evaluations that would encourage computing scientists to understand how algorithms are applied. For example, detection of a rare species, species richness estimation, abundance estimation, and call rate estimation. I further recommend closer working relationships between the two disciplines to reap benefits on both sides of the equation. For example, I read deeply through the computing science literature as part of my thesis and worked with a team of computing scientists and a linguist on a multispecies recognition project (Knight et al. 2019b). Those experiences and relationships greatly improved my understanding of deep learning, signal detection, and acoustic theory and played a strong role in shaping my thesis. The Ecoacoustics Lab at Queensland University of Technology is an excellent example of scientists and ecologists working in-house together to design purpose-driven algorithms (Burivalova et al. 2019).

Throwing biostatisticians into the collaborative mix would further benefit PAM because it would provide yet another perspective from the PAM pipeline. Given the increasing emphasis on statistical integration in ecology, input from experts in data integration on statistical assumptions would be particularly valuable when making decisions about training and using recognizers. For example, I showed in Chapter 3 of this thesis that recognizers that are not trained with clips recorded at close range do not satisfy the assumptions of density estimation. Biostatistical focus on continuing to develop methods to integrate PAM with other data types, especially large-scale community science projects like eBird will be key to expanding the temporal and spatial scale of how PAM is used (la Sorte et al. 2018). The use of statistical offsets to account for differences in survey area is a simple and elegant approach to integration that is gaining increasing traction and is applicable across a variety of analysis approaches (Matsuoka et al. 2012; Sòlymos et al. 2013). Other approaches also exist for the integration of PAM data for species distribution modelling (Isaac et al. 2019; Jr. et al. 2019) and density estimation (Doser et al. 2021; Sebastián-González et al. 2018).

The success of PAM and its ability to help solve global ecological crises is dependent upon the willingness of scientists to collaborate. Other disciplines that it would be helpful to have closer relationships with include the sound engineers that design ARUs, and linguists, whose work in speech classification (e.g., Siri, Alexa) is always several years ahead of bioacoustics. Ultimately, however, we need PAM researchers from across the globe to bring datasets together to answer some of the pressing questions in ecology and conservation. Understanding the health and maintenance of global pollinator and kelp forest diversity, the efficacy of varying protected area strategies, and effects of human settlement strategies, are just a few of the pressing questions in conservation that could be answered by pooling remotely sensed acoustic datasets (Sutherland et al. 2019, 2009). The ongoing culture shift towards open data and collaboration is extremely encouraging (Pannell et al. 2019; Aubin et al. 2020; Tenopir et al. 2015); it is a challenging but exciting time to be an ecologist.

Media Features Referenced

Struzik E (2017) Firestorm. Island Press, Washington, DC, USA.

CBC News (2017) Wildlife returns as forest regenerates after Fort McMurray wildfire. http://www.cbc.ca/news/canada/edmonton/wildlife-returns-as-forest-regenerates-after-fort-mcmurray-wildfire-1.3665041.

CBC Radio (2020) Hidden Frequencies: How the invisible powers of sound shape our lives and world. https://www.cbc.ca/radio/how-the-invisible-powers-of-sound-shape-our-lives-and-world-1.5753088.

American Birding Association (2021) Unravelling a nighthawk migration mystery. https://www.aba.org/mysterious-nighthawk-migrations-with-elly-knight/.

The Globe and Mail (2018) ’Little backpacks’: GPS used to track nighthawks' migration from northern Alberta to Brazil. https://www.theglobeandmail.com/canada/article-little-backpacks-gps-used-to-track-nighthawks-migration-from/.

Global News (2021) Learn more about the Canadian Nightjar Survey. https://globalnews.ca/video/7965717/learn-more-about-the-canadian-nightjar-survey/

References

Abadi M, Agarwal A, Barham P, Brevdo E, Chen Z, Citro C, Corrado G, Davis A, Dean J, Devin M, Ghemawat S, Goodfellow I, Harp A, Irving G, M I, Jozefowicz R, Jia Y, Kaiser L, Kudlur M, Levenberg J, Aavik T, Mané D, Schuster M, Monga R, Moore S, Murray D, Olah C, Shlens J, Steiner B, Sutskever I, Talwar K, Tucker P, Vanhoucke V, Vasudevan V, Viégas F, Vinyals O, Warden P, Wattenberg M, Wicke M, Yu Y, Zheng X (2015) Tensorflow: large-scale machine learning on heterogeneous systems. Software available from tensorflow.org.

Acevedo MA, Corrada-Bravo CJ, Corrada-Bravo H, Villanueva-Rivera LJ, Aide TM (2009) Automated classification of bird and amphibian calls using machine learning: A comparison of methods. Ecological Informatics 4:206–214.

Adams J, Pendlebury D, Potter R, Szomszor M (2019) Multi-authorship and research analytics. Institute for Scientific Information. www.webofsciencegroup.com/isi.

Aide TM, Corrada-Bravo C, Campos-Cerqueira M, Milan C, Vega G, Alvarez R (2013) Real-time bioacoustics monitoring and automated species identification. PeerJ 1:e103-19.

Akaike H (1987) Factor analysis and AIC. Psychometrika 52, 317–332.

Alberta Agriculture and Forestry (2020) Spatial Wildfire Data. https://wildfire.alberta.ca/resources/historical-data/spatial-wildfire-data.aspx

Aldridge H, Brigham RM (1991) Factors influencing foraging time in two aerial insectivores: the bird *Chordeiles minor* and the bat *Eptesicus fuscus*. Canadian Journal of Zoology 69:62-69.

Allen MC, Peters KA (2012) Nest Survival, Phenology, and Nest-Site Characteristics of Common Nighthawk in a New Jersey Pine Barrens Grassland. The Wilson Journal of Ornithology 124:113–118.

Alpac (2015) Alberta-Pacific FMA 2015 forest management plan. Alberta-Pacific Forest Industries Inc. https://open.alberta.ca/publications/alberta-pacific-fma-area-2015-forest-management-plan#summary

Amarasekare P, Possingham HP (2001) Patch Dynamics and Metapopulation Theory: the Case of Successional Species. Journal of Theoretical Biology 209:333–344.

Amrhein V, Korner P, Naguib M (2002) Nocturnal and diurnal singing activity in the nightingale: correlations with mating status and breeding cycle. Animal Behaviour 64:939–944.

Anich NM, Benson TJ, Bednarz JC (2009) Estimating Territory and Home-range Sizes: Do Singing Locations Alone Provide an Accurate Estimate of Space Use? Auk 126:626–634.

Armitage DW, Ober HK (2010) A comparison of supervised learning techniques in the classification of bat echolocation calls. Ecological Informatics 5:465–473.

Armstrong JT (1965) Breeding home range in the nighthawk and other birds: its evolutionary and ecological significance. Ecology 46:619–629.

Arneill GE, Critchley EJ, Wischnewski S, Jessopp MJ, Quinn JL (2020) Acoustic activity across a seabird colony reflects patterns of within-colony flight rather than nest density. Ibis 162:416–428.

Aubin I, Cardou F, Boisvert-Marsh L, Garnier E, Strukelj M, Munson AD (2020) Managing data locally to answer questions globally: The role of collaborative science in ecology. Journal of Vegetation Science 31:509–517.

Balantic CM, Donovan TM (2019) Statistical learning mitigation of false positives from template-detected data in automated acoustic wildlife monitoring. Bioacoustics 29:296-321.

Bardeli R, Wolff D, Kurth F, Koch M, Tauchert KH, Frommolt KH (2010) Detecting bird sounds in a complex acoustic environment and application to bioacoustic monitoring. Pattern Recognition Letters 31:1524–1534.

Barg JJ, Jones J, Roberston RJ (2005) Describing breeding territories of migratory passerines: suggestions for sampling, choice of estimator, and delineation of core areas. Journal of Applied Ecology 74:139–149.

Barker NKS, Fontaine PC, Cumming SG, Stralberg D, Westwood A, Bayne EM, Sólymos P, Schmiegelow FKA, Song SJ, Rugg DJ (2015) Ecological monitoring through harmonizing existing data: Lessons from the boreal avian modelling project. Wildlife Society Bulletin 39:480–487.

Barré K, Viol IL, Julliard R, Pauwels J, Newson SE, Julien J-F, Claireau F, Kerbiriou C, Bas Y (2019) Accounting for automated identification errors in acoustic surveys Gutierrez, VZ, editor. Methods in Ecology and Evolution 3:992–18.

Bart J, Schoultz JD (1984) Reliability of singing bird surveys: changes in observer efficiency with avian density. The Auk 101:307–318.

Bas JM, Pons P, Gómez C (2016) Home range and territory of the Sardinian Warbler *Sylvia melanocephalain* Mediterranean shrubland. Bird Study 52:137–144.

Bates D, Mächler M, Bolker B, Walker SC (2015). Fitting linear mixed-effects models using lme4. Journal of Statistical Software 67:1–48.

Berzins L (2020) Research, Conservation and Outreach Priorities for Conserving Aerial Insectivore Populations in Canada. Report from March 2020 Aerial Insectivore Workshop, Saskatoon, SK, Canada. 180 pp.

Beyer HL, Haydon DT, Morales JM, Frair JL, Hebblewhite M, Mitchell M, Matthiopoulos J (2010) The interpretation of habitat preference metrics under use-availability designs. Philosophical Transactions of the Royal Society B: Biological Sciences 365:2245–2254.

Bivand R, Rundel C (2020) rgeos: Interface to Geometry Enging – Open Source ('GEOS'). R package version 0.5-5. https://CRAN.R-project.org/package=rgeos

Bivand R, Keitt T, Rowlingson B (2021 rgdal: Bindings for the 'Geospatial" Data Abstraction Library. R package version 1.5-23. https://CRAN.R-project.org/package=rgdal

Blumstein DT, Mennill DJ, Clemins P, Girod L, Yao K, Patricelli G, Deppe JL, Krakauer AH, Clark C, Cortopassi KA, Hanser SF, McCowan B, Ali AM, Kirschel ANG (2011) Acoustic monitoring in terrestrial environments using microphone arrays: applications, technological considerations and prospectus. Journal of Applied Ecology 48:758–767.

