INVESTIGATIONS OF ECOACOUSTIC BIODIVERSITY CHARACTERIZED WITH REMOTE SENSING DATA USING MACHINE AND DEEP LEARNING APPROACHES

By Colin Aidan Quinn

A Dissertation

Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

in Informatics & Computing

Northern Arizona University

May 2023

Approved:

Scott J. Goetz, Ph.D., Chair

Matthew L. Clark, Ph.D.

Bret Pasch, Ph.D.

Patrick Jantz, Ph.D.

ABSTRACT

INVESTIGATIONS OF ECOACOUSTIC BIODIVERSITY CHARACTERIZED WITH REMOTE SENSING DATA USING MACHINE AND DEEP LEARNING APPROACHES

COLIN AIDAN QUINN

Animal biodiversity assessment provides key measurements on the health and dynamics of animals in relation to habitat setting, human presence, and disturbances. As these factors continue to change due to climate change, introducing more intense and frequent wildfires, for example, assessing animal dynamics allows for proper protection and scientific understanding of the indicators and influences leading to animal community change. Ecoacoustic recording of the landscape is a rapidly evolving new approach in capturing the dynamics of animals and human impact. Recording passive soundscapes, the collection of noises emanating from the landscape, amounts to large quantities of acoustic data that require informed methods for summarizing underlying patterns so methods are generalizable and reliable. Many of these efforts over the previous decade and a half focused on a suite of metrics called acoustic indices that summarize dimensions of the acoustic signals such as frequency and amplitude. However, acoustic indices lack generalizability when compared among study regions particularly in relation to biotic noise sources (Biophony) and anthropogenic noise sources (Anthropophony). These gaps in understanding and lack of confidence in applying acoustic indices to our dataset of over 700,000 minutes of recordings from Sonoma County, California and the Soundscapes to Landscapes project led to the creation of the first research project in this dissertation where I classified soundscapes into informative types of sound called soundscape components using convolutional

neural networks (Chapter 2). The second project then analyzed the patterns in classified soundscapes with a set of fifteen acoustic indices to investigate in what sonic conditions indices may be more reliable and how Biophony along with indices can model bird species richness (Chapter 3). From these two chapters I provided a classification approach to distill sounds types in 700,000 minutes of acoustic data and related these sounds to acoustic indices to assess their ability to reflect signals of biodiversity and robustness in the presence of confounding sound sources. The final project combined the above work to analyze spatio-temporal patterns of acoustic indices and soundscape components across Sonoma County (Chapter 4). This included modeling the ecoacoustic metrics with a suite of remote sensing variables in a species distribution modeling framework. Products from these models were used to understand the most important metrics, structural and climate metrics, in predicting ecoacoustic metrics. Maps were used to investigate increased Biophony levels following wildlife activity across increasing burn severities and patterns related to human impact. Combined, this work demonstrates the ability to leverage machine learning approaches to understand the complex nature of soundscape dynamics using large datasets in a regional-scale analysis.

ACKNOWLEDGEMENTS

I am deeply appreciative of the support from my committee members, Drs Scott J. Goetz, Patrick Jantz, Bret Pasch, and Matthew L. Clark, and committee-adjacent mentors who provided perspective on the work from day 1. Dr. Scott Goetz, my advisor and committee chair, supported exploring my interests in my first year of the Ph.D., which contributed to laying a solid foundation that evolved into this dissertation. He provided a supportive environment with invaluable input on research that enabled a space for work I could not have imagined and the ultimate success of this research. Even though the Ph.D. committee is inherently limited in size, I feel folks included as co-authors on these manuscripts provided significant input on my research starting in earnest the first summer of COVID. This includes Patrick Burns, Dr. Gurman Gill, Shrishail Baligar, Rose L. Snyder, Dr. Leonardo Salas, Dr. Christopher R. Hakkenberg, and the academic community I found myself in was critical in supporting this work. Specifically, Dr. Matthew Clark and Pat Burns were critical to the development of the work in this dissertation, frequently meeting to discuss progress and provide input on analyses over the past five years. My academic career and drive to continue research was sparked by Dr. Deborah Nichols, who was endlessly supportive and patient when I was trying to navigate being a reliable undergraduate researcher.

All of my research rests upon the mountains of work that field technicians and Soundscapes to Landscape volunteers completed from 2017-2022. Without the extensive project management, field seasons, and volunteer-oriented soundscape identification sessions, none of the data included in this dissertation would exist for these analyses. I don’t think any of this work would have been physically possible in my timeline without the support from the folks behind NAU’s high-performance computing cluster, Monsoon. That sentiment is quite literal, as the

total computing time used for this project and associated analyses is a mind-boggling 50,000,000 hours roughly. As with any published work, I am grateful to the anonymous reviewers for my first two manuscripts, who helped highlight the robustness of the peer-review process, ensuring that science is as sound as possible.

Throughout my work days, the GEODE group helped maintain a work-life balance, whether chatting about research, a recent adventure, or just sharing a coffee over our workhorse of an espresso machine. It isn't possible to name all of my friends and family who supported and encouraged me as I continued this journey. Whether unwinding with a competitive round of trivia, standing around a fire trying to "social distance" in frigid temperatures, or getting out on a hike or bike around Flagstaff, it is all appreciated and helped with the excellent work-life balance. Finally, I would like to thank my best friend and partner, Singer, whose patience and support persisted through too many ups and downs associated with my graduate work as well as an acknowledgement of Smoulder. I'm excited to see what new adventures we have around the corner.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

Chapters 2, 3, and 4 of this dissertation were prepared as academic manuscripts that used various machine-learning techniques to develop, test, and apply novel soundscape analyses in Sonoma County, CA. These chapters have or will be submitted to peer-reviewed academic journals for publication and have been reviewed by co-authors. Some redundancy will occur due to formatting these manuscripts to adhere to university formatting requirements. Each chapter has a reference section to allow for clarity. Chapter 1 consists of a review of literature and contextualization related to all manuscripts while chapters 2, 3, and 4 contain more targeted introductions. Chapter 2 focuses on developing soundscape components to assess biodiversity in the Soundscapes to Landscape acoustic dataset and has been accepted in the journal *Ecological Indicators* (doi.org/10.1016/j.ecolind.2022.108831). Chapter 3 contextualizes the novel soundscape components with a suite of traditional ecoacoustic indices in relation to bird species richness and is accepted in the journal of *Frontiers in Remote Sensing Acoustic Remote Sensing* and awaiting publication at the time of this dissertation. Chapter 4 applies the soundscape metrics in a species distribution modeling framework to understand their relation to remote sensing-derived ecosystem characteristics, their spatiotemporal patterns, and the potential for Biophony to capture acoustic community response following wildfires and is being prepared for submission. Chapter 5 summarizes the findings in the dissertation and provides the implications and anticipated future directions for related ecoacoustic studies.

CHAPTER 1: OVERVIEW AND INTRODUCTION

1.1. BIOMONITORING

Biological diversity, or biodiversity, is recognized as an essential component of ecosystem health and highly valued by society (Martens et al., 2003). In addition, greater species-level biodiversity has been linked with increases in ecosystem productivity and stability (Burger & Gochfeld, 2004; Martens et al., 2003). Biodiversity has been described as "the sum total of all biotic variation from the level of genes to ecosystems" (Purvis & Hector, 2000) and here I use it specifically to reflect the number of avian species and community diversity in an ecosystem.

Research continues to demonstrate the sensitivity of biodiversity to human presence and climate change (Barnosky et al., 2011; Johnson et al., 2017; Pimm et al., 2014). Species extinction has occurred at rates higher than the expected background rates since humans began altering their environments over 2 million years ago and has accelerated over the past 500 years (Barnosky et al., 2011; Dirzo et al., 2014; Johnson et al., 2017; Joppa et al., 2016; Pimm et al., 2014). Not only is the loss of biodiversity an indicator of human impact, but environmental change is easily observed and intricately tied with loss (Dirzo et al., 2014). Certain species, for example birds, have been used as indicators of environmental degradation by linking their change in abundance with informative environmental measures (Burger & Gochfeld, 2004; Morrison, 1986; Radford & Bennett, 2007). Changes in landscapes (e.g. habitat loss) can now be rapidly assessed using remote sensing imagery, but direct assessment of fauna in these environments remains difficult (Dirzo et al., 2014).

Biodiversity assessment leads to the ability to monitor or assess species communities over space and time (Debinski and Humphry 1997). Relevant and meaningful indicators are required

to measure biodiversity. Noss (1990) summarizes seven key factors for what an indicator should be: (1) sensitive to change and provides early warning; (2) broadly geographically applicable; (3) can provide continuous assessment across a gradient of stress; (4) independent of sample size; (5) easy and cost effective for data collection; (6) sensitive to natural cycles versus those induced by anthropogenic impact; and (7) relevant to ecologically meaningful events (Noss, 1990). Once an indicator or ideally a suite of indicators are identified for biodiversity assessment they can be implemented for long-term monitoring (Gómez et al., 2018).

Measurements of diversity are particularly useful when comparing across multiple places or time and the choice of indicators should be based on the priority of research objectives (e.g. changes in bird species across habitat types or change in a specific species at one location; Purvis & Hector, 2000). At the species and community levels, this is considered compositional biodiversity or the frequency/relative abundance of species that are measurable in a community, both through observation and genetic diversity (e.g. not considered structural or functional biodiversity; Duelli & Obrist, 2003; Franklin, 1988; Noss, 1990). Different measurements target unique levels of compositional biodiversity (e.g. species, genus, or family) to account for differences among assemblages both locally and ultimately across space and time (Gotelli & Colwell, 2011; Magurran, 2004). For this research I focus on species diversity and 'community' diversity summarized using acoustic recordings, where the community is composed of class taxonomic levels (e.g. amphibians, birds, insects, and mammals).

There are three commonly recognized measures (indicators) to assess biodiversity: (1) species richness – the number of unique species in a community; (2) species evenness – the measure of how similar observed abundances are among species; and (3) species dominance –

the extent to which a single species is notably more prevalent in a community (Magurran, 2004). From these measures, levels of biodiversity are usually equated with higher values of evenness or diversity and low values of dominance, but this is dependent on the application and composite indices combine these metrics for more robust assessments (Magurran, 2004; Morris et al., 2014). In the scope of this research, continued biodiversity monitoring involves empirically recording and modeling acoustically derived diversity metrics related to environmental factors considered as proxies that influence aspects of the target species' ecology (Morrison, 1986). However, the complexity of measuring biodiversity means that the use of a single indicator value is frequently insufficient to represent the true magnitude of diversity and instead the use of multiple indicators tends to be more appropriate (Gómez et al., 2018; Purvis & Hector, 2000; Scarpelli et al., 2021). To apply new indices and achieve proper biodiversity monitoring extensive empirical fieldwork and observation is required, which is where the utility of low-cost and low-effort acoustic recording units (ARUs) provides value (Dumyahn & Pijanowski, 2011).