Bognounou F, Venier LA, Wilgenburg SLV, Aubin I, Candau J-N, Arsenault A, Hebert C, Ibarzabal J, Song SJ, Grandpré LD (2021) Early avian functional assemblages after fire, clearcutting, and post-fire salvage logging in North American forests. Canadian Journal of Forest Research 51:393–407.

Bonan GB, Shugart HH (1989) Environmental factors and ecological processes in boreal forests. Annual Reviews of Ecology, Evolution, and Systematics 20:1-28.

Borker AL, McKown MW, Ackerman JT, Eagles-Smith CA, Tershy BR, Croll DA (2014) Vocal Activity as a Low Cost and Scalable Index of Seabird Colony Size. Conservation Biology 28:1100–1108.

Bormann FH, Likens GE (1979) Catastrophic Disturbance and the Steady State in Northern Hardwood Forests: A new look at the role of disturbance in the development of forest ecosystems suggests important implications for land-use policies. American Scientist 67:660–669.

Boucaud ICA, Perez EC, Ramos LS, Griffith SC, Vignal C (2017) Acoustic communication in zebra finches signals when mates will take turns with parental duties. Behavioral Ecology 28:645–656.

Bowler DE, Nilsen EB, Bischof R, Hara RBO x02019, Yu TT, Oo T, Aung M, Linnell JDC, Linnell JDC (2019) Integrating data from different survey types for population monitoring of an endangered species: the case of the Eld's deer. Scientific Reports 9:7766.

Bowles JH (1921) Nesting habits of the nighthawk at Tacoma, Wash. The Auk 38:2013–2217.

Boyce MS, McDonald LL (1999) Relating populations to habitats using resource selection functions. Trends in Ecology & Evolution 14:268–272.

Brambilla M, Saporetti F (2014) Modelling distribution of habitats required for different uses by the same species: Implications for conservation at the regional scale. Biological Conservation 174:39–46.

Brauer CL, Donovan TM, Mickey RM, Katz J, Mitchell BR (2016) A comparison of acoustic monitoring methods for common anurans of the northeastern United States. Wildlife Society Bulletin 40:140–149.

Brigham RM (1990) Prey selection by big brown bats (*Eptesicus fuscus*) and common nighthawks (*Chordeiles minor*). American Midland Naturalist 124:73–80.

Brigham RM, Barclay RMR (1995) Prey detection by common nighthawks: Does vision impose a constraint? Ecoscience 2:276-279.

Brigham RM, Barclay RMR (1992) Lunar Influence on Foraging and Nesting Activity of Common Poorwills (*Phalaenoptilus Nuttallii*). The Condor 109:315–320.

Brigham RM, Fenton MB (1991) Convergence in foraging strategies by two morphologically and phylogenetically distinct nocturnal aerial insectivores. Journal of Zoology 223:475–489.

Brigham RM, Ng J, Poulin RG, Grindal SD (2011) Common Nighthawk (*Chordeiles minor*) Poole, A, ed. The Birds of North America Online.

Brown JH (1984) On the Relationship between Abundance and Distribution of Species. The American Naturalist 124:255–279.

Büchi L, Vuilleumier S (2016) Ecological strategies in stable and disturbed environments depend on species specialisation. Oikos 125:1408–1420.

Buckland ST, Anderson DR, Burnham KP, Laake JL (2005) Distance Sampling. In Encyclopedia of Biostatistics, P. Armitage and T. Colton, eds.

Buckland ST, Rexstad EA, Marques TA, Oedekoven CS (2015) Distance sampling: methods and applications. Springer, Cham, Switzerland.

Burivalova Z, Purnomo, Wahyudi B, Boucher TM, Ellis P, Truskinger A, Towsey M, Roe P, Marthinus D, Griscom B, Game ET (2019) Using soundscapes to investigate homogenization of tropical forest diversity in selectively logged forests. Journal of Applied Ecology 56:2493–2504.

Burnham K. P., and D. R. Anderson (2003) Model selection and multimodel inference: a practical information-theoretic approach. Springer Science & Business Media, New York, NY, USA.

Burt WH (1943) Territoriality and home range concepts as applied to mammals. The Auk.

Caccamise DF (1974) Competitive relationships of the common and lesser nighthawks. The Condor 76:1–20.

Calenge C (2006) The package adehabitat for the R software: a tool for the analysis of space and habitat use by animals. Ecological Modelling 197:516-519.

Campbell M, Francis CM (2012) Using microphone arrays to examine effects of observers on birds during point count surveys. Journal of Field Ornithology 83:391–402.

Campbell P, Pasch B, Pino JL, Crino OL, Phillips M (2010) Geographic variation in the songs of neotropical singing mice: testing the relative importance of drift and local adaptation. Evolution 60:1874.

Campos-Cerqueira M, Aide TM (2016) Improving distribution data of threatened species by combining acoustic monitoring and occupancy modelling. Methods in Ecology and Evolution 7:1340–1348.

Canadian Centre for Remote Sensing (2010) 2020 Land Cover of Canada. https://open.canada.ca/data/en/dataset/c688b87f-e85f-4842-b0e1-a8f79ebf1133

Cardinale BJ, Duffy JE, Gonzalez A, Hooper DU, Perrings C, Venail P, Narwani A, Mace GM, Tilman D, Wardle DA, Kinzig AP, Daily GC, Loreau M, Grace JB, Larigauderie A, Srivastava DS, Naeem S (2012) Biodiversity loss and its impact on humanity. Nature 486:59–67.

Catchpole CK, Slater P (2008) Bird song: biological themes and variations. Cambridge University Press, New York, NY, USA.

Chambert T, Waddle JH, Miller DAW, Walls SC, Nichols JD (2017) A new framework for analysing automated acoustic species detection data: Occupancy estimation and optimization of recordings post-processing. Methods in Ecology and Evolution 9:560–570.

Charif RA, Waack AM, Strickman LM (2010) Raven Pro 1.4 User’s Manual. 379 pp.

Chen HY, Popadiouk RV (2002) Dynamics of North American boreal mixedwoods. Environmental Reviews 10:137–166.

Chu W, Blumstein DT (2011) Noise robust bird song detection using syllable pattern-based hidden Markov models. In: ICASSP 2011 - 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 345–348

Cleere N (2010) Nightjars, Potoos, Frogmouts, Oilbird and Owlet-nightjars of the World. Princeton University Press, Princeton, NJ, USA.

Cohen J (1960) A coefficient of agreement for nominal scales. Educational and Psychological Measurement 20:37-46.

COSEWIC (2017) Common Nighthawk threats calculator draft for post-call review 4 March 2017.

COSEWIC (2018) COSEWIC assessment and status report on the Common Nighthawk *Chordeiles minor* in Canada. Committee on the Status of Endangered Wildlife in Canada. Ottawa, ON, Canada. 50 pp.

Crump PS, Houlahan J (2017) Designing better frog call recognition models. Ecology and Evolution 7:3087–3099.

Cumming S, Stralberg D, Lefevre KL, Sólymos P, Bayne EM, Fang S, Fontaine T, Mazerolle D, Schmiegelow FKA, Song SJ (2013) Climate and vegetation hierarchically structure patterns of songbird distribution in the Canadian boreal region. Ecography 37:137–151.

Darras K, Furnas B, Fitriawan I, Mulyani Y, Tscharntke T (2018) Estimating bird detection distances in sound recordings for standardizing detection ranges and distance sampling Methods in Ecology and Evolution 9:1928–1938.

Davis GE, Baumgartner MF, Bonnell JM, Bell J, Berchok C, Thornton JB, Brault S, Buchanan G, Charif RA, Cholewiak D, Clark CW, Corkeron P, Delarue J, Dudzinski K, Hatch L, Hildebrand J, Hodge L, Klinck H, Kraus S, Martin B, Mellinger DK, Moors-Murphy H, Nieukirk S, Nowacek DP, Parks S, Read AJ, Rice AN, Risch D, Širović A, Soldevilla M, Stafford K, Stanistreet JE, Summers E, Todd S, Warde A, Parijs SMV (2017) Long-term passive acoustic recordings track the changing distribution of North Atlantic right whales (*Eubalaena glacialis*) from 2004 to 2014. Scientific Reports 7:13460.

Davis J, Goadrich M (2006) The relationship between Precision-Recall and ROC curves. In Proceedings of the 23rd International Conference on Machine Learning, Pittsburgh, PA, USA.

Dawson K., Efford MG (2009) Bird population density estimated from acoustic signals. Journal of Applied Ecology 46:1201–1209.

Deans AM, Malcolm JR, Smith SM, Bellocq MI (2005) Edge effects and the responses of aerial insect assemblages to structural-retention harvesting in Canadian boreal peatland forests. Forest Ecology and Management 204:249–266.

De'Ath G, Fabricius KE (2000) Classification and regression trees: a powerful yet simple technique for ecological data analysis. Ecology 81:3178–3192.

Demšar J (2006) Statistical comparisons of classifiers over multiple data sets. Journal of Machine learning research 7:1–30.

Derryberry EP, Seddon N, Derryberry GE, Claramunt S, Seeholzer GF, Brumfield RT, Tobias JA (2018) Ecological drivers of song evolution in birds: Disentangling the effects of habitat and morphology. Ecology and Evolution 8:1890–1905.

Desrochers A, Renaud C, Hochachka WM, Cadman M (2010) Area-sensitivity by forest songbirds: theoretical and practical implications of scale-dependency. Ecography 33:921–931.

Dexter RW (1961) Further studies on nesting of the Common Nighthawk. Bird-banding 32:79-85.

Didham RK, Barbero F, Collins CM, Forister ML, Hassall C, Leather SR, Packer L, Saunders ME, Stewart AJA (2019) Spotlight on insects: trends, threats and conservation challenges. Insect Conservation and Diversity 13:99–102.

Dietterich TG (1998) Approximate statistical tests for comparing supervised classification learning algorithms. Neural computation 10:1895–1923.

Digby A, Towsey M, Bell BD, Teal PD (2013) A practical comparison of manual and autonomous methods for acoustic monitoring. Methods in Ecology and Evolution 4:675–683.

Doser JW, Finley AO, Weed AS, Zipkin EF (2021) Integrating automated acoustic vocalization data and point count surveys for estimation of bird abundance. Methods in Ecology and Evolution 12:1040–1049.

Drake KL, Frey M, Hogan D, Hedley R (2016) Using digital recordings and sonogram analysis to obtain counts of yellow rails. Wildlife Society Bulletin 40:346–354.

Driscoll DA, Lindenmayer DB, Bennett AF, Bode M, Bradstock RA, Cary GJ, Clarke MF, Dexter N, Fensham R, Friend G, Gill M, James S, Kay G, Keith DA, MacGregor C, Russell-Smith J, Salt D, Watson JEM, Williams RJ, York A (2010) Fire management for biodiversity conservation: Key research questions and our capacity to answer them. Biological Conservation 143:1928–1939.

Duan R-Y, Kong X-Q, Huang M-Y, Fan W-Y, Wang Z-G (2014) The Predictive Performance and Stability of Six Species Distribution Models. PLoS ONE 9:e112764-8.

Duan S, Zhang J, Roe P, Wimmer J, Dong X (2013) Timed Probabilistic Automaton: Bridge between Automatic Species Recognition. In: 25th Innovative Applications of Artificial Intelligence Conference.

Ehnes M, Foote JR (2014) Comparison of autonomous and manual recording methods for discrimination of individually distinctive Ovenbird songs. Bioacoustics 24:111–121.

Elgin AS, Clark RG, Morrissey CA (2020) Tree Swallow selection for wetlands in agricultural landscapes predicted by central-place foraging theory. Condor 67:311–12.

Elith J, Leathwick JR, Hastie T (2008) A working guide to boosted regression trees. Journal of Animal Ecology 77:802–813.

Environment Canada (2016a) Recovery Strategy for the Common Nighthawk (*Chordeiles minor*) in Canada. Species at Risk Act Recovery Strategy Series. Environment Canada, Ottawa, ON, Canada. 49 pp.

Environment Canada (2016b) Recovery Strategy for the Olive-sided Flycatcher (*Contopus cooperi*) in Canada. Species at Risk Act Recovery Strategy Series. Environment Canada, Ottawa, ON, Canada. 52 pp.

Erbs F, Elwen SH, Gridley T (2017) Automatic classification of whistles from coastal dolphins of the southern African subregion. The Journal of the Acoustical Society of America 141:2489–2500.

Evens R, Beenaerts N, Neyens T, Witters N, Smeets K, Artois T, Artois (2018) Proximity of breeding and foraging areas affects foraging effort of a crepuscular, insectivorous bird. Scientific Reports 8:3008.

Evens R, Conway GJ, Henderson IG, Cresswell B, Jiguet F, Moussy C, Sénécal D, Witters N, Beenaerts N, Artois T (2017) Migratory pathways, stopover zones and wintering destinations of Western European Nightjars *Caprimulgus europaeus.* Ibis 159:680-686.

Evens R, Jacot A, Artois T, Ulenaers E, Neyens T, Rappaz L, Theux C, Pradervand JN (2020) Improved ecological insights commission new conservation targets for a crepuscular bird species. Animal Conservation 89:1–13

Farley SS, Dawson A, Goring SJ, Williams JW (2018) Situating Ecology as a Big-Data Science: Current Advances, Challenges, and Solutions. BioScience 68:563–576.

Farrell CE, Fahrig L, Mitchell G, Wilson S (2019) Local habitat association does not inform landscape management of threatened birds. Landscape Ecology 34:1313-1327.

Farrell CE, Wilson S, Mitchell G (2017) Assessing the relative use of clearcuts, burned stands, and wetlands as breeding habitat for two declining aerial insectivores in the boreal forest. Forest Ecology and Management 386:62–70.

Fieberg J, O'Kochanny C (2005) Quantifying home-range overlap: the importance of the utilization distribution. Journal of Wildlife Management 69:1346–1359

Fielding AH, Bell JF (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental Conservation 24:38–49

Fisher JT, Anholt B, Volpe JP (2011) Body mass explains characteristic scales of habitat selection in terrestrial mammals. Ecology and Evolution 1:517–528.

Fisher RJ, Fletcher QE, Willis C (2004) Roost selection and roosting behavior of male common nighthawks. The American Midland Naturalist 151:79–87.

Fiske I, Chandler R (2011) unmarked: An R Package for Fitting Hierarchical Models of Wildlife Occurrence and Abundance. Journal of Statistical Software 43:1-23.

Foley G (2018) Habitat use by Common Nighthawks (*Chordeiles minor*) in Canada's boreal forest. M.Sc. thesis, University of Regina, Regina, SK, Canada.

Forcey GM, Anderson JT, Ammer FK, Whitmore RC (2006) Comparison of Two Double-Observer Point-Count Approaches for Estimating Breeding Bird Abundance. Journal of Wildlife Management 70:1674–1681.

Frans VF, Augé AA, Edelhoff H, Erasmi S, Balkenhol N, Engler JO (2017) Quantifying apart what belongs together: A multi-state species distribution modelling framework for species using distinct habitats. Methods in Ecology and Evolution 54:412–11.

Frantz MW, Aldinger KR, Wood PB, Duchamp J, Nuttle T, Larkin AV and JL (2016) Space and Habitat Use of Breeding Golden-winged Warblers in the Central Appalachian Mountains. Studies in Avian Biology 49:81-94.

Furnas BJ, Callas RL (2015) Using automated recorders and occupancy models to monitor common forest birds across a large geographic region. The Journal of Wildlife Management 79:325–337.

Galán-Acedo C, Arroyo-Rodríguez V, Cudney-Valenzuela SJ, Fahrig L (2019) A global assessment of primate responses to landscape structure. Biological Reviews 94:1605–1618.

Galán-Acedo C, Arroyo-Rodríguez V, Estrada A, Ramos-Fernández G (2018) Drivers of the spatial scale that best predict primate responses to landscape structure. Ecography 41:2027–2037.

Ganchev TD, Jahn O, Marques MI, Figueiredo JM de, Schuchmann K-L (2015) Automated acoustic detection of *Vanellus chilensis lampronotus*. Expert Systems With Applications 42:6098–6111.

Garnett ST, Burgess ND, Fa JE, Fernández-Llamazares Á, Molnár Z, Robinson CJ, Watson JEM, Zander KK, Austin B, Brondizio ES, Collier NF, Duncan T, Ellis E, Geyle H, Jackson MV, Jonas H, Malmer P, McGowan B, Sivongxay A, Leiper I (2018) A spatial overview of the global importance of Indigenous lands for conservation. Nature Sustainability 1:369–374.

Génier CSV, Guglielmo CG, Mitchell GW, Falconer M, Hobson KA (2021) Nutritional consequences of breeding away from riparian habitats in Bank Swallows: new evidence from multiple endogenous markers. Conservation Physiology 9:e101218-12.

Gibb R, Browning E, Glover-Kapfer P, Jones KE (2018) Emerging opportunities and challenges for passive acoustics in ecological assessment and monitoring. Methods in Ecology and Evolution 3:992–17.

Gillespie D, Caillat M, Gordon J, White P (2013) Automatic detection and classification of odontocete whistles. The Journal of the Acoustical Society of America 134:2427–2437.

Government of Canada (2016) Critical habitat identification toolbox: Species at Risk Act guidance Version 2.3. https://www.canada.ca/en/environment-climate-change/services/species-risk-public-registry/critical-habitat-descriptions/identification-toolbox-guidance.html

Grau J., I. Grosse, and J. Keilwagen. (2015) PRROC: computing and visualizing precision-recall and receiver operating characteristic curves in R. Bioinformatics 3:2595–2597.

Griffith SC, Owens I, Thuman KA (2002) Extra pair paternity in birds: a review of interspecific variation and adaptive function. Journal of Applied Ecology 11:2195–2212.

de Groot WJ, Flannigan MD, Cantin AS (2013) Climate change impacts on future boreal fire regimes. Forest Ecology and Management 294:35–44.

Gross AO (1940) Eastern Nighthawk. In: Bent, AC, editor. Life Histories of North American Cuckoos, Goatsuckers, Hummingbirds, and their Allies. Orders Psittaciformes, Cuculiformes, Trogoniformes, Coraciiformes, Caprimulgiformes, and Micropodiiformes. 506 pp.

Guerrero AM, Mcallister RRJ, Wilson KA (2015) Achieving Cross-Scale Collaboration for Large Scale Conservation Initiatives. Conservation Letters 8:107–117.

Haché S, Sòlymos P, Fontaine PC, Bayne EM, Cumming S, Schmiegelow FKA, Stralberg D (2014) Analyses to support critical habitat identification for Canada Warbler, Olive-sided Flycatcher, and Common Nighthawk Final Report. Boreal Avian Modelling Project, Edmonton, AB, Canada.

Hafner SD, Katz J (2015) A short introduction to acoustic template matching with monitoR. https://cran.r-project.org/web/packages/monitoR/vignettes/monitoR_QuickStart.pdf.

Hagar J, Howlin S, Ganio L (2004) Short-term response of songbirds to experimental thinning of young Douglas-fir forests in the Oregon Cascades. Forest Ecology and Management 199:333–347.