1.2. ECOACOUSTICS

To properly monitor bird diversity, robust empirical observations of species are required (Bush et al., 2018). This is where the recording of sound produced by vocalizing species becomes an invaluable method (Haselmayer & Quinn, 2000; Sueur et al., 2008). Sound as an indicator of animal presence has been beneficial for 'humans' for over 2 million years (Ploog, 2002). In the modern era, these sounds or vocalizations have been used to measure the behavioral patterns in animals (bioacoustics) and more recently, to estimate biodiversity of vocalizers in an ecosystem (ecoacoustics). Here, I focus on the latter, specifically the

foundational knowledge that allows us to use digitally recorded animal vocalizations to estimate measures of biodiversity.

Sound provides a valuable source of information on animal behavior and more generally, animal presence. For birds, sound may have developed along with vision as two essential means of communicating over other senses such as scent which is limited in its spatial effect and slow propagation speed (Catchpole & Slater, 1995). However, it seems that sound may have many behavioral benefits over even vision. For example, dependence on vision for communicating is difficult in densely vegetated or complex ecosystems leading to a prominence in non-visual signaling preferences (Boncoraglio & Saino, 2007; Riede, 1993; Stoddard & Prum, 2011). Additionally, sound has the benefit of travelling longer distances than vision in some cases (e.g. dense forest, around obstacles) allowing for cryptic or difficult to visually identify signaling species to still be observed (Boncoraglio & Saino, 2007).

Recorded sounds from the environment have been shown to yield higher rates of bird detection probability than visual point counts, especially for species in forested or structurally complex habitats where visual identification is limited (point counts performed better in more open habitat types) or when richness is high during the dawn chorus (Celis-Murillo et al., 2012; Haselmayer & Quinn, 2000). However, implementation of sound recordings to measure the bird species presence in an effort to assess and monitor biodiversity across habitat types remains a non-trivial task (Clark et al., 2023; Snyder et al., 2022; Sueur et al., 2014; Wiley & Richards, 1978). Fortunately, sound recording for the aforementioned purposes has the following benefits over traditional sampling methodologies: (1) more cost efficient (with recent technology); (2) provides a permanent record to detect *all* species; (3) reduces repeat field visits and impact from

researchers on field sites; (4) useful in remote locations; (5) can target rare or cryptic species; in addition to many others (Depraetere et al., 2012; Shonfield & Bayne, 2017). There are drawback though, such as reduced distance of detection compared to human observation in some cases, a need for large storage space, sufficient computing hardware for processing and interpreting data, and non-trivial transferability among study domains (Gibb et al., 2018; Shonfield & Bayne, 2017).

1.3. ECOACOUSTICS FOR BIOMONITORING

Ecoacoustics focuses on treating sound as an indicator of ecosystem status especially at organizational levels coarser than the individual species (i.e. community or taxa levels) (Eldridge et al., 2018; Pijanowski et al., 2011a). The aforementioned classical measures of biodiversity represent an aspect of the within or between community assemblages of diversity estimated using methods such as visual or auditory observations recorded by a researcher. In ecoacoustics, biodiversity measurements are approximated using patterns in acoustic energy (i.e. variation in frequency, amplitude, and time) (Eldridge et al., 2018). Since 2007, over 69 unique measurements of sound diversity (i.e. acoustic indices) have been described to measure biological and anthropogenic components of soundscapes (Buxton et al., 2018a). These acoustic indices quantify unique properties of the soundscape, can be used as distilled representations of the soundscape, and approximate biodiversity or animal acoustic activity (Sueur et al., 2008). They also allow for the characterization of acoustic communities without the need for species identification by experts or the application of computationally intensive automated recognition (Gasc et al., 2013).

The term soundscape draws on a half-century old concept from urban planning described as "the collection of sounds that emanate from landscapes" (Southworth, 1969). Southworth (1969) emphasized the need to increase the informative-ness in the urban soundscape to improve sensory awareness and quality of life for residents. It took until 1977 for these concepts to spread to ecology, not yet for biodiversity research, with the goal of using naturally occurring sounds to compose music (Schafer, 1993). By the 1980s, components of the soundscape were more formally defined to include biophony, the assemblage of animal created sounds, and geophony, or those sounds created by wind, rain, and other abiotic or weather production (Krause, 1987). Pijanowski et al., (2011b) further defined another component, anthropophony as the sounds created by humans. Biophony is an important component of the soundscape as it is closely related to the acoustic animal community and is the "collection of sounds produced by all living organisms in a given habitat" (Gasc et al., 2013; Pijanowski et al., 2011b). Anthropophony has been used to assess the impact humans have on an environment and the natural soundscape (Mckenna, 2020; Shannon et al., 2016). However, these soundscape components (Anthropophony, Biophony, and Geophony) are typically abstractly summarized of acoustic activity using acoustic indices or frequency bands, often applied as proxies for biodiversity.

The first acoustic index created in 2007, the bioacoustic index (BI), converted digital sound recordings to interpretable numerical summaries of bird-specific activity of sound level and frequency space (Boelman et al., 2007). BI was directly related to gradients in avian abundance across multiple ecosystems and compared to avifauna point counts as a viable indicator of biodiversity (Boelman et al., 2007). In 2008, a suite of acoustic indices (acoustic entropy) were created using concepts from Shannon's index (Chao & Shen, 2003; Sueur et al., 2008). For

acoustic entropy (H), the Shannon's index category is the units of time and probability is a function of the amplitude across the frequency range which is insensitive to the number of time categories (i.e. the length of recording) (Sueur et al., 2008). H measures evenness across the amplitude envelope which can be thought of as the change in amplitude observed in an amplitude time. Further, H describes community level diversity and was created because of the high sampling effort required to derive traditional biodiversity metrics (Sueur et al., 2008). BI and the acoustic entropy measures were the first indices to be described and have been followed by more than 60 other unique indices.

The vast number of acoustic indices has led to (1) variation in methods to select indices for analysis and (2) variation in results and informative indices found across studies, possibly as an unintended byproduct of the unique assemblages used in different studies. For example, Buxton et al. (2018b), applied 36 indices to sites in the Western Ghats, India and found two indices, the acoustic complexity index (ACI) and roughness to be informative for predicting bird species vocalization (Buxton et al., 2018b). Comparatively, Machado et al. (2017) applied two indices to sites in the Brazilian Cerrado and found the acoustic diversity index (ADI) was associated with bird species richness and normalized difference soundscape index (NDSI) was lower for areas close to human impact (Machado et al., 2017). Buxton et al., (2018b) implemented ADI and NDSI, but these indices were not found to be important for predicting avian vocal diversity (Buxton et al., 2018b). These differences in index findings may be due to unique assemblages of biodiversity or they may be a byproduct of confounding methodological designs, differences between studies in the surrounding ecosystem structure, or the presence of natural and anthropogenic background sounds.

Sueur et al., (2014) state “To find a single index that summarizes all biodiversity facets is undoubtedly utopian,” which echoes prior biodiversity sentiments (Purvis & Hector, 2000). They go further though in demonstrating the success of α indices – indices that estimate within group diversity (e.g. BI, H, NDSI, ACI, ADI, etc.) – because they yield a single value intended to represent a soundscape property similar to richness or abundance (Sueur et al., 2014). Comparatively, other indices that instead focus on dissimilarity (between-group diversity; β-indices) have not been linked directly with bird species community composition and tend to require pair-wise comparisons, but are still used (Lellouch et al., 2014; Sueur et al., 2014). Combining acoustic indices to take advantage of individual weaknesses and strengths is a key methodological consideration in properly applying ecacoustic methods (Bradfer-Lawrence et al., 2019; Gómez et al., 2018; Scarpelli et al., 2021). Differences in acoustic index application motivated the first and second manuscripts in this dissertation to investigate the effects of soundscape components on commonly used acoustic indices.

1.4. ENVIRONMENTAL CONTEXT OF ECOACOUSTICS

The effect of environmental surroundings directly on acoustic indices, not necessarily on sound propagation, has been discussed, but is infrequently measured. Here is a non-exhaustive summary of studies linking environmental characteristics with acoustic dynamics:

- Depraetere et al., (2012) demonstrate the acoustic richness (AR) index and the dissimilarity index (D) vary with increased microhabitat space (i.e. higher in structurally complex young forest or forest edge) and with distinct habitat structure, respectively. Additionally, these indices are suggested to provide robust estimation of animal diversity

assessment, even when considerable background noise is present, across multiple seasons or years and landscape scales (Depraetere et al., 2012).

- Buxton et al., (2018b) provide another line of evidence that sound-informed avian species richness is directly related to more complex habitat and land-use types such as the shade-grown crops – coffee and cardamom – that grow beneath the shade of rainforest canopies compared to lower avian richness in intensively managed tea plantations (Buxton et al., 2018b).
- In another study, ACI and BI were found to be related to avian song intensity, but not to landscape or biodiversity characteristics, however three indices (NDSI, H, and AEI – acoustic evenness index) were related to landscape configuration, ecological condition, and bird species richness (Fuller et al., 2015).
- Bradfer-Lawrence et al. (2019) suggest the use of seven indices in combination permits discrimination among habitat-specific soundscapes (Bradfer-Lawrence et al., 2019).
- Gómez et al., (2018) employed a suite of acoustic indices to measure the quality of habitats using acoustic activity. They provide a method that integrates several α-indices sensitive to differences in habitat using a mulitlayer neural network and support vector machine to classify three habitat types with different degrees of disturbance. The predictive modeling identified 14 indices as good discriminators between different habitats: AEI, ADI, ACI, BI, NDSI, P – ratio of biophony to anthrophony, M, NP – number of peaks, MID – mid-band activity, ESM – entropy of spectral maxima, ESV – entropy of spectral variance, H_t, H, and AR (Gómez et al., 2018).

- Other indices have been linked to differences in bird assemblage composition and habitat structural differences (Boelman et al., 2007; Pekin et al., 2012; Rodriguez et al., 2014). Machado et al., (2017) compared environmental measures (e.g. normalized difference vegetation index - NDVI) between two ecosystems (cerrado/closed forest and gallery forest) with ADI and NDSI and found both acoustic indices were higher in the more structurally complex gallery forest associated with higher bird richness (Machado et al., 2017). Boelman et al., (2007) found that avian abundance indicated using BI was highest in habitats with higher canopy cover and height, measured using LiDAR and AVIRIS hyperspectral data (Boelman et al., 2007).