Hampton SE, Strasser CA, Tewksbury JJ, Gram WK, Budden AE, Batcheller AL, Duke CS, Porter JH (2013) Big data and the future of ecology. Frontiers in Ecology and the Environment 11:156–162.

Hannah KG, Leston LL, Knight EC, Weeber R (In review) In the twilight zone: patterns in Common Nighthawk (*Chordeiles minor*) acoustic signals during the breeding season and recommendations for surveys. Submitted to Avian Conservation and Ecology ACE-ECO-2020-1698.

Harris CM (1966) Absorption of Sound in Air versus Humidity and Temperature. The Journal of the Acoustical Society of America 40:148–159.

Harrison A (2006) Who's Who in Conservation Biology—an Authorship Analysis. Conservation Biology 20:652–657.

Haselmayer J, Quinn JS (2000) A comparison of point counts and sound recording as bird survey methods in Amazonian southeast Peru. The Condor 102:887.

He F, Gaston K, Wu J (2016) On species occupancy-abundance models. Écoscience 9:119–126.

Hebblewhite M, Haydon DT (2010) Distinguishing technology from biology: a critical review of the use of GPS telemetry data in ecology. Philosophical Transactions of the Royal Society B: Biological Sciences 365:2303–2312.

Hedley RW, Wilson SJ, Yip DA, Li K, Bayne EM (2020) Distance truncation via sound level for bioacoustic surveys in patchy habitat. Bioacoustics 30:1–21.

Heidorn PB (2008) Shedding Light on the Dark Data in the Long Tail of Science. Library Trends 57:280–299.

Heinicke S, Kalan AK, Wagner OJJ, Mundry R, Lukashevich H, Kühl HS (2015) Assessing the performance of a semi-automated acoustic monitoring system for primates Jones. Methods in Ecology and Evolution 6:753–763.

Hijmans RJ. (2020) raster: Geographic Data Analysis and Modeling. R package version 3.4-5. https://CRAN.R-project.org/package=raster

Hijmans RJ, Phillips S, Lethwick J, Elith J (2020). dismo: Species Distribution Modeling. R package version 1.3-3. https://CRAN.R-project.org/package=dismo

Hobson KA, Schieck J (1999) Changes in bird communities in boreal mixedwood forest: harvest and wildfire effects over 30 years. Ecological Applications 9:849–863.

Holland JD, Bert DG, Fahrig L (2004) Determining the spatial scale of species' response to habitat. BioScience 54:227.

Holmes SB, McIlwrick KA, Venier LA (2014) Using automated sound recording and analysis to detect bird species-at-risk in southwestern Ontario woodlands. Wildlife Society Bulletin 38:591–598.

Holyoak DT (2001) Nightjars and their Allies. Oxford University Press, Oxford, UK.

Hudson M-AR, Francis CM, Campbell KJ, Downes CM, Smith AC, Pardieck KL (2017) The role of the North American Breeding Bird Survey in conservation. Condor 119:526–545.

Hunter ML (1993) Natural fire regimes as spatial models for managing boreal forests. Biological Conservation 65:115–120.

Hutto RL (2002) Composition of Bird Communities Following Stand-Replacement Fires in Northern Rocky Mountain (U.S.A.) Conifer Forests. Conservation Biology 9:1041–1058.

Bird Life International (2017) Bird species distribution maps of the world.

IPBES (2019) Global assessment report of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services, Brondízio, E. S., Settele, J., Díaz, S., Ngo, H. T. (eds). IPBES secretariat, Bonn, Germany. 1144 pp.

IPCC (2014) Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer eds.]. IPCC, Geneva, Switzerland. 151 pp.

Isaac NJB, Jarzyna MA, Keil P, Dambly LI, Boersch-Supan PH, Browning E, Freeman SN, Golding N, Guillera-Arroita G, Henrys PA, Jarvis S, Lahoz-Monfort J, Pagel J, Pescott OL, Schmucki R, Simmonds EG, O'Hara RB (2019) Data Integration for Large-Scale Models of Species Distributions. Trends in Ecology & Evolution 35:56-67.

IUCN (2012) IUCN Red List Categories and Criteria: Version 3.1. Second edition. IUCN.

Jackson HB, Fahrig L (2015) Are ecologists conducting research at the optimal scale? Global Ecology and Biogeography 24:52–63.

Jackson HB, Fahrig L (2012) What size is a biologically relevant landscape? Landscape Ecology 27:929–941.

Jahn O, Ganchev TD, Marques MI, Schuchmann K-L (2017) Automated Sound Recognition Provides Insights into the Behavioral Ecology of a Tropical Bird. PLoS ONE 12:e0169041-29.

Jeliazkov A, Bas Y, Kerbiriou C, Julien J-F, Penone C, Viol IL (2016) Large-scale semi-automated acoustic monitoring allows to detect temporal decline of bush-crickets. Global Ecology and Conservation 6:208–218.

Jennifer J-A (2015) Breeding ecology of the Florida Nighthawk (*Chordeiles minor chapmani*). Florida Field Naturalist 43:12–18.

Jennings N, Parsons S, Pocock MJO (2008) Human vs. machine: identification of bat species from their echolocation calls by humans and by artificial neural networks. Canadian Journal of Zoology 86:371–377.

Jirinec V, Varian CP, Smith CJ, Leu M (2016) Mismatch between diurnal home ranges and roosting areas in the Wood Thrush (*Hylocichla mustelina*): Possible role of habitat and breeding stage. Auk 133:1–12.

Johnson MD (2007) Measuring habitat quality: a review. The Condor 109:489–504.

Johst K, Brandl R, Eber S (2002) Metapopulation persistence in dynamic landscapes: the role of dispersal distance. Oikos 98:263-270.

Joo R, Boone ME, Clay TA, Patrick SC, Trullas SC, Basille M (2019) Navigating through the rpackages for movement. Journal of Animal Ecology 58:227–20.

Joshi KA, Mulder RA, Rowe KMC (2017) Comparing manual and automated species recognition in the detection of four common south-east Australian forest birds from digital field recordings. Emu-Austral Ornithology 117:233–246.

Kahl S, Clapp M, Hopping W, Goëau H, Glotin H, Planqué R, Vellinga W-P, Joly A (2020) Bird sound recognition in complex acoustic environments. In: CLEF 2020 - 11th International Conference of the Cross-Language Evaluation Forum for European Languages, Thessaloniki, Greece.

Kahl S, Wood CM, Eibl M, Klinck H (2021) BirdNET: A deep learning solution for avian diversity monitoring. Ecological Informatics 61:101236.

Kalan AK, Mundry R, Wagner OJJ, Heinicke S, Boesch C, Kühl HS (2015) Towards the automated detection and occupancy estimation of primates using passive acoustic monitoring. Ecological Indicators 54:217–226.

Katz J, Hafner SD, Donovan T (2016) Assessment of Error Rates in Acoustic Monitoring with the R package monitoR. Bioacoustics 25:177–196.

Kingma D, Ba J (2014) Adam: A method for stochastic optimization. In 3rd International Conference for Learning Representations, San Diego, CA, USA.

Knaggs MD (2018) Effects of burn severity and time since fire on songbird communities in the northern boreal forest. M.Sc. thesis, University of Alberta, Edmonton, AB, Canada.

Knight EC, Hannah K, Foley G, Scott C, Brigham RM, Bayne EM (2017) Recommendations for acoustic recognizer performance assessment with application to five common automated signal recognition programs. Avian Conservation and Ecology 12:14.

Knight EC, Bayne EM (2018) Classification threshold and training data affect the quality and utility of focal species data processed with automated audio recognition software. Bioacoustics 28:539-554.

Knight EC, Ng JW, Mader CE, Brigham RM, Bayne EM (2018) An inordinate fondness for beetles: first description of Common Nighthawk (*Chordeiles minor*) diet in the boreal biome. The Wilson Journal of Ornithology 130:525–531.

Knight EC, Hannah KC, Brigham RM, McCracken J, Falardeau G, Julien M-F, Guénette J-S, Manthorne A, Sidler A, Foley GJ, McKnight J, Laurent KS, Sinclair P, Pankratz R (2019a) Canadian Nightjar Survey Protocol. http://wildresearch.ca/wp-content/uploads/2019/05/National-Nightjar-Survey-Protocol-WildResearch-2019.pdf

Knight EC, Hernandez SP, Bayne EM, Bulitko V, Tucker BV (2019b) Pre-processing spectrogram parameters improve the accuracy of bioacoustic classification using convolutional neural networks. Bioacoustics 29:337-355.

Knight EC, Sòlymos P, Scott C, Bayne EM (2020) Validation prediction: a flexible protocol to increase efficiency of automated acoustic processing for wildlife research. Ecological Applications 30:e02140.

Knight EC., Harrison A, Scarpignato AL, Wilgenburg SLV, Bayne EM, Ng JW, Angell E, Bowman R, Brigham RM, Drolet B, Easton WE, Forrester TR, Foster JT, Haché S, Hannah KC, Hick KG, Ibarzabal J, Imlay TL, Mackenzie SA, Marsh A, McGuire LP, Newberry GN, Newstead D, Sidler A, Sinclair PH, Stephens JL, Swanson DL, Tremblay JA, Marra PP (2021a) Comprehensive estimation of spatial and temporal migratory connectivity across the annual cycle to direct conservation efforts. Ecography 44:665-679.

Knight Elly C, Smith AC, Brigham RM, Bayne EM (2021b) Combination of targeted monitoring and Breeding Bird Survey data improves population trend estimation and species distribution modeling for the Common Nighthawk. Ornithological Applications 123:duab005.

Knight, EC, Brigham RM, Bayne EM (In review) The Big Boom Theory: Common Nighthawks use wing-boom displays to defend nesting territories. Ornithology ORNITH-21-054.

Knowles JE and Frederick C (2020) merTools: Tools for analyzing mixed effect regression models. R package version 0.5.2. https://CRAN.R-project.org/package=merTools.