Many of these studies linked a small set of environmental characteristics, usually restricted to vegetation metrics or land use/land cover (LULC) metrics but did not incorporate an exhaustive set of environmental characteristics (e.g., climate and geomorphology with vegetation metrics). Furthermore, fewer studies attempt to interpolate or spatially understand the patterns in acoustic activity (e.g., Hao et al., 2021; Matsinos and Tsaligopoulos, 2018; Mennitt et al., 2014; Turner et al., 2018), though continuous mapping of acoustic activity can facilitate a better understanding of soundscape dynamics across heterogeneous landscapes.

1.5. MOTIVATIONS AND OBJECTIVES

The initial goal of this dissertation was to use a large acoustic recording archive to understand the spatial dynamics of animal and human communities in the context of disturbance. However, I quickly realized that in order to address these questions in a rigorous manner required further investigating the underlying assumptions and patterns of the ecoacoustic methods I planned to use. At the time, this was primarily centered around investigating acoustic

indices, the numerical summaries of acoustic energy in a recording. Their non-trivial interpretation in novel ecosystems and datasets meant a more thorough understanding of their patterns was needed, beyond an aural validation. Initially, this was going to be related to acoustically derived bird species richness, however, the models used to generate these summaries were delayed which led to my investigation of soundscape dynamics and Chapter 2.

In Chapter 2 of this dissertation, I trained a convolutional neural network (CNN) with the goal of classifying broad components of soundscapes, namely Anthropophony, Biophony, Geophony, Quiet, and ARU Interference (ABGQI). This investigation was undertaken because of the unknown patterns of biodiversity and human impact that acoustic indices would reflect in the S2L dataset. The chapter required extensive model testing and validation as the model was trained on a much smaller set of data than the S2L dataset, therefore, choices during training had to ensure that overfitting was minimal. A high-accuracy CNN was developed that achieved average classification accuracy of 0.88, 0.94 precision, and 0.80 recall across the five soundscape components while assessing error. I was then able to investigate if we observed anticipated patterns in relation to hour of day and LULC.

Chapter 3 combined the work of classifying the S2L dataset into interpretable sound sources and related these to a suite of fifteen acoustic indices and acoustically derived bird species richness. The dataset also increased from approximately 487,000 minutes of data to 725,000 minutes of data. These analyses involved modeling each acoustic index with the five soundscape components to investigate the effects of soundscape components on acoustic index values. I was able to model the fifteen acoustic indices with a high degree of performance (avg. adj-R^2 = 0.61 ± 0.16) while maintaining the interpretability of the covariate effects by using generalized

additive models. I first demonstrated that Biophony alone was a robust indicator of biotic activity when compared to acoustic indices and then provided recommendations as to which indices perform well for representing Biophony and which indices should be used with caution when Anthropophony and Geophony are present. These relationships were then extended to model bird species richness finding that Biophony-alone is able to reflect bird species richness with the same performance as fifteen acoustic indices when targeting morning periods. Combining Biophony with acoustic indices led to the best ability to predict bird species richness and the morning-only targeted performance was comparable in this case with the 24-hr average value model. These findings provide suggestions (1) when ecoacoustic metrics more reliably indicate bird species richness and (2) that quantify sonic conditions that adversely affect acoustic index interpretations.

Chapter 4 used the classified soundscape components, three unique acoustic indices, and bird species richness to understand the spatiotemporal patterns of soundscapes. I developed a set of predictive random forest models that linked remote sensing and climate environmental characteristics with soundscapes. From these models, I was able to quantify the changes in soundscapes across four period of the day and a range of habitat types to understand effects of human impact and wildfire disturbance and identify the importance of forest structure derived from the spaceborne Lidar sensor GEDI (Global Ecosystem Dynamics Investigation) and climate in modeling soundscapes. These analyses support the quantification of spatial and temporal ecoacoustics monitoring techniques that can facilitate the understanding of vocalizing species dynamics by highlighting core and edge communities and reorganization of species following wildfire.

The following is a summary of the primary objectives related to this research from the three analyses:

1) Develop a flexible deep learning solution to quantify soundscape dynamics to inform human activity and biodiversity and identify confounding noise sources in recordings.
2) Compare the reliability of the newly developed ecoacoustic metrics (soundscape components) to traditional ecoacoustic indices and assess both sets of indices to reflect bird species richness, a measure of biodiversity.
3) Use soundscape components to understand spatiotemporal patterns of sound across Sonoma County and their relationships with ecosystem characteristics, including the effects of wildfire burn severity, described using remote sensing and climate data.

1.6. REFERENCES

Barnosky, A. D., Matzke, N., Tomiya, S., Wogan, G. O. U., Swartz, B., Quental, T. B., … Ferrer, E. A. (2011). Has the Earth's sixth mass extinction already arrived? *Nature*, *471*(7336), 51–57. doi.org/10.1038/nature09678

Boelman, N. T., Asner, G. P., Hart, P. J., & Martin, R. E. (2007). Multi-trophic invasion resistance in Hawaii: Bioacoustics, field surveys, and airborne remote sensing. *Ecological Applications*, *17*(8), 2137–2144. doi.org/10.1890/07-0004.1

Boncoraglio, G., & Saino, N. (2007). Habitat structure and the evolution of bird song: A meta-analysis of the evidence for the acoustic adaptation hypothesis. *Functional Ecology*, *21*(1), 134–142. doi.org/10.1111/j.1365-2435.2006.01207.x

Bradfer-Lawrence, T., Gardner, N., Bunnefeld, L., Bunnefeld, N., Willis, S. G., & Dent, D. H. (2019). Guidelines for the use of acoustic indices in environmental research. *Methods in Ecology and Evolution*, *00*, 1–12. doi.org/10.1111/2041-210X.13254

Burger, J., & Gochfeld, M. (2004). Marine Birds as Sentinels of Environmental Pollution. *EcoHealth*, *1*(3), 263–274. doi.org/10.1007/s10393-004-0096-4

Bush, A., Sollmann, R., Wilting, A., Bohmann, K., Cole, B., Balzter, H., … Yu, D. W. (2018). The Promise and Practice of Connecting Earth Observation to Biodiversity and Ecosystem Services. *Nature Ecology and Evolution*. doi.org/10.1038/s41559-017-0176

Buxton, R. T., Agnihotri, S., Robin, V. V., Goel, A., & Balakrishnan, R. (2018b). Acoustic indices as rapid indicators of avian diversity in different land-use types in an Indian biodiversity hotspot. *Journal of Ecoacoustics*, *2*, 1–17. doi.org/10.22261/jea.gwpzvd

Buxton, R. T., McKenna, M. F., Clapp, M., Meyer, E., Stabenau, E., Angeloni, L. M., … Wittemyer, G. (2018a). Efficacy of extracting indices from large-scale acoustic recordings to monitor biodiversity. *Conservation Biology*, *32*(5), 1174–1184. doi.org/10.1111/cobi.13119

Catchpole, C. K., & Slater, P. J. (1995). *Bird song: biological themes and variations*. Cambridge University Press.

Celis-Murillo, A., Deppe, J. L., & Ward, M. P. (2012). Effectiveness and utility of acoustic recordings for surveying tropical birds. *Journal of Field Ornithology*, *83*(2), 166–179. doi.org/10.1111/j.1557-9263.2012.00366.x

Chao, A., & Shen, T. J. (2003). Nonparametric estimation of Shannon's Index of Diversity When There Are Unseen Species in Sample. *Environmental and Ecological Statistics, 10,* 429–443. Retrieved from link.springer.com/content/pdf/10.1023%2FA%3A1026096204727.pdf

Clark, M. L., Salas, L., Baligar, S., Quinn, C. A., Snyder, R. L., Leland, D., … Newsam, S. (2023). Ecological Informatics The effect of soundscape composition on bird vocalization classification in a citizen science biodiversity monitoring project. Ecological Informatics, 75(March), 102065. doi.org/10.1016/j.ecoinf.2023.102065

Debinski, D. M., & Humphrey, P. S. (1997). An integrated approach to biological diversity assessment. *Natural Areas Journal*, *17*(4), 355–365.

Depraetere, M., Pavoine, S., Jiguet, F., Gasc, A., Duvail, S., & Sueur, J. (2012). Monitoring animal diversity using acoustic indices: Implementation in a temperate woodland. *Ecological Indicators*, *13*(1), 46–54. doi.org/10.1016/j.ecolind.2011.05.006

Dirzo, R., Young, H. S., Galetti, M., Ceballos, G., Isaac, N. J. B., & Collen, B. (2014). Defaunation in the Anthropocene. *Science*, *345*(6195), 401–406. doi.org/10.1126/science.1251817

Duelli, P., & Obrist, M. K. (2003). Biodiversity indicators: The choice of values and measures. *Agriculture, Ecosystems and Environment*, *98*(1–3), 87–98. doi.org/10.1016/S0167-8809(03)00072-0

Dumyahn, S. L., & Pijanowski, B. C. (2011). Soundscape conservation. *Landscape Ecology*, *26*(9), 1327–1344. doi.org/10.1007/s10980-011-9635-x

Eldridge, A., Guyot, P., Moscoso, P., Johnston, A., Eyre-Walker, Y., & Peck, M. (2018). Sounding out ecoacoustic metrics: Avian species richness is predicted by acoustic indices in temperate but not tropical habitats. *Ecological Indicators*, *95*(September), 939–952. doi.org/10.1016/j.ecolind.2018.06.012

Franklin, J. F. (1988). Structural and temporal diversity in temperate forests. In *Wilson E.O. Biodiversity* (pp. 166–175). Washington, D.C.: National Academy Press.

Fuller, S., Axel, A. C., Tucker, D., & Gage, S. H. (2015). Connecting soundscape to landscape: Which acoustic index best describes landscape configuration? *Ecological Indicators*, *58*, 207–215. doi.org/10.1016/j.ecolind.2015.05.057

Gasc, A., Sueur, J., Jiguet, F., Devictor, V., Grandcolas, P., Burrow, C., … Pavoine, S. (2013). Assessing biodiversity with sound: Do acoustic diversity indices reflect phylogenetic and functional diversities of bird communities? *Ecological Indicators*, *25*, 279–287. doi.org/10.1016/j.ecolind.2012.10.009

Gibb, R., Browning, E., Glover-Kapfer, P., & Jones, K. E. (2018). Emerging opportunities and challenges for passive acoustics in ecological assessment and monitoring. Methods in Ecology and Evolution, 2019(September 2018), 169–185. doi.org/10.1111/2041-210X.13101

Gómez, W. E., Isaza, C. V, & Daza, J. M. (2018). Ecological Informatics Identifying disturbed habitats : A new method from acoustic indices. *Ecological Informatics*, *45*(March), 16–25. doi.org/10.1016/j.ecoinf.2018.03.001

Gotelli, N. J., & Colwell, R. K. (2011). Chapter 4: Estimating species richness. In *Biological diversity: frontiers in measurement and assessment* (pp. 39–54).

Hao, Z., Wang, C., Sun, Z., van den Bosch, C.K., Zhao, D., Sun, B., … Pei, N. (2021). Soundscape mapping for spatial-temporal estimate on bird activities in urban forests. *Urban Forestry Urban Greening*, 57, 126822. doi.org/10.1016/j.ufug.2020.126822

Haselmayer, J., & Quinn, J. S. (2000). A Comparison of Point Counts and Sound Recording as Bird Survey Methods in Amazonian Southeast Peru. *The Condor*, *102*(4), 887–893. doi.org/10.2307/1370317

Johnson, C. N., Balmford, A., Brook, B. W., Buettel, J. C., Galetti, M., Guangchun, L., & Wilmshurst, J. M. (2017). Biodiversity losses and conservation responses in the Anthropocene. *Science*, *356*(6335), 270–275. doi.org/10.1126/science.aam9317

Joppa, L., O'Connor, B., Visconti, P., Smith, C., Geldmann, J., Hoffmann, M., … Ahmed, S. (2016). Filling in biodiversity threat gaps. *Science*, *326*, 1154–1155.

Krause, Bernard. (1987). Bioacoustics, habitat ambience in ecological balance. *Whole Earth Reviews*, *57*, 14–18.

Lellouch, L., Pavoine, S., Jiguet, F., Glotin, H., & Sueur, J. (2014). Monitoring temporal change of bird communities with dissimilarity acoustic indices. *Methods in Ecology and Evolution*, *5*(6), 495–505. doi.org/10.1111/2041-210X.12178

Machado, R. B., Aguiar, L., & Jones, G. (2017). Do acoustic indices reflect the characteristics of bird communities in the savannas of Central Brazil ? *Landscape and Urban Planning*, *162*, 36–43. doi.org/10.1016/j.landurbplan.2017.01.014

Magurran, A. E. (2004). *Measuring Biological Diversity*. Blackwell Publishing. Retrieved from http://www.wiley.com/WileyCDA/WileyTitle/productCd-0632056339.html

Martens, P., Rotmans, J., & De Groot, D. (2003). Biodiversity: Luxury or necessity? *Global Environmental Change*, *13*(2), 75–81. doi.org/10.1016/S0959-3780(02)00089-4

Matsinos, Y.G. & Tsaligopoulos, A. (2018). Hot spots of ecoacoustics in Greece and the issue of background noise. *Journal of Ecoacoustics, 2(2),* 1–10. doi.org/10.22261/jea.u3xbiy

Mckenna, M. F. (2020). The sounds around us. *Physics Today*, *73*(1), 28–34. doi.org/10.1063/PT.3.4387

Mennitt, D., Sherrill, K., & Fristrup, K. (2014). A geospatial model of ambient sound pressure levels in the contiguous United States. Journal of Acoustic Society of America, 135, 2746–2764. doi.org/10.1121/1.4870481

Morris, E. K., Caruso, T., Buscot, F., Fischer, M., Hancock, C., Maier, T. S., … Rillig, M. C. (2014). Choosing and using diversity indices: Insights for ecological applications from the

German Biodiversity Exploratories. *Ecology and Evolution*, *4*(18), 3514–3524. doi.org/10.1002/ece3.1155

Morrison, M. L. (1986). Bird populations as indicators of environmental change. In *Current Ornithology1* (pp. 429–451). Boston, MA: Springer.

Noss, R. F. (1990). Indicators for Monitoring Biodiversity : A Hierarchical Approach. *Conservation Biology*, *4*(4), 355–364.

Pekin, B. K., Jung, J., Villanueva-Rivera, L. J., Pijanowski, B. C., & Ahumada, J. A. (2012). Modeling acoustic diversity using soundscape recordings and LIDAR-derived metrics of vertical forest structure in a neotropical rainforest. *Landscape Ecology*, *27*(10), 1513–1522. doi.org/10.1007/s10980-012-9806-4

Pijanowski, B. C., Farina, A., Gage, S. H., Dumyahn, S. L., & Krause, B. L. (2011b). What is soundscape ecology? An introduction and overview of an emerging new science. *Landscape Ecology*, *26*(9), 1213–1232. doi.org/10.1007/s10980-011-9600-8

Pijanowski, B. C., Villanueva-Rivera, L. J., Dumyahn, S. L., Farina, A., Krause, B. L., Napoletano, B. M., … Pieretti, N. (2011a). Soundscape Ecology: The Science of Sound in the Landscape. *BioScience*, *61*(3), 203–216. doi.org/10.1525/bio.2011.61.3.6

Pimm, S. L., Jenkins, C. N., Abell, R., Brooks, T. M., Gittleman, J. L., Joppa, L. N., … Sexton, J. O. (2014). The biodiversity of species and their rates of extinction, distribution, and protection. *Science*, *344*(6187), 1246752. doi.org/10.1126/science.1246752

Ploog, D. (2002). Is the neural basis of vocalisation different in non-human primates and Homo sapiens? In T. Crow (Ed.), The speciation of modern Homo Sapiens (pp. 121–135). Oxford University Press.

Purvis, A., & Hector, A. (2000). Getting the measure of biodiversity. *Nature*, *405*, 212–219.

Radford, J. Q., & Bennett, A. F. (2007). The relative importance of landscape properties for woodland birds in agricultural environments. *Journal of Applied Ecology*, *44*(4), 737–747. doi.org/10.1111/j.1365-2664.2007.01327.x

Riede, K. (1993). Monitoring biodiversity: Analysis of Amazonian rainforest sounds. *Ambio*, *22*(8), 546–548. doi.org/10.2307/4314145

Rodriguez, A., Gasc, A., Pavoine, S., Grandcolas, P., Gaucher, P., & Sueur, J. (2014). Temporal and spatial variability of animal sound within a neotropical forest. *Ecological Informatics*, *21*, 133–143. doi.org/10.1016/j.ecoinf.2013.12.006

Scarpelli, M.D.A., Liquet, B., Tucker, D., Fuller, S., Roe, P. (2021). Multi-Index Ecoacoustics Analysis for Terrestrial Soundscapes: A New Semi-Automated Approach Using Time-Series Motif Discovery and Random Forest Classification. *Frontiers in Ecology and Evolution*, 9, 1–14, doi.org/10.3389/fevo.2021.738537

Schafer, R. M. (1993). *The soundscape: our sonic environment and the tuning of the world.* Simon and Schuster.

Shannon, G., McKenna, M. F., Angeloni, L. M., Crooks, K. R., Fristrup, K. M., Brown, E., … Wittemyer, G. (2016). A synthesis of two decades of research documenting the effects of noise on wildlife. *Biological Reviews*, *91*(4), 982–1005. doi.org/10.1111/brv.12207

Shonfield, J., & Bayne, E. M. (2017). Autonomous recording units in avian ecological research: current use and future applications. *Avian Conservation and Ecology*, *12*(1). doi.org/10.5751/ace-00974-120114

Snyder, R., Clark, M., Salas, L., Schackwitz, W., Leland, D., Stephens, T., … Clas, K. (2022). The Soundscapes to Landscapes Project: Development of a Bioacoustics-Based Monitoring Workflow with Multiple Citizen Scientist Contributions. Citizen Science: Theory and Practice, 7(1), 24. doi.org/10.5334/cstp.391

Southworth, M. (1969). The Sonic Environment of Cities. *Environemnt and Behavior*, *1*(1), 22.

Stoddard, M. C., & Prum, R. O. (2011). How colorful are birds? Evolution of the avian plumage color gamut. *Behavioral Ecology*, *22*(5), 1042–1052. doi.org/10.1093/beheco/arr088

Sueur, J., Farina, A., Gasc, A., Pieretti, N., & Pavoine, S. (2014). Acoustic indices for biodiversity assessment and landscape investigation. *Acta Acustica United with Acustica*, *100*(4), 772–781. doi.org/10.3813/AAA.918757

Sueur, J., Pavoine, S., Hamerlynck, O., & Duvail, S. (2008). Rapid acoustic survey for biodiversity appraisal. *PLoS ONE*, *3*(12), 1–10. doi.org/10.1371/journal.pone.0004065

Turner, A., Fischer, M., & Tzanopoulos, J. (2018). Sound-mapping a coniferous forest—Perspectives for biodiversity monitoring and noise mitigation. *PLoS One*, 13, 1–21. doi.org/10.1371/journal.pone.0189843

Wiley, R. H., & Richards, D. G. (1978). Physical Constraints on Acoustic Communication in the Atmosphere : Implications for the Evolution of Animal Vocalizations. *Behavioral Ecology and Sociobiology*, *3*(1), 69–94.

CHAPTER 2: SOUNDSCAPE CLASSIFICATION WITH CONVOLUTIONAL NEURAL NETWORKS REVEALS TEMPORAL AND GEOGRAPHIC PATTERNS IN ECOACOUSTIC DATA.

2.1. ABSTRACT

Interest in ecoacoustics has resulted in an influx of acoustic data and novel methodologies to classify and relate landscape sound activity to biodiversity and ecosystem health. However, indicators used to summarize sound and quantify the effects of disturbances on biodiversity can be inconsistent when applied across ecological gradients. This study used an acoustic dataset of 487,148 minutes from 746 sites collected over 4 years across Sonoma County, California, USA, by citizen scientists. We built a custom labeled dataset of soundscape components and applied a deep learning framework to test our ability to predict these soundscape components: human noise (Anthropophony), wildlife vocalizations (Biophony), weather phenomena (Geophony), Quiet periods, and microphone Interference. These soundscape components allowed us to balance predicting variation in environmental recordings and relative time to build a custom labeled dataset. We used these data to quantify soundscape patterns across space and time that could be useful for environmental planning, ecosystem conservation and restoration, and biodiversity monitoring. We describe a pre-trained convolutional neural network, fine-tuned with our sound reference data, with classification achieving an overall F0.75-score of 0.88, precision of 0.94, and recall of 0.80 across the five target soundscape components. We deployed the model to predict soundscape components for all acoustic data and assess their hourly patterns. We noted an increase in Biophony in the early morning and evening, coinciding with peak animal community vocalization (e.g., dawn chorus). Anthropophony increased during morning/daylight

hours and was lowest in the evenings, coinciding with diurnal patterns in human activity. Further, we examined soundscape patterns related to geographic properties at recording sites. Anthropophony decreased with increasing distance to major roads, while Quiet increased. Biophony and Quiet were comparable to Anthropophony at more urban/developed and agriculture/barren sites, while Biophony and Quiet were significantly higher than Anthropophony at less-developed shrubland, oak woodland, and conifer forest sites. These results demonstrate that acoustic classification of broad soundscape components is possible with small datasets, and classifications can be applied to a large acoustic dataset to gain ecological knowledge.