Kock N (2007) Emerging e-Collaboration Concepts and Applications. Advances in E-Collaboration. CyberTech Publishing, Hershey, PA, USA.

Koops H. V., J. Van Balen, and F. Wiering (2014) A deep neural network approach to the lifeclef 2014 bird task. CLEF2014 Working Notes 1180:634-642.

Kramer GR, Chalfoun AD (2012) Growth Rate and Relocation Movements of Common Nighthawk (*Chordeiles minor*) Nestlings in Relation to Age. The Wilson Journal of Ornithology 124:793–797.

Kranstauber B, Cameron A, Weinzerl R, Fountain T, Tilak S, Wikelski M, Kays R (2011) The Movebank data model for animal tracking. Environmental Modelling & Software 26:834–835.

Kuuluvainen T, Grenfell R (2012) Natural disturbance emulation in boreal forest ecosystem management — theories, strategies, and a comparison with conventional even-aged management. Canadian Journal of Forest Research 42:1185–1203.

La Sort F, Lepczyk CA, Burnett JL, Hurlbert AH, Tingley MW, Zuckerberg B (2018) Opportunities and challenges for big data ornithology. The Condor 120:414–426.

Lambert D (1992) Zero-Inflated Poisson Regression, with an Application to Defects in Manufacturing. Technometrics 34:1

Lankau H (2015) Autonomous Recording Unit (ARU) Deployment Protocol. University of Alberta, Edmonton, AB, Canada.

Law BS, Dickman CR (1998) The use of habitat mosaics by terrestrial vertebrate fauna: implications for conservation and management. Biodiversity and Conservation 7:323-333.

Lawrence TB, Gardner N, Bunnefeld L, Bunnefeld N, Willis SG, Dent DH (2019) Guidelines for the use of acoustic indices in environmental research. Methods in Ecology and Evolution 26:32–12.

Lawton JH (1999) Are there general laws in ecology? Oikos 84:177–192

LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521:436–444.

Leonard TD, Taylor PD, Warkentin IG (2008) Landscape structure and spatial scale affect space use by songbirds in naturally patchy and harvested boreal forests. The Condor 110:467–481.

Levin SA (1992) The problem of pattern and scale in ecology: the Robert H. MacArthur award lecture. Ecology 73:1943-1967.

Lindenmayer DB, Westgate MJ, Scheele BC, Foster CN, Blair DP (2019) Key perspectives on early successional forests subject to stand-replacing disturbances. Forest Ecology and Management 454:117656–10.

Linhart P, Šálek M (2017) The assessment of biases in the acoustic discrimination of individuals. PLoS ONE 12:e0177206-16.

Linnes C (2020) Embracing the Challenges and Opportunities of Change Through Electronic Collaboration. International Journal of Information Communication Technologies and Human Development 12:37–58.

Lohnes R (2010) Nest site selection and nest thermal properties of Common Nighthawks on the tallgrass prairie of Kansas. M.Sc. thesis, Cornell University, Ithaca, NY, USA.

Long JN (2009) Emulating natural disturbance regimes as a basis for forest management: A North American view. Forest Ecology and Management 257:1868–1873.

Luther DA, Derryberry EP (2012) Birdsongs keep pace with city life: changes in song over time in an urban songbird affects communication. Animal Behaviour 83:1059–1066.

MacKenzie DI, Nichols JD, Lachman GB, Droege S (2002) Estimating site occupancy rates when detection probabilities are less than one. Ecology 83:2248–2255.

MacKenzie DI, Nichols JD, Royle JA, Pollock KH, Bailey LL, Hines JE (2006) Occupancy Estimation and Modeling. Academic Press, Cambridge, MA, USA.

MacKenzie DI, Royle JA (2005) Designing occupancy studies: general advice and allocating survey effort. Journal of Applied Ecology 42:1105–1114.

Maher CR, Lott DF (2000) A review of ecological determinants of territoriality within vertebrate species. American Midland Naturalist 143:1-29.

Maher CR, Lott DF (1995) Definitions of territoriality used in the study of variation in vertebrate spacing systems. Animal Behaviour 49:1581-1597.

Manly B, McDonald LL, Thomas DL, McDonald TL, Erickson WP (2002) Resource selection by animals. Second Edition. Kluwer Academic Publishers, Boston, MA, USA.

Marques TA, Thomas L, Martin SW, Mellinger DK, Ward JA, Moretti DJ, Harris D, Tyack PL (2012) Estimating animal population density using passive acoustics. Biological Reviews 88:287–309.

Marques TA, Thomas L, Ward J, DiMarzio N, Tyack PL (2009) Estimating cetacean population density using fixed passive acoustic sensors: An example with Blainville's beaked whales. The Journal of the Acoustical Society of America 125:1982–1994.

Marsh A, Bayne EM, Wellicome TI (2014) Using vertebrate prey capture locations to identify cover type selection patterns of nocturnally foraging Burrowing Owls. Ecological Applications 24:950–959.

MARSH H, SINCLAIR DF (1989) Correcting for Visibility Bias in Strip Transect Aerial Surveys of Aquatic Fauna. Journal of Wildlife Management 53:1017–1024.

Martin AE (2018) The Spatial Scale of a Species' Response to the Landscape Context Depends on which Biological Response You Measure. Current Landscape Ecology Reports 3:23-33.

Martin AE, Fahrig L (2012) Measuring and selecting scales of effect for landscape predictors in species–habitat models. Ecological Applications 22:2277–2292.

Matsuoka SM, Bayne EM, Sòlymos P, Fontaine PC, Cumming SG, Schmiegelow FKA, Song SJ (2012) Using binomial distance-sampling models to estimate the effective detection radius of point-count surveys across boreal Canada. The Auk 129:268–282.

Matsuoka SM, Mahon CL, Handel CM, Sólymos P (2014) Reviving common standards in point-count surveys for broad inference across studies.

Mayor SJ, Schneider DC, Schaefer JA, Mahoney SP (2009) Habitat Selection at Multiple Scales. Ecoscience 16:238–247.

Mays SE, Newberry GN, Riley L, Britten HB, Swanson DL (2019) Common Nighthawks (*Chordeiles Minor*) in Agricultural Landscapes: Genetic Structure of Populations Restricted to Urban Rooftop Nesting. The American Midland Naturalist 181:29–39.

McClintock BT, Bailey LL, Pollock KH, Simons TR (2010) Unmodeled observation error induces bias when inferring patterns and dynamics of species occurrence via aural detections. Ecology 91:2446–2454.

McFee B, Raffel C, Liang D, Ellis DP, McVicar M, Battenberg E, Nieto O (2017) librosa: Audio and music signal analysis in python. In Proceedings of the 14th python in science conference.

McGarigal K, Wan HY, Zeller KA, Timm BC, Cushman SA (2016) Multi-scale habitat selection modeling: a review and outlook. Landscape Ecology 31:1161–1175.

McGuire LP, Boyles JG, Brigham RM (2021) Lack of Foraging Site Fidelity Between Years by Common Nighthawks (*Chordeiles minor*). American Midland Naturalist 185:139-144.

McGuire LP, Brigham RM (2017) Common Nighthawks (*Chordeiles minor*) Can Take Off from Water. The Canadian Field-Naturalist 131:125–3.

McLauchlan KK, Higuera PE, Miesel J, Rogers BM, Schweitzer J, Shuman JK, Tepley AJ, Varner JM, Veblen TT, Adalsteinsson SA, Balch JK, Baker P, Batllori E, Bigio E, Brando P, Cattau M, Chipman ML, Coen J, Crandall R, Daniels L, Enright N, Gross WS, Harvey BJ, Hatten JA, Hermann S, Hewitt RE, Kobziar LN, Landesmann JB, Loranty MM, Maezumi SY, Mearns L, Moritz M, Myers JA, Pausas JG, Pellegrini AFA, Platt WJ, Roozeboom J, Safford H, Santos F, Scheller RM, Sherriff RL, Smith KG, Smith MD, Watts AC (2020) Fire as a fundamental ecological process: Research advances and frontier. Journal of Ecology 108:2047–2069.

McLean CM, Bradstock R, Price O, Kavanagh RP (2015) Tree hollows and forest stand structure in Australian warm temperate Eucalyptus forests are adversely affected by logging more than wildfire. Forest Ecology and Management 341:37–44.

McRae DJ, Duchesne LC, Freedman B, Lynham TJ, Woodley S (2001) Comparisons between wildfire and forest harvesting and their implications in forest management. Environmental Reviews 9:223–260.

Mesaros A, Heittola T, Virtanen T (2016) Metrics for Polyphonic Sound Event Detection. Applied Sciences 6:162–17.

Metcalf OC, Lees AC, Barlow J, Marsden SJ, Devenish C (2020) hardRain_ An R package for quick, automated rainfall detection in ecoacoustic datasets using a threshold-based approach. Ecological Indicators 109:105793.

Michel NL, Smith AC, Clark RG, Morrissey CA, Hobson KA (2015) Differences in spatial synchrony and interspecific concordance inform guild-level population trends for aerial insectivorous birds. Ecography 39:774–786.

Michelson CI, Clark RG, Morrissey CA (2018) Agricultural land cover does not affect the diet of Tree Swallows in wetland-dominated habitats. Condor 120:751–764.

Miguet P, Jackson HB, Jackson ND, Martin AE, Fahrig L (2016) What determines the spatial extent of landscape effects on species? Landscape Ecology 31:1177–1194

Miller AH (1925) The boom-flight of the Pacific Nighthawk. The Auk 27:141–143.

Miller DAW, Pacifici K, Sanderlin JS, Reich BJ (2018) The recent past and promising future for data integration methods to estimate species' distributions. Methods in Ecology and Evolution 10:22–37.

Moraga AD, Martin AE, Fahrig L (2019) The scale of effect of landscape context varies with the species' response variable measured. Landscape Ecology 34:703–715.

Morissette J, Bayne E, Kardynal K, Hobson K (2019) Regional variation in responses of wetland-associated bird communities to conversion of boreal forest to agriculture. Avian Conservation and Ecology 14:12.

Morissette JL, Cobb TP, Brigham RM, James PC (2002) The response of boreal forest songbird communities to fire and post-fire harvesting. Canadian Journal of Forest Research 32:2169–2183.

Nadeau CP, Conway CJ (2012) Field evaluation of distance-estimation error during wetland-dependent bird surveys. Wildlife Research 39:311–320.

Naimi B, Hamm NAS, Groen TA, Skidmore AK, Toxopeus AG (2013) Where is positional uncertainty a problem for species distribution modelling? Ecography. 37:191–203.

National Recovery Working Group (2005) Recovery Handbook (ROMAN). Recovery of Nationally Endangered Wildlife, Ottawa, ON, Canada.

Natural Regions Committee (2006) Natural Regions and Subregions of Alberta. Compiled by DJ Downing and WW Pettapiece. Government of Alberta. Pub. No. T/852.

Nebel S, Casey J, Cyr M-A, Kardynal KJ, Krebs EA, Purves EF, Bélisle M, Brigham RM, Knight EC, Morrissey C, Clark RG (2020) Falling through the policy cracks: implementing a roadmap to conserve aerial insectivores in North America. Avian Conservation and Ecology 15:23.

Nebel S, Mills A, McCracken J, Taylor P (2010) Declines of Aerial Insectivores in North America Follow a Geographic Gradient. Avian Conservation and Ecology 5:1.

Newberry GN, Swanson DL, 2018 (2018) Common Nighthawks (*Chordeiles minor*) in the Western Corn Belt: habitat associations and population effects of grassland and rooftop nesting habitat conversion. American Midland Naturalist 180:216-232.

Ng J (2009) Habitat use and home range characteristics of Common Nighthawks in mixed-grass prairie. M.Sc. thesis, University of Regina, Regina, SK, Canada.

Ng JW, Knight EC, Scarpignato AL, Harrison AL, Bayne EM, Marra PP (2018) First full annual cycle tracking of a declining aerial insectivorous bird, the Common Nighthawk (*Chordeiles* minor), identifies migration routes, nonbreeding habitat, and breeding site fidelity. Canadian Journal of Zoology 96:869–875.

Nice MM (1941) The role of territory in bird life. American Midland Naturalist 26:441–487.

Nicholson D (2016) Comparison of machine learning methods applied to birdsong element classification. In: Proceedings of the 15th Python in Science Conference.

North American Bird Conservation Initiative Canada. 2019. The State of Canada's Birds, 2019. Environment and Climate Change Canada, Ottawa, Canada. 12 pp.

NWT Centre for Geomatics (2020) Firehistory. http://www.geomatics.gov.nt.ca/Downloads/Vector/Biota/FireHistory.zip?h=1

Odum EP, Kuenzler EJ (1955) Measurement of Territory and Home Range Size in Birds. The Auk 72:128–137.

Oliveira AG de, Ventura TM, Ganchev TD, Figueiredo JM de, Jahn O, Marques MI, Schuchmann K-L (2015) Bird acoustic activity detection based on morphological filtering of the spectrogram. Applied Acoustics 98:34–42.

Orben RA, Fleishman AB, Borker AL, Bridgeland W, Gladics AJ, Porquez J, Sanzenbacher P, Stephensen SW, Swift R, McKown MW, Suryan RM (2019) Comparing imaging, acoustics, and radar to monitor Leach's storm-petrel colonies. PeerJ 7:e6721.

Osko T, MacFarlane A (2001) Natural reforestation on seismic lines and wellsites in comparison to natural burns or logged sites. Alberta-Pacific Forest Industries, Boyle, AB, Canada.

Palmer MW (1993) Putting things in even better order: the advantages of canonical correspondence analysis. Ecology 74:2215–2230.

Pannell JL, Dencer-Brown AM, Greening SS, Hume EA, Jarvis RM, Mathieu C, Mugford J, Runghen R (2019) An early career perspective on encouraging collaborative and interdisciplinary research in ecology. Ecosphere 10:e02899.

Partners in Flight (2020) Population Estimates Database, version 3.1. Available at http://pif.birdconservancy.org/PopEstimates.

Pasher J, Seed E, Duffe J (2014) Development of boreal ecosystem anthropogenic disturbance layers for Canada based on 2008 to 2010 Landsat imagery. Canadian Journal of Remote Sensing 39:42–58.

Peiman KS, Robinson BW (2010) Ecology and evolution of resource-related heterospecific aggression. The Quarterly Review of Biology 85:133–158.

Pérez-Granados C, Barrero A, Traba J, Rosa DB la, Reverter M, Gómez-Catasús J (2021) Assessment of cue counting for estimating bird density using passive acoustic monitoring: recommendations for estimating a reliable cue rate. Avian Conservation and Ecology 16:11.

Pérez-Granados C, Bota G, Giralt D, Barrero A, Gómez-Catasús J, Rosa DBL, Traba J (2019) Vocal activity rate index: a useful method to infer terrestrial bird abundance with acoustic monitoring. Ibis 161:901–907.

Pérez-Granados C, Traba J (2021) Estimating bird density using passive acoustic monitoring: a review of methods and suggestions for further research. Ibis 163:765–783.

Petrie M, Kempenaers B (1998) Extra-pair paternity in birds: explaining variation between species and populations. Trends in Ecology and Evolution 13:52–58.

Pickell PD, Andison DW, Coops NC (2013) Characterizations of anthropogenic disturbance patterns in the mixedwood boreal forest of Alberta, Canada. Forest Ecology and Management 304:243–253.

Pickell PD, Andison DW, Coops NC, Gergel SE, Marshall PL (2015) The spatial patterns of anthropogenic disturbance in the western Canadian boreal forest following oil and gas development. Canadian Journal of Forest Research 45:732–743.

Pickell PD, Hermosilla T, Coops NC, Masek JG, Franks S, Huang C (2014) Monitoring anthropogenic disturbance trends in an industrialized boreal forest with Landsat time series. Remote Sensing Letters 5:783–792.

Ploton P, Mortier F, Réjou-Méchain M, Barbier N, Picard N, Rossi V, Dormann C, Cornu G, Viennois G, Bayol N, Lyapustin A, Gourlet-Fleury S, Pélissier R (2020) Spatial validation reveals poor predictive performance of large-scale ecological mapping models. Nature Communications 11:4540.

Pollock KH, Nicholes JD, Brownie C, Hines JE (1990) Statistical Inference for Capture-Recapture Experiments. Wildlife Monographs 107:3–97.

Potamitis I, Ntalampiras S, Jahn O, Riede K (2014) Automatic bird sound detection in long real-field recordings: Applications and tools. Applied Acoustics 80:1–9.

Powers SM, Hampton SE (2019) Open science, reproducibility, and transparency in ecology. Ecological Applications 29:e01822.

Priyadarshani N, Marsland S, Castro I (2018) Automated birdsong recognition in complex acoustic environments: a review. Journal of Avian Biology 49:jav-01447.

Pröhl H, Koshy RA, Mueller U, Rand AS, Ryan MJ (2006) Geographic variation of genetic and behavioral traits in northern and southern túngara frogs. Evolution 60:1669-1679.

Provost FJ, Fawcett T, Kohavi R (1998) The case against accuracy estimation for comparing induction algorithms. In: Proceedings of the 15th International Conference on Machine Learning.

R Core Team (2016) R: a language and environment for statistical computing. Vienna, Austria. https://www.R-project.org/

R Core Team (2017) R: a language and environment for statistical computing. Vienna, Austria. https://www.R-project.org/

R Core Team (2020) R: a language and environment for statistical computing. Vienna, Austria. https://www.R-project.org/

Raffel C, McFee B, Humphrey EJ, Salamon J (2014) mir_eval: A transparent implementation of common MIR metrics. In Proceedings of the 15th International Society for Music Information Retrieval Conference.

Raghavan V, Bollmann P, Jung GS (1989) A critical investigation of recall and precision as measures of retrieval system performance. ACM Transactions on Information Systems 7:205–229.

Ralph CJ, Geupel GR, Pyle P, Martin TE, DeSante DF (1993) Handbook of field methods for monitoring landbirds. Albany, CA: U.S. Department of Agriculture, Forest Service, Pacific Southwest Research Station. Gen. Tech. Rep. PSW-GTR-144. 41 pp.

Reeves S, Xyrichis A, Zwarenstein M (2017) Teamwork, collaboration, coordination, and networking: Why we need to distinguish between different types of interprofessional practice. Journal of Interprofessional Care 32:1–3.

Reichman OJ, Jones MB, Schildhauer MP (2011) Challenges and Opportunities of Open Data in Ecology. Science 331:703–705.

Ricci B, Franck P, Valantin-Morison M, Bohan DA, Lavigne C (2013) Do species population parameters and landscape characteristics affect the relationship between local population abundance and surrounding habitat amount? Ecological Complexity 15:62–70.

Robertson BA, Hutto RL (2007) Is selectively harvested forest an ecological trap for Olive-sided Flycatchers? The Condor 109:109–121,].

Robinson OJ, Ruiz-Gutierrez V, Fink D (2017) Correcting for bias in distribution modelling for rare species using citizen science data. Diversity and Distributions 24:460–472.

Robinson, OJ, Ruiz-Gutierrez V, Reynolds MD, Golet GH, Strimas-Mackey M, Fink D (2019) Integrating citizen science data with expert surveys improves the accuracy and range of inference of species distribution models. Biodiversity Research 26:976-986.

Roever CL, Beyer HL, Chase MJ, Aarde RJ van (2013) The pitfalls of ignoring behaviour when quantifying habitat selection Roura-Pascual. Diversity and Distributions 20:322–333.