2.2. INTRODUCTION

The value of different sounds across the landscape has long been recognized as socially valuable (Schafer, 1993; Southworth, 1969), and acoustic data are becoming more economical and efficient to collect, permitting characterization of spatial and temporal patterns of biodiversity, human activity, and other sounds (Depraetere et al., 2012; Shonfield and Bayne, 2017). The acoustic quality of habitats is also recognized as a vital dimension of conservation (Dumyahn and Pijanowski, 2011; Schafer, 1993), as increasingly excessive human noise can have a range of direct deleterious effects on biodiversity (e.g., acoustic masking from overlapping communication frequency ranges) (Doser et al., 2019; Francis et al., 2017). Identifying naturally quiet landscapes and relating patterns in anthropogenic and biotic noise is essential in understanding the effects of changing human activity on biodiversity and noise reduction on conservation and management efforts of protected areas (Newport et al., 2014; Rice et al., 2020).

The unique assemblage of sounds across a landscape is collectively referred to as a soundscape (Krause, 2002; Pijanowski et al., 2011) and is treated as an ecological characterization of landscapes (Pavan, 2017, Sethi et al., 2020). Recorded soundscapes consist of anthropogenic activity (anthropophony), wildlife vocalizations (biophony), and weather-related phenomena (geophony; Pijanowski et al., 2011), along with quiet (the ambient sound or lack of acoustic events) at a given timeframe. Because soundscapes integrate the acoustic dynamics of an ecosystem, they can be considered as "community phenotypes," integrating vocalizing species, anthropogenic noise, and natural phenomena (Lellouch et al., 2014). Capturing soundscape snapshots provide meaningful insights related to biodiversity and human impacts, highlighting changes such as degraded habitats (Bush et al., 2018; Dumyahn and Pijanowski, 2011; Sueur et al., 2008). Recent work has related soundscape activity to geographic characteristics such as land-use change (Eldridge et al., 2018; Sethi et al., 2020), as well as gradients in vegetation and forest structure (Boelman et al., 2007; Dröge et al., 2021; Farina and Pieretti, 2014). Human impact has been linked to soundscape variation, including increased anthropogenic noise near high traffic roads (Doser et al., 2019), complex interactions between biophony and anthropophony in urban soundscapes (e.g., overlapping frequencies) (Fairbrass et al., 2017), soundscape similarity in oil palm production with forested soundscapes (Furumo and Aide, 2019), and impacts of snowmobile activity on winter quiet (Mullet et al., 2017b). Additional research linking patterns in sounds across landscapes and time will improve the utility of ecoacoustic methods for informing conservation and land management.

Two standard methods of assessing the information in soundscapes include deriving acoustic indices and manual identification of acoustic events (i.e., done by a human, not an algorithm).

Acoustic indices describe the acoustic energy in amplitude, time, and frequency space (Sueur et al., 2014), while manual identification provides specific time and frequency data (Grant and Samways, 2016; Rose et al., 2018). Both methods summarize sound into meaningful, interpretable ecological indicators (e.g., species diversity or acoustic complexity). However, acoustic indices vary in their ability to convey meaningful biodiversity information, making comparisons between studies or geographic regions non-trivial (Bradfer-Lawrence et al., 2019; Metcalf et al., 2021). Poor performance of acoustic indices has been attributed to the presence of confounding, background sounds in recordings, making identification and filtering of non-biophonic noise a necessary step for indices to be applied for consistent biodiversity and human impact monitoring (Depraetere et al., 2012; Eldridge et al., 2018; Fairbrass et al., 2017). Likewise, manual identification of sound sources is highly time-intensive and requires detailed knowledge of target animal vocalizations, thus, limiting this approach to smaller datasets (Pérez-Granados and Traba, 2021; Shonfield and Bayne, 2017).

Due to recent computational innovations, soundscape dynamics and their associated patterns of biodiversity can be derived with less effort and time using deep learning identification (Christin et al., 2019; Fairbrass et al., 2019), source separation (Lin and Tsao, 2020), and unsupervised classification (Sethi et al., 2020). Efforts in environmental sound classification established the effectiveness of convolutional neural networks (CNN), a type of deep learning architecture (Lecun et al., 2015), for soundscape classification (Piczak 2015; Salamon and Bello, 2017). In ecoacoustics, CNNs have been applied in species-specific vocalization (Christin et al., 2019; Kahl et al., 2018; Ruff et al., 2021) and targeted sound classification (e.g., gunshots or rain) (Metcalf et al., 2020; Sánchez-Giraldo et al., 2020). Alternatively, source separation

methods have successfully identified specific types of sound by separating sound mixtures into individual sources (Eldridge et al., 2018; Lin and Tsao, 2020). Few ecoacoustic studies have attempted to classify entire soundscapes using broadly defined sound categories such as anthropophony, biophony, geophony, and quiet. One study that classified broader soundscape components developed two CNN classifiers that modeled biophony and anthropophony in urban London (Fairbrass et al., 2019). They found that CNN models trained on a limited amount of expertly annotated sound samples outperformed multiple acoustic indices in representing human and animal activity patterns in urban soundscapes (Fairbrass et al., 2019). Other efforts have successfully modeled acoustic-environmental relationships of anthropophony, biophony, and geophony, but this required manual identification of sounds in almost 60,000 recordings (Mullet et al., 2016).

Ecoacoustic analyses typically focus on one or a few target sounds, but to characterize these sounds with confidence, other, confounding sounds need to be identified and, in some cases, removed. Some methods to account for unwanted sound include avoiding sites close to roads (Duchac et al., 2020), using pre-programmed amplitude or frequency audio filters (Bedoya et al., 2017; Duchac et al., 2020; Towsey, 2013), relating meteorological data to recordings to filter weather events (Desjonquères et al., 2018; Gasc et al., 2018), or manual identification of noises (Bradfer-Lawrence et al., 2019; Gordon et al., 2018). Comparatively, deep learning solutions can efficiently and accurately extract important data features, allowing for noise from signal separation. Products from deep learning can address many of the above issues in accounting for confounding noises and classifying soundscapes in one modeling framework (Christin et al., 2019). For example, these modeling efforts can negate the need for acoustic filtering, which is

broadly applied but is ineffective in specific environments (e.g., frequency filtering in urban landscapes) (Fairbrass et al., 2017). Deep learning models can also be more broadly applicable than traditional acoustic indicators because they can be updated when presented with new acoustic data (Fairbrass et al., 2019; Ruff et al., 2021).

Here, we present methods to classify soundscape components using a deep learning approach in Sonoma County, California, USA, an area that includes a gradient of natural and anthropogenic ecosystems. We aim to demonstrate the ability to classify broadly inclusive soundscape components: Anthropophony, Biophony, Geophony, Quiet, and physical or electronic recorder Interference (ABGQI) while accounting for modeling uncertainty using deep-learning practices and accuracy on par with current modeling efforts (Christin et al., 2019; Ruff et al., 2021). We investigate the relationship between these soundscape indicators with land use/land cover and road distance to understand soundscape variation across human impact and geographic gradients. Soundscape classification allows for (a) automated identification of unwanted sounds in large amounts of data, (b) modeling the effects and interactions of different sounds, and (c) use of modeling products themselves (e.g., classified acoustic samples) to understand spatio-temporal patterns in sound activity.

2.3. METHODS

2.3.1. Study region and acoustic data collection

The study area was Sonoma County, California, USA (38.51°N, 122.93°W), covering 4,152 km^2 north of the San Francisco metro area (Fig. 2-1). The county has a Mediterranean climate with average annual precipitation of 1,040 mm (Appendix 2A). Non-native annual grasses dominate grasslands and can be unmanaged or support beef and dairy cows. Urban areas and

agriculture (primarily vineyards) are concentrated in valleys (Fig. 2-1). The county's common vocalizing animals include multiple bird species, amphibians, and invertebrates (e.g., crickets, katydids, cicadas).