Rosenberg KV, Blancher PJ, Stanton JC, Panjabi AO (2017) Use of North American Breeding Bird Survey data in avian conservation assessments. Condor 119:594–606.

Rosenberg KV, Dokter AM, Blancher PJ, Sauer JR, Smith PA, Stanton JC, Panjabi AO, Helft L, Parr M, Marra PP (2019) Decline of the North American avifauna. Science 366:120–124.

Rosenfeld JS, Hatfield T (2006) Information needs for assessing critical habitat of freshwater fish. Canadian Journal of Fisheries and Aquatic Sciences 63:683–698.

Rota CT, Jr RJF, Dorazio RM, Betts MG (2009) Occupancy estimation and the closure assumption. Journal of Applied Ecology 46:1173–1181.

Roy C, Michel NL, Handel CM, Wilgenburg SLV, Burkhalter JC, Gurney KEB, Messmer DJ, Princé K, Rushing CS, Saracco JF, Schuster R, Smith AC, Smith PA, Sòlymos P, Venier LA, Zuckerberg B (2019) Monitoring boreal avian populations: how can we estimate trends and trajectories from noisy data? Avian Conservation and Ecology 14:8.

Ruegg KC, Anderson EC, Paxton KL, Apkenas V, Lao S, Siegel RB, DeSante DF, Moore F, Smith TB (2014) Mapping migration in a songbird using high-resolution genetic markers. Molecular Ecology 23:5726–5739.

Rushing CS (2019) Estimability of migration survival rates from integrated breeding and winter capture–recapture data. Ecology and Evolution 9:849–858.

Russo D, Voigt CC (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.

Rust HJ (1947) Migration and nesting of nighthawks in northern Idaho. Condor 49:177–188.

Rutz, C, Loretto M-C, Bates AE, Davidson SC, Duarte CM, Jetz W, Johnson M, Kato A, Kays R, Mueller T, Primack RB, Ropert-Coudert Y, Tucker MA, Wikelski M, Cagnacci F (2020). COVID-19 lockdown allows researchers to quantify the effects of human activity on wildlife. Nature Ecology & Ecolution 4:1156-1159.

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

Saab VA, Russell RE, Dudley JG (2007) Nest densities of cavity-nesting birds in relation to postfire salvage logging and time since wildfire. The Condor 109:97–108.

Salamon J, Bello JP (2017) Deep Convolutional Neural Networks and Data Augmentation for Environmental Sound Classification. IEEE Signal Processing Letters 24:279–283.

Salzberg SL (1997) On comparing classifiers: Pitfalls to avoid and a recommended approach. Data Mining and Knowledge Discovery 1:317–328.

Sauer JR, Peterjohn BG, Link WA (1994) Observer Differences in the North American Breeding Bird Survey. The Auk 111:50–62.

Schaub M, Abadi F (2011) Integrated population models: a novel analysis framework for deeper insights into population dynamics. Journal of Ornithology 152:227–237.

Schieck J, Song SJ (2006) Changes in bird communities throughout succession following fire and harvest in boreal forests of western North America: literature review and meta-analyses. Canadian Journal of Forest Research 36:1299–1318.

Schowalter TD, Webb JW, Crossley DA (1981) Communtiy Structure and Nutrient Content of Canopy Arthropods in Clearcut and Uncut Forest Ecosystems. Ecology 62:1010–1019.

Schuster R, Germain RR, Bennett JR, Reo NJ, Arcese P (2019) Vertebrate biodiversity on indigenous-managed lands in Australia, Brazil, and Canada equals that in protected areas. Environmental Science & Policy 101:1–6.

Seaman DE, Millspaugh JJ, Kernohan BJ, Brundige GC, J RK, A GR (1999) Effects of sample size on kernel home range estimates. The Auk 63:739–747.

Searcy WA, Beecher MD (2009) Song as an aggressive signal in songbirds. Animal Behaviour 78:1281–1292.

Sebastián-González E, Camp R, Tanimoto A, Oliveira P de, Lima B, Marques T, Hart P (2018) Density estimation of sound-producing terrestrial animals using single automatic acoustic recorders and distance sampling. Avian Conservation and Ecology 13:7.

Seedre M, Shrestha BM, Chen HYH, Colombo S, Jõgiste K (2011) Carbon dynamics of North American boreal forest after stand replacing wildfire and clearcut logging. Journal of Forest Research 16:168–183.

Seibold S, Gossner MM, Simons NK, Blüthgen N, Müller J, Ambarlı D, Ammer C, Bauhus J, Fischer M, Habel JC, Linsenmair KE, Nauss T, Penone C, Prati D, Schall P, Schulze E-D, Vogt J, Wöllauer S, Weisser WW (2019) Arthropod decline in grasslands and forests is associated with landscape-level drivers. Nature 574:671-674.

Service UF and W (2021) Birds of Conservation Concern 2021. United States Department of the Interior, US Fish and Wildlife Service, Migratory Birds, Falls Church, VA, USA

Shonfield J, Bayne EM (2017) Autonomous recording units in avian ecological research: current use and future applications. 12:14.

Shonfield J, Heemskerk S, Bayne EM (2018) Utility of Automated Species Recognition For Acoustic Monitoring of Owls. Journal of Raptor Research 52:42–55.

Sidie-Slettedahl AM, Jensen KC, Johnson RR, Arnold TW, Austin JE, Stafford JD (2015) Evaluation of autonomous recording units for detecting 3 species of secretive marsh birds. Wildlife Society Bulletin 39:626–634.

Sidler A (2017) Common Nighthawk (*Chordeiles minor*) habitat associations and activity patterns in the northern boreal forest. M.Sc. thesis, University of Regina, Regina, SK, Canada.

Simmonds EG, Jarvis SG, Henrys PA, Isaac NJB, O'Hara RB (2020) Is more data always better? A simulation study of benefits and limitations of integrated distribution models. Ecography 43:1-10.

Simpson BS (1985) Effects of location in territory and distance from neighbours on the use of song repertoires by Carolina wrens. Animal Behaviour 33:793–804.

Sing T., O. Sander, N. Beerenwinkel, and T. Lengauer (2005) ROCR: visualizing classifier performance in R. Bioinformatics 21:3940–3941.

Skowronski MD, Fenton MB (2009) Detecting Bat Calls: An Analysis of Automated Methods. Acta Chiropterologica 11:191–203.

Skowronski MD, Harris JG (2006) Acoustic detection and classification of microchiroptera using machine learning: Lessons learned from automatic speech recognition. The Journal of the Acoustical Society of America 119:1817–1833.

Slabbekoorn H, Smith TB (2002) Bird song, ecology and speciation. Philosophical Transactions of the Royal Society B: Biological Sciences 357:493–503.

Smith AC, Fahrig L, Francis CM (2011) Landscape size affects the relative importance of habitat amount, habitat fragmentation, and matrix quality on forest birds. Ecography 34:103–113.

Smith AC, Hudson M-AR, Downes CM, Francis CM (2015) Change Points in the Population Trends of Aerial-Insectivorous Birds in North America: Synchronized in Time across Species and Regions. PLoS ONE 10:e0130768.

Sokolova M, Japkowicz N, Szpakowicz and S (2006) Beyond Accuracy, F-score, and ROC: A Family of Discriminant Measures for Performance Evaluation. Lecture Notes in Computer Science, AI 2006: Advances in Artificial Intelligence. Pp 1015-1021.

Sokolova M, Lapalme G (2009) A systematic analysis of performance measures for classification tasks. Information Processing and Management 45:427–437.

Sòlymos P, Lele S, Bayne EM (2012) Conditional likelihood approach for analyzing single visit abundance survey data in the presence of zero inflation and detection error. Environmetrics 23:197–205.

Sòlymos P, Matsuoka SM, Bayne EM, Lele SR, Fontaine P, Cumming SG, Stralberg D, Schmiegelow FKA, Song SJ (2013) Calibrating indices of avian density from non-standardized survey data: making the most of a messy situation. Methods in Ecology and Evolution 4:1047–1058.

Sòlymos P, Matsuoka SM, Cumming SG, Stralberg D, Fontaine P, Schmiegelow FKA, Song SJ, Bayne EM (2018) Evaluating time-removal models for estimating availability of boreal birds during point count surveys: Sample size requirements and model complexity. Condor 120:765–786.

Sòlymos P, Toms JD, Matsuoka SM, Cumming SG, Barker NKS, Thogmartin WE, Stralberg D, Crosby AD, Dnes FV, Hach S, Mahon CL, Schmiegelow FKA, Bayne EM (2020) Lessons learned from comparing spatially explicit models and the Partners in Flight approach to estimate population sizes of boreal birds in Alberta, Canada. The Condor 122:1–22.

Southwood T (1977) Habitat, the Templet for Ecological Strategies? Journal of Animal Ecology 46:336-365.

Spiller KJ, Dettmers R (2019) Evidence for multiple drivers of aerial insectivore declines in North America. Condor 121:42–13.

Stagliano DM, Benke AC, Anderson DH (1998) Emergence of Aquatic Insects from 2 Habitats in a Small Wetland of the Southeastern USA: Temporal Patterns of Numbers and Biomass. Journal of the North American Benthological Society 17:37–53.

Stevenson Ben C, Borchers DL, Altwegg R, Swift RJ, Gillespie DM, Measey GJ (2015) A general framework for animal density estimation from acoustic detections across a fixed microphone array. Methods in Ecology and Evolution 6:38–48.

Stowell D, Wood M, Stylianou Y, Glotin H (2016) Bird detection in audio: A survey and a challenge. In 2016 IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP) pp. 1–6.

Stowell D, Wood MD, Pamuła H, Stylianou Y, Glotin H (2018) Automatic acoustic detection of birds through deep learning: The first Bird Audio Detection challenge. Methods in Ecology and Evolution 10:368-380.