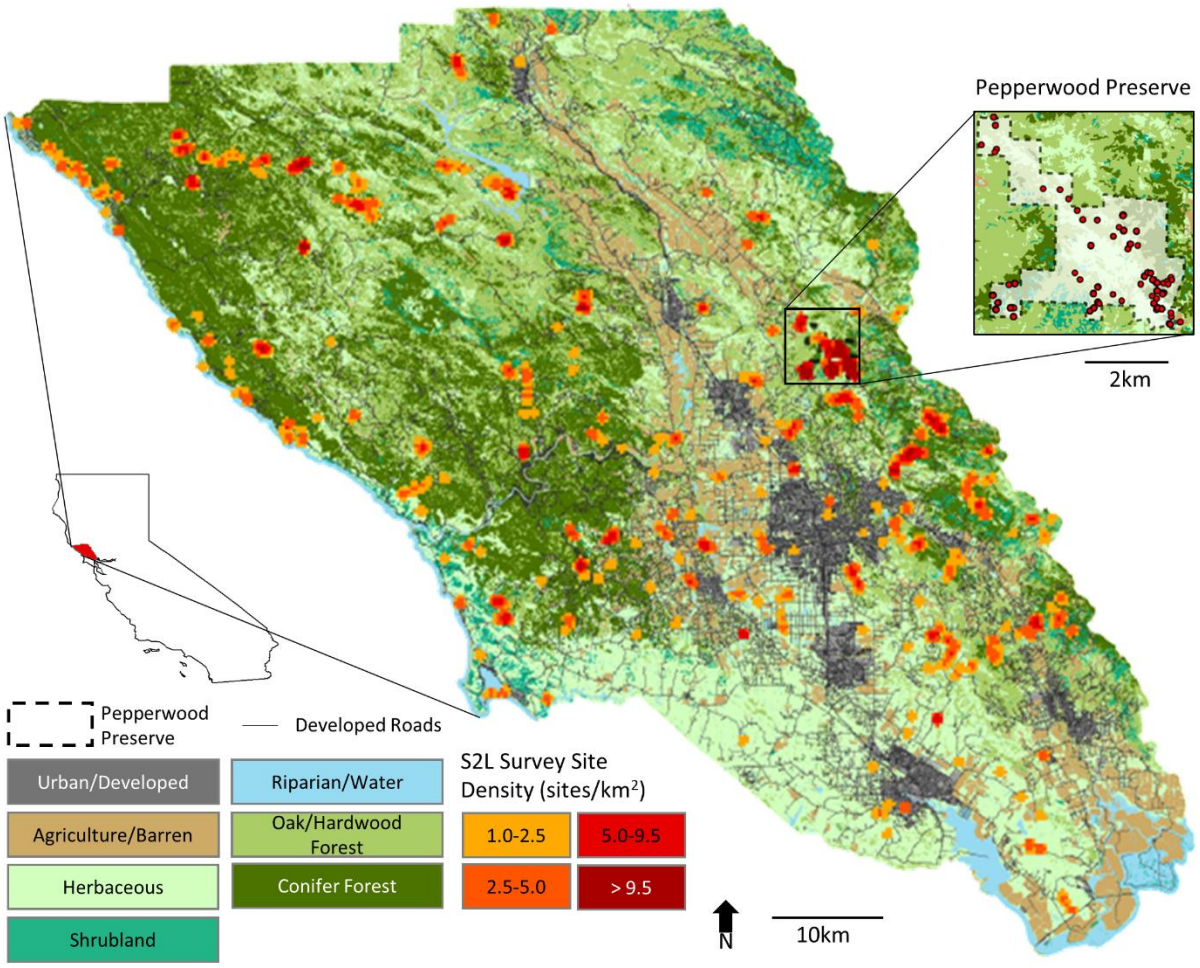


Figure 2-1. Sonoma County, California, USA land use/land cover classes derived from Sonoma County Fine-scale Vegetation and Habitat Map. Inset of Pepperwood Preserve sites. We show Soundscapes to Landscapes site location densities from 2017 – 2020 (n = 746).

Autonomous recording units (ARUs) were deployed across Sonoma County in 2017-2020 as part of the Soundscapes to Landscapes (S2L; soundscapes2landscapes.org) citizen science project. Sites were selected across the county based on a topographic lowland and highland

stratification, then broad land use/land cover (LULC) types, such as forest, shrubland, herbaceous, urban areas, and agriculture (sonomavegmap.org; Appendix 2A). This stratification scheme provided field sites with heterogeneous vegetation types, various vocalizing species, and a range of human impacts to capture diverse acoustic settings.

At each site, citizen scientists deployed a single ARU, either an Android-LG smartphone with an attached microphone in a waterproof case or an AudioMoth recording device (Hill et al., 2018) in a vinyl protective pouch. ARUs were deployed for at least 72 hours, with programming to record 60-s every 10 minutes, resulting in six minutes per hour and 144 minutes per 24-hour period. Each minute recording was saved in a waveform audio file format (.wav) with 16-bit digitization depth and 44.1 kHz or 48 kHz sampling rate for LG ARUs and AudioMoth ARUs, respectively. This sampling rate allowed for an effective upper frequency of 22.05 kHz for LG ARUs and 24 kHz for AudioMoth ARUs.

We collected 487,148 minutes (~ 8,000 hours) of sound data during bird breeding seasons (March-July) across 746 sites with an average of 642 ± 370 minutes per site (Fig. 1). The 746 sites were distributed across years unevenly: 2017 (n = 122), 2018 (n = 89), 2019 (n = 345), and 2020 (n = 190) and ARUs: LG (n = 201) and AM (n = 545). Sampling effort was not consistent across years due to an early project prototyping phase with less data (2017) and COVID-19 lockdown (2020), which delayed recorder deployment. Furthermore, sites were not evenly distributed across LULC types: agriculture/barren (n = 11), conifer forest (n = 222), oak/hardwood forest (n = 283), herbaceous (n = 147), riparian/wetland (n = 21), shrubland (n = 47), and urban/developed (n = 15) (Appendix 2A). Uneven distribution of sites across LULC types is primarily due to land accessibility in Sonoma.

2.3.2. Deep learning classification of soundscape components

Deep learning, a branch of machine learning, uses multiple hidden, non-linear transformations to automatically derive abstracted representations of raw data (e.g., images or words). Comparatively, non-deep learning modeling techniques are limited in their ability to interpret raw data without prior feature extraction by a human (e.g., amplitude or frequency) (Christin et al., 2019; Lecun et al., 2015). Here, we are interested in identifying specific types of sounds in audio recordings, represented as spectrogram images (Fig. 2). We used a CNN to which we supplied a training dataset composed of acoustic images labeled present or absent for a given class (e.g., is human noise present?). Training the model (learning) begins with extracting basic "features" (e.g., shape and texture) from the image set. These extraction steps become abstracted within the model until the final step, which attempts to classify the image by relating learned features to the known image label. After training, we used another set of labeled images withheld during model training (test set) for model selection and accuracy evaluation. We then predict the presence and absence of sounds in all recordings with the best performing model.

To identify soundscape components across Sonoma County, we used a CNN to classify five broad target classes: Anthropophony, Biophony, Geophony, Quiet, and Interference (ABGQI) (Fig. 2). The Quiet class represented periods without other soundscape components below 11 kHz (i.e., no discernible acoustic events), resulting in periods with minimal change in sound from baseline ARU self-noise levels (Fig. 2 – Quiet; Appendix 2A). Interference is defined here as broadband, short duration, electronic or physical microphone interference events, for example, when a branch hits a recorder during a strong wind gust. We classified Interference to represent recording error, while ABGQ relay ecologically meaningful information. Converting raw

acoustic data to spectrograms allowed us to use the well-established pre-trained MobileNetV2 architecture, a lightweight CNN trained on Google search imagery (ImageNet; github.com/tensorflow/models/tree/master/research/slim/nets/mobilenet) for transfer learning (Christin et al., 2019; Yosinski et al., 2014). Other ecoacoustic recognition tasks have successfully demonstrated the effectiveness of spectrograms and CNNs (Fairbrass et al., 2019; Sethi et al., 2020).

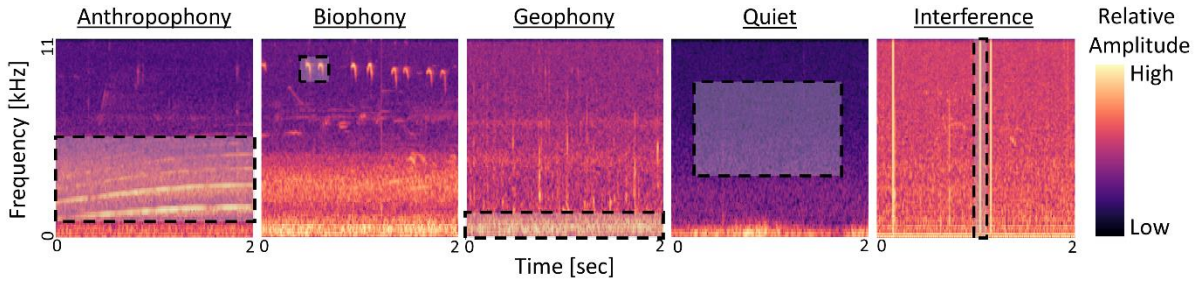


Figure 2-2. The five broad soundscape classes shown here are examples of the acoustic activity represented in Mel spectrograms. The Mel spectrograms are 2-s long (x-axis), span from 0 kHz to 11.025 kHz (y-axis), and display relative amplitude (z-axis). Manually annotated regions of interest (ROIs) are represented as black hashed and shaded boxes. See Section 2.2.3 for Mel spectrogram methods.

2.3.3. Training dataset collection

To create training data for ABGQI classification, we randomly sampled audio files from the entire S2L dataset, first stratified by LULC type and then by site. An upper limit of 350 audio files was sampled from each LULC type amounting to 2,367 sampled recordings. Sites within each LULC type were sampled equally (i.e., if urban/developed land cover contained 15 sites, 23 audio files were randomly selected from each site for 345 total audio files). We verified that random sampling within LULC and sites resulted in an even sampling across 24 hours (Appendix 2A).

Citizen scientists and team members were randomly assigned unique recordings to identify ABGQI sounds from the audio file subset by loading the recording in Raven Lite 2 audio software program (Cornell Lab of Ornithology, Cornell University, Ithaca, NY; ravensoundsoftware.com). Audio files were then annotated with regions of interest (ROIs; Figs. 2 & 3), or discrete sound events, by visualizing spectrograms with a standard frequency range of 0-12 kHz and spectrogram window size of 512 samples. We chose a cutoff duration of 5-s because some sound events are not temporally discrete (e.g., constant wind or vehicle traffic) and limited the length of ROIs to no more than 5 ROIs of the same class per audio file to account for oversampling and biasing individual sound files. C.Q. then verified that every ROI was a true presence. This process resulted in 5,396 ROI annotations in 1,194 of the 2,367 subset recordings with an average count of 2.77 ± 3.56 ROIs per recording with a presence (Table 1). A list of the target sounds is in Appendix 2A and ROI collection methods at (doi.org/10.5281/zenodo.6027024).

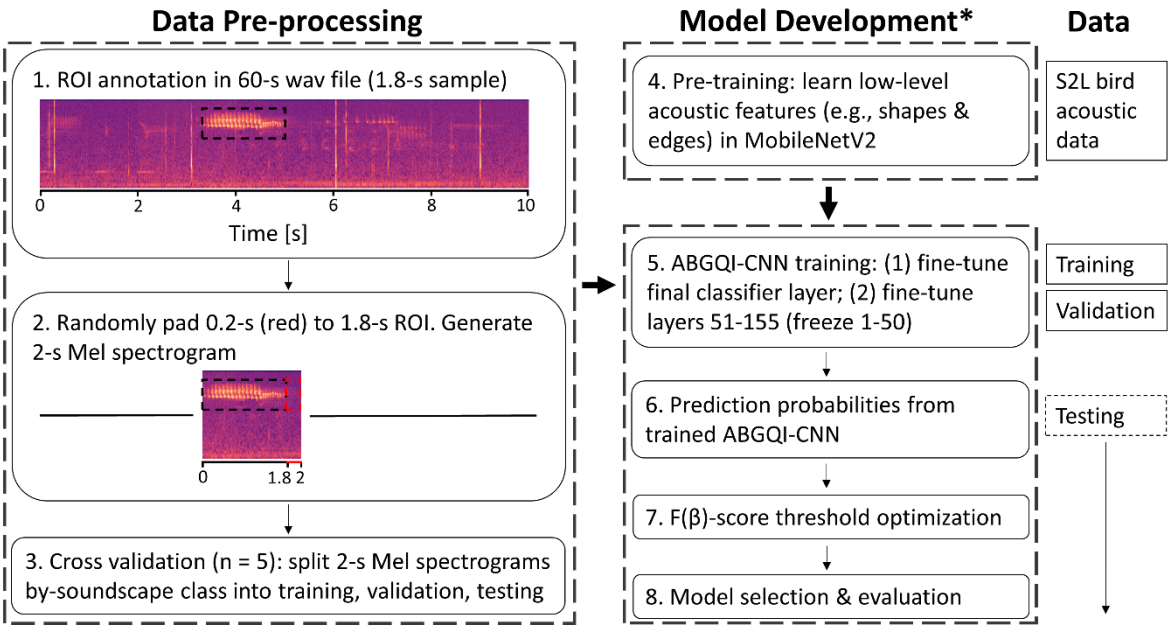


Figure 2-3. Our workflow for data preprocessing (1-3) and model development (4-8). *(5-8) performed on each cross-validation data split.

Table 2-1. The number of regions of interest (ROIs) from S2L and Freesound data, the number of 2-s Mel spectrogram (Mel spec) samples created from all ROIs, and the final 2-s training set size for each soundscape class.

Sound Class	S2L ROIs	Freesound ROIs	Mean ROI length (s)	Total 2-s Mel spec	Training set size
Anthropophony	916	883	1.67	2,170	1,920
Biophony	2,372	556	1.88	3,336	3,086
Geophony	957	801	1.60	1,955	1,705
Quiet	765	0	3.06	1,023	773
Interference	386	0	2.05	430	330
TOTAL	**5,396**	**2,240**	**-**	**8,914**	**7,814**

2.3.3.1. Auxiliary Freesound data

CNN classification training typically requires thousands to millions of images or hundreds of hours of acoustic data (Christin et al., 2019; Knight et al., 2017). We added ROI samples using open-access Freesound data because there were fewer than 1,000 ROIs per class in our ABGQI ROI data (freesound.org; file data at doi.org/10.5281/zenodo.6027024). C.Q. listened to each Freesound file for quality (e.g., clarity and presence of single, unmixed sound) because recordings were long and variable (60.25 ± 75.13-s) and were sometimes poorly labeled. We used the autodetec function in the package warbleR (Araya-Salas and Smith-Vidaurre, 2017) in the statistical software program R (R Core Team, 2020) to isolate sound events in recordings. If an event was evident in the spectrogram (e.g., consistent with the general characteristics of the soundscape component), it was included in the ROI dataset resulting in the addition of 2,240 Freesound ROIs (Table 1).

2.3.3.2. Spectrogram generation and cross-validation

We chose to use 2-s Mel spectrogram representations of ROIs to balance short-duration acoustic events (e.g., bird chirps) with longer-duration events (e.g., vehicle traffic) (Zhong et al., 2020). We clipped and combined all ROIs into a single synthetic acoustic file for each soundscape component. ROIs > 2-s were directly added to these recordings, whereas ROIs < 2-s were padded randomly around the ROI center time for a total of 2-s (Fig. 3; e.g., if an ROI was 1.5-s, 0.5-s from the audio file would be randomly added before or after the ROI). This process meant short ROIs had additional surrounding context, and padding could introduce unknown sounds, reflecting the model's real-world application. Because we directly added ROIs > 2-s to each soundscape component's synthetic recording, some 2-s spectrogram segments contained multiple unique ROIs. Synthetic soundscape component recordings were spliced into 2-s segments and converted into Mel spectrograms using python 3.7.7's librosa 0.6.3 library (McFee et al., 2015; Python Software Foundation, 2016). Although ARUs sampled at a rate of 44.1 kHz or higher (effective reproduction of frequencies below 22.05 kHz), spectrograms were generated from 0 kHz to 11.025 kHz (half the Nyquist frequency), regardless of ROI frequency range. The upper-frequency limit overlapped with most acoustic indices and allowed for frequency shifts above the 0-8 kHz frequency window where most anthropogenic and low-frequency animal vocalizations occur (Boelman et al., 2007; Kasten et al., 2012; Sueur et al., 2014). The resulting image was 432x432 pixels in red-green-blue color bands representing amplitude levels (Fig. 2). Spectrograms were down-sampled to 224x224 pixel, 3-banded images for CNN training, testing, and deployment; the default input size for the pre-trained CNN.

In CNN modeling practices, training data are used for learning data features, while validation data are used for model assessment and tuning hyper-parameters during training. In contrast,

testing data are withheld to perform an independent model performance evaluation and generate accuracy estimates. We randomly sampled spectrograms from each sound class into validation (n = 200), testing (n = 50), and training (all remaining component spectrograms; Table 1) datasets five times for a five-fold cross-validation training approach (Fig. 3). Cross-validation provided a measure of model performance to strengthen model evaluation given the small dataset size. We decreased the number of Interference validation samples to n = 50 because of the class's low sample count. We had 850 total validation samples and 250 total testing samples. The model with the highest class-specific accuracy in cross-fold validation was selected for evaluation and prediction to the countywide dataset. We recognized the small size of the test set but prioritized training performance, provided evaluation metrics in a cross-fold validation framework, and provided an independent accuracy test on a soundscape dataset in Appendix 2A.

2.3.3.3. CNN transfer learning

We developed the CNN in Python (v.3.7.7, Python Software Foundation, 2016) using the TensorFlow v.2.3.0 framework (Abadi et al., 2015). The fully trained CNN model and related code are available at doi.org/10.5281/zenodo.6112713. CNN training is most effective when sizeable labeled training datasets are available (Christin et al., 2019). However, when large labeled datasets are unavailable, or we desire faster training time, features developed in existing CNNs can be "transferred" to the new dataset through transfer learning (Yosinski et al., 2014). We tested pitch shift augmentation but ultimately did not include any form of data augmentation in the final CNN as it required increased computation with minimal change in accuracy (Salamon and Bello, 2017; Piczak, 2015) (Appendix 2A).

We applied concepts from transfer learning using the pre-trained MobileNetV2 architecture. First, we pre-trained the ImageNet weighted MobileNetV2 using existing labeled acoustic data of calls from 54 bird species from the S2L project, some recordings overlapping ABGQI data (Fig. 3). We used S2L bird data ROIs to learn low-level spectrogram features from an existing, large dataset similar to the domain of interest (Appendix 2A). Following acoustic feature pre-training, we trained the S2L-MobileNetV2 CNN with ABGQI data (Appendix 2A). This modeling framework leveraged previously learned low-level spectrogram features from S2L bird data while modifying weights to learn high-level, ABGQI-specific features. The fully trained CNN produced a vector of five probabilities using a binary cross-entropy classifier with sigmoid activation (one value for ABGQI each). We refer to the fully-trained model as the ABGQI-CNN.

2.3.4. Classification threshold optimization

During model training, we used a custom threshold process, and testing data to (1) create binary classified values (present=1 / absent=0) for all classes in each 2-s spectrogram and (2) assess model performance. Threshold value choices should be made on a study-to-study basis depending on the research question (Knight et al., 2017). Each class was optimized individually using the 50 presences and 200 absences from the testing data by setting threshold values from 0 to 1 in 0.0001 increments, forcing probabilities below the threshold to zero/absent and probabilities above the threshold to one/present, and selecting the threshold where the F(β)-score was maximized. F(β)-score (Eq.1):

$$F(\beta) = \frac{(1+\beta^2) * (precision * recall)}{(\beta^2 * precision) + recall} \quad \text{[Eq.1]}$$

Where precision = TP/(TP+FP) and recall = TP/(TP+FN); TP is true positive, FP is false positive, and FN is false negative.

Generally, prioritizing precision is favored to minimize false positives (Kahl et al., 2021; Knight et al., 2017; LeBien et al., 2020), and automated audio classification tasks can result in excess false positives (Balantic and Donovan, 2020). Prioritizing precision over recall means that classifications are conservative estimates for each class; i.e., there are fewer false positives and many true positives, but at the cost of increased false negatives. We used soundscape component CNN predictions and threshold values based on maximum F(β)-scores to produce binary classifications (Fig. 5). This framework allowed each 2-s spectrogram to have more than one class present, resulting in some 2-s spectrograms without a present soundscape prediction. In these cases, we created a sixth post-classification category, "Unidentified," which occurred when ABGQI were absent in a spectrogram. We tested three β values: 0.50, 0.75, and 1.00. β values below 1.00 weight precision higher than recall while 1.00 weights recall equal to precision (Kahl, 2020). Following CNN threshold selection, we used changes in evaluation metrics (F(β)-score, precision, recall) and the rate of Unidentified samples in the test dataset to select the optimal F(β)-score model.

2.3.5. ABGQI-CNN model deployment

We deployed the modeling framework for inference on the entire S2L acoustic dataset to classify ABGQI and the Unidentified class (Fig. 4). Recordings used in the ABGQI-CNN training process (n = 1,195 recordings) constituted 0.25% of the entire S2L audio dataset. In our subsequent analyses, we removed these recordings to account for any bias in model deployment. The model deployment included the generation of 2-s spectrograms, ABGQI-CNN probability

predictions, and classification with optimized prediction thresholds (Fig. 4). Deployment resulted in binary ABGQI classifications for every 2-s spectrogram (n = 14,578,590), indicating the presence or absence of each of the five classes or unidentified. We calculated percent time present, the number of predicted 2-s class presences relative to total 2-s spectrograms, using presence/absence values aggregated hourly and by-site for a rate of sound present. For example, if a site with 100 2-s clips had five 2-s presence, the rate was 0.05 or 5%.

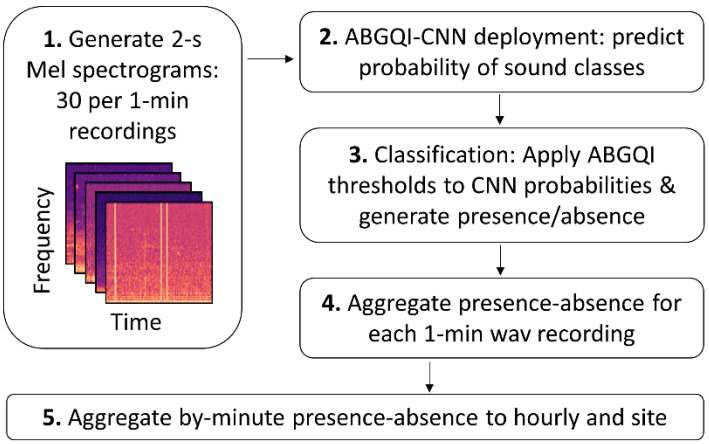


Figure 2-4. Workflow chart for the deployment of ABGQI-CNN on the entire S2L recording dataset.