Streby HM, Loegering JP, Andersen DE (2012) Spot-mapping underestimates song-territory size and use of mature forest by breeding golden-winged warblers in Minnesota, USA. Wildlife Society Bulletin 36:40–46

Subalusky AL, Post DM (2019) Context dependency of animal resource subsidies. Biological Reviews 94:517–538.

Sueur J, Pavoine S, Hamerlynck O, Duvail S (2008) Rapid acoustic survey for biodiversity appraisal. PLoS ONE. 3:e4065.

Sugai LSM, Silva TSF, Ribeiro JW, Llusia D (2018) Terrestrial Passive Acoustic Monitoring: Review and Perspectives. BioScience 69:15–25.

Sun K, Luo L, Kimball RT, Wei X, Jin L, Jiang T, Li G, Feng J (2013) Geographic Variation in the Acoustic Traits of Greater Horseshoe Bats: Testing the Importance of Drift and Ecological Selection in Evolutionary Processes. PLoS ONE 8:e70368-10.

Sutherland WJ, Adams WM, Aronson RB, Aveling R, Blackburn TM, Broad S, Ceballos G, Côté IM, Cowling RM, Da Fonseca GA, Dinerstein E, Ferraro PJ, Fleishman E, Gascon C, Hunter M Jr, Hutton J, Kareiva P, Kuria A, Macdonald DW, Mackinnon K, Madgwick FJ, Mascia MB, McNeely J, Milner-Gulland EJ, Moon S, Morley CG, Nelson S, Osborn D, Pai M, Parsons EC, Peck LS, Possingham H, Prior SV, Pullin AS, Rands MR, Ranganathan J, Redford KH, Rodriguez JP, Seymour F, Sobel J, Sodhi NS, Stott A, Vance-Borland K, Watkinson AR. (2009) One Hundred Questions of Importance to the Conservation of Global Biological Diversity. Conservation Biology 23:557–567.

Sutherland CA (1963) Notes on the behavior of Common Nighthawks in Florida. Living Bird 2:31-39.

Sutherland WJ, Dias MP, Dicks LV, Doran H, Entwistle AC, Fleishman E, Gibbons DW, Hails R, Hughes AC, Hughes J, Kelman R, Roux XL, LeAnstey B, Lickorish FA, Maggs L, Pearce-Higgins JW, Peck LS, Pettorelli N, Pretty J, Spalding MD, Tonneijck FH, Wentworth J, Thornton A (2019) A Horizon Scan of Emerging Global Biological Conservation Issues for 2020. Trends in Ecology & Evolution 35:81–90.

Swiston KA, Mennill DJ (2009) Comparison of Manual and Automated Methods for Identifying Target Sounds in Audio Recordings of Pileated, Pale-Billed, and Putative Ivory-Billed Woodpeckers. Journal of Field Ornithology 80:42–50.

Tachibana RO, Oosugi N, Okanoya K (2014) Semi-Automatic Classification of Birdsong Elements Using a Linear Support Vector Machine. PLoS ONE 9:e92584-8.

Tenopir C, Dalton ED, Allard S, Frame M, Pjesivac I, Birch B, Pollock D, Dorsett K (2015) Changes in Data Sharing and Data Reuse Practices and Perceptions among Scientists Worldwide. PLOS ONE 10:e0134826.

Thessen A (2016) Adoption of Machine Learning Techniques in Ecology and Earth Science. One Ecosystem 1:e8621-38.

Therneau TM, Grambsch PM (2000) Modelling survival data: extending the Cox model. Springer, New York, NY, USA.

Thogmartin WE (2010) Sensitivity analysis of North American bird population estimates. Ecological Modelling 221:173–177.

Thogmartin WE, Howe FP, James FC, Johnson DH, Reed ET, Sauer JR, Thompson FR (2006) A review of the population estimation approach of the North American landbird conservation plan. The Auk 123:892–904 .

Thornton DH, Jr RJF (2013) Body size and spatial scales in avian response to landscapes: a meta-analysis. Ecography 37:454-463.

Timoney K (2003) The changing disturbance regime of the boreal forest of the Canadian Prairie Provinces. The Forestry Chronicle 79:502–516.

Tingley MW, Ruiz-Gutierrez V, Wilkerson RL, Howell CA, Siegel RB (2016) Pyrodiversity promotes avian diversity over the decade following forest fire. Proceedings of the Royal Society B: Biological Sciences 283:20161703–9.

Tingley MW, Stillman AN, Wilkerson RL, Howell CA, Sawyer SC, Siegel RB (2018) Cross-scale occupancy dynamics of a postfire specialist in response to variation across a fire regime. Journal of Animal Ecology 87:1484–1496.

Tingley MW, Stillman AN, Wilkerson RL, Sawyer SC, Siegel RB (2020) Black-backed woodpecker occupancy in burned and beetle-killed forests: Disturbance agent matters. Forest Ecology and Management 455:117694.

Tittler R (2008) Source-sink dynamics, dispersal, and landscape effects on North American songbirds. Ph.D. thesis, Carleton University, Ottawa, ON, Canada.

Turner MG (2010) Disturbance and landscape dynamics in a changing world. Journal of Applied Ecology 91:2833–2849.

Turner MG, Romme WH, Gardner RH, Hargrove WW (1997) Effects of fire size and pattern on early succession in Yellowstone National Park. Ecological Monographs 67:411–433.

Twining CW, Shipley JR, Winkler DW (2018) Aquatic insects rich in omega-3 fatty acids drive breeding success in a widespread bird. Ecology Letters 21:1812–1820.

Ulloa JS, Gasc A, Gaucher P, Aubin T, Réjou-Méchain M, Sueur J (2016) Screening large audio datasets to determine the time and space distribution of Screaming Piha birds in a tropical forest. Ecological Informatics 31:91–99.

Upham-Mills EJ, Reimer JR, Haché S, Lele SR, Bayne EM (2020) Can singing rate be used to predict male breeding status of forest songbirds? A comparison of three calibration models. Ecosphere 11:e03005.

Van Wilgenburg SL, Mahon CL, Campbell G, McLeod L, Campbell M, Evans D, Easton W, Francis CM, Haché S, Machtans CS, Mader C, Pankratz RF, Russell R, Smith AC, Thomas P, Toms JD, Tremblay JA (2020) A cost efficient spatially balanced hierarchical sampling design for monitoring boreal birds incorporating access costs and habitat stratification. PLOS ONE 15:e0234494.

Van Wilgenburg SL, Sòlymos P, Kardynal KJ, Frey MD (2017) Paired sampling standardizes point count data from humans and acoustic recorders. Avian Conservation and Ecology 12:13.

Venier LA, Mazerolle MJ, Rodgers A, McIlwrick KA, Holmes S, Thompson D (2017) Comparison of semiautomated bird song recognition with manual detection of recorded bird song samples. 12:2.

Viel JG, Donnelly A, Fredlund G, Mueller W (2020) Common Nighthawk (*Chordeiles minor*) Occurrence in Urban Areas Southeastern Wisconsin, USA. American Midland Naturalist 183:233-245.

Waddle JH, Thigpen TF, Glorioso BM (2009) Efficacy of automatic vocalization recognition software for anuran monitoring. 4:384–388.

Wagner DL (2020) Insect Declines in the Anthropocene. Annual Review of Entomology 65:457–480.

Wedgwood J (1973) Nighthawks in the city. Blue Jay 31:65-238.

Westneat DF, Sherman PW (1997) Density and extra-pair fertilizations in birds: a comparative analysis. Behavioural Ecology and Sociobiology 41:205–215.

Whitaker DM, Warkentin IG (2010) Spatial Ecology of Migratory Passerines on Temperate and Boreal Forest Breeding Grounds. The Auk 127:471–484.

Wiens JA (1989) Spatial Scaling in Ecology. Functional Ecology 3:385.

Wildlife Acoustics (2011) Song Scope Bioacoustics Software Version 4.0 Documentation. Wildlife Acoustics, Maynard, MA, USA.

Wildlife Acoustics (2016) Converting Song Scope recognizer to Kaleidoscope Cluster-based Classifiers.

Will T, Stanton JC, Rosenberg KV, Panjabi AO, Camfield AF, Shaw AE, Thogmartin WE, Blancher PJ (2020) Handbook to the Partners in Flight Population Estimates Database, Version 3.1. PIF Technical Series No. 7.1.

Wood SN (2011) Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. Journal of the Royal Statistical Society (B) 73:3-36.

Xie J, Towsey M, Zhang J, Roe P (2016) Frog call classification: a survey. Artificial Intelligence Review 49:375-391.

Yip DA, Bayne EM, Sòlymos P, Campbell J, Proppe D (2017a) Sound attenuation in forest and roadside environments: Implications for avian point-count surveys. Condor 119:73–84.

Yip DA, Leston L, Bayne EM, Sòlymos P, Grover A (2017b) Experimentally derived detection distances from audio recordings and human observers enable integrated analysis of point count data. Avian Conservation and Ecology 12:11.

Yip DA (2020) Using autonomous recording units for estimating detectability in animal populations. PhD thesis, University of Alberta, Edmonton, AB, Canada.

Yip DA, Knight EC, Audet EH, Wilson SJ, Charchuk C, Scott CD, Sòlymos P, Bayne EM (2020) Sound level measurements from audio recordings provide objective distance estimates for distance sampling wildlife populations. Remote Sensing in Ecology and Conservation 71:2759–15.

York JE, Young AJ, Radford AN (2014) Singing in the moonlight: dawn song performance of a diurnal bird varies with lunar phase. Biology Letters 10:20130970.

Youden WJ (1950) Index for rating diagnostic tests. Cancer 3:32–35.

Zipkin EF, DeWan A, Royle JA (2009) Impacts of forest fragmentation on species richness: a hierarchical approach to community modelling. Journal of Applied Ecology 46:815–822.