2.3.6. Sound pattern statistical analyses

2.3.6.1. Covariate data

We used ABGQI percent time present to examine temporal variation in soundscape recording site characteristics. The goal of these analyses was to examine how classified sounds reflect (1) expected patterns across the landscape, (2) site characteristics (e.g., does Anthropophony vary with distance from road), and (3) temporal sound variation (e.g., diel sound patterns). Covariates in these analyses included the majority LULC type within a 50-m radius of recorder location, distance to the nearest road, temporally relevant measures of recording (i.e., diel, monthly, and

annual patterns), and meteorological (MET) station data. We extracted LULC data using aggregated classes from the Sonoma County Fine-scale Vegetation and Habitat Map (sonomavegmap.org). Road distance was calculated as the linear distance from a site to the nearest improved road (i.e., paved or high-use roads; gis-sonomacounty.hub.arcgis.com). We used meteorological data from Pepperwood Preserve's five MET stations (Ferrell et al., 2021a, 2021b) with 15-min resolution data co-occurring near 56 S2L sites at Pepperwood (n = 39,552 recordings).

2.3.6.2. Comparative analyses

We used hourly estimates of ABGQI to analyze diurnal patterns in rates of sounds (24-hour and night vs. day) among LULC types. We used the by-site estimates of ABGQI daytime data (5 a.m. to 8 p.m.), a temporal subset that focused on animal and human activity co-occurrence, to compare ABGQI percent time present with overlapping deployment dates (day of the year), LULC, and road distance. Overlapping deployment dates (each site's first day of recording) were used to analyze whether there were annual differences in soundscape activity. This aspect is essential because, for example, annual start dates may overlap each year differently with events such as bird breeding season and therefore capture different amounts of Biophony. Comparisons among LULC types helped discern any diagnostic ABGQI patterns within these ecosystems. We further explored how these relationships are modulated by distance from the nearest road. Road distance was treated as a grouped variable to examine patterns in sound in 100-m intervals (e.g., 0-99m, 100-199m) from 0-1000m and >1000m. We took hourly averages of wind speed data from Pepperwood MET stations to the nearest recording site and compared these data to hourly

amounts of soundscape components using linear regression (i.e., does Geophony co-vary with wind speed).

Because sample sizes and variances among treatment categories were not homogeneous, we used a Kruskal-Wallis test to evaluate effects for tests with more than two samples. If a Kruskal-Wallis test was significant (large χ^2 statistic and z-values and $p < 0.05$ indicate considerable differences are present), we then applied Dunn's simultaneous multiple-comparison test with a Bonferroni correction to account for the inflated error rate to identify which pairwise relationships were significantly different (when p-adj < 0.05). We used Mann-Whitney's U test for two-sample tests (i.e., day and night soundscape activity). We reported all pairwise test results in Appendix 2A.

2.3.6.3. Modeling of soundscape components

We used main-effects-only linear regression models to find underlying factors and drivers of the amount of predicted soundscape components. Descriptive covariates included road distance (log-transformed), number of total 2-s samples at the site (log-transformed), LULC types, ARU type, year, and type of sound (i.e., ABQGI). The year was included to capture year-unique effects, such as deployment date or changes in human activity. The response, amount of sound, was the percent of positively predicted 2-s samples for each ABGQI as a logit score and included all soundscape data. For example, if Biophony was predicted in 80 of 100 2-s samples, this would result in a 0.8 positive rate (80%), converted to a logit value of 1.39. The initial model included all candidate variables, quadratic and cubic road distances, and interactions between road distance and soundscape type and LULC. We used the function stepAIC from the package MASS in R (Venables and Ripley, 2002) to select the optimal model (Appendix 2A).

2.4. RESULTS

2.4.1. Model performance

We used the highest performing ABGQI model that included pre-training with bird vocalization and auxiliary Freesound data (F0.75-score = 0.883 ± 0.035) for evaluation and deployment (optimal model results in Table 2). Performance was determined from cross-validation results. Cross-validation had consistent accuracy across iterations (F0.75-score = 0.838 ± 0.03; precision = 0.902 ± 0.03; recall = 0.754 ± 0.03) indicating that our sample sizes were sufficient to learn spectrogram features (i.e., consistent performance across folds). Pre-training with S2L bird vocalization data resulted in the default MobileNetV2 learning acoustic-specific features and increased average F0.75-score from 0.659 to 0.756, average precision from 0.643 to 0.790, and a decrease in average recall from 0.712 to 0.704. Adding Freesound data to ABGQI data training, the default MobileNetV2 resulted in an overall increase in average F0.75-score from 0.659 to 0.760, average precision increased from 0.643 to 0.847, and average recall decreased from 0.712 to 0.672. Combined, pre-training with S2L bird vocalization and auxiliary Freesound data increased the average F0.75-score from 0.659 to 0.883. The optimal F(β)-score value of $\beta = 0.75$ was chosen for threshold values (Fig. 5) based on a balance between prioritizing precision over recall and accounting for the number of unidentified 2-s spectrograms (see Appendix 2A 1 for cross-validation threshold evaluation). We have provided an additional measure of model accuracy and generalizability in Appendix 2A.

Table 2-2. Optimal model performance was based on independent test data for each soundscape class using recommended evaluation metrics. We assessed model

classification performance with the F0.75-score. We also report F1-score and area under the curve (AUC) for comparison with other studies.					
Soundscape Class	Precision	Recall	F0.75-score	F1-score	AUC
Anthropophony	0.975	0.780	0.894	0.867	0.939
Biophony	0.913	0.840	0.885	0.875	0.984
Geophony	0.923	0.720	0.838	0.809	0.959
Quiet	0.891	0.820	0.864	0.854	0.984
Interference	0.977	0.860	0.932	0.915	0.980
AVERAGE	**0.936**	**0.804**	**0.883**	**0.865**	**0.969**

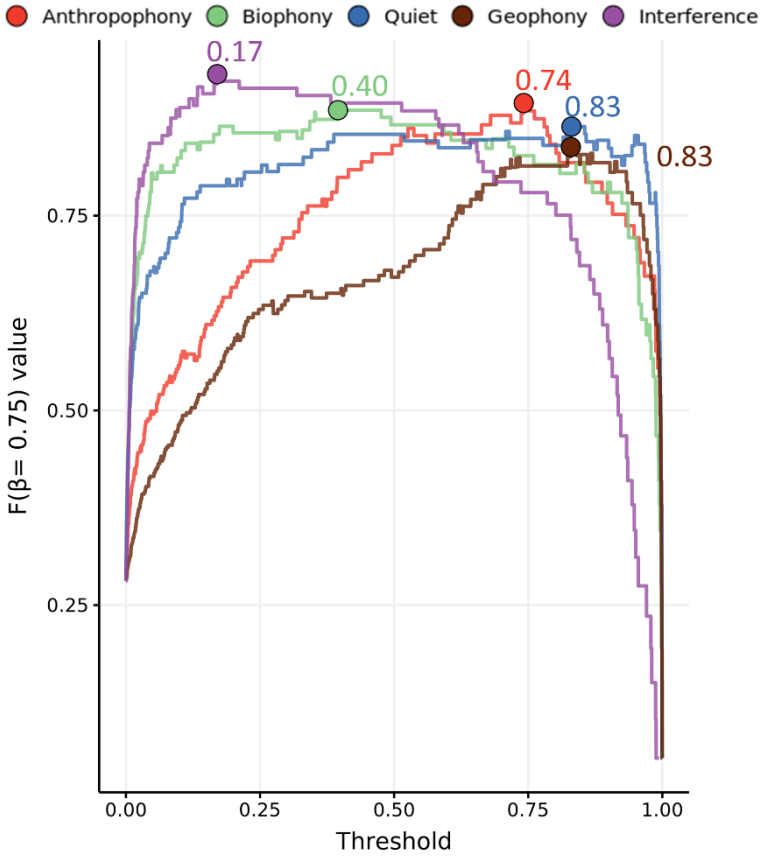


Figure 2-5. Threshold optimization values for maximum F0.75-score (circles), optimized independently for each soundscape class.

2.4.2. Statistical analyses of soundscape components

The final ABGQI-CNN predicted the highest site average hourly amount of Quiet (μ = 24.6, σ = 27.7%), followed by Biophony (μ = 24.2, σ = 24.8%), Interference (μ = 9.8, σ = 17.4%),

Geophony ($\mu = 8.5$, $\sigma = 12.9\%$), and Anthropophony ($\mu = 5.5$, $\sigma = 9.8\%$), while Unidentified averaged 29.3 ± 17.5%. See Appendix 2A for a summary of sounds in Unidentified samples.

2.4.2.1. Diurnal LULC patterns

Mann-Whitney's U tests revealed significant differences in the amount of Anthropophony, Biophony, Geophony, and Quiet during nighttime hours (8 p.m. to 5 a.m.) compared to daytime hours (5 a.m. to 8 p.m.; sample size and test results in Appendix 2A). Anthropophony was lowest during nighttime hours on average and was approximately two times higher during the day (μ = 6.79%; $U = 2.364 \times 10^7$, $p < 0.001$). Similar doubling patterns in activity between night and day were observed for Biophony (night = 15.4% and day = 29.5%; $U = 1.917 \times 10^7$, $p < 0.001$), while Quiet had the opposite pattern (night = 37.3% and day = 16.6%; $U = 5.032 \times 10^7$, $p < 0.001$). These patterns were consistent when visualizing data partitioned by LULC, but we did not test for within-LULC differences.

Biophony peaked between 5-8 a.m., depending on LULC, and gradually decreased until a local daytime minimum at 3-4 p.m. for all LULCs (Fig. 6). There was a slight rise in activity at 8 p.m. for all LULC types except oak/hardwood and conifer forests, and lower activity throughout all nighttime hours to a global minimum at 4 a.m. for all LULCs. Quiet was highest from 3-4 a.m. and lowest (i.e., there was the highest soundscape activity) at 10 a.m., opposite Biophony and Anthropophony activity. Maximum Geophony generally occurred in the afternoon between 12-3 p.m., while the minimum occurred during early-morning (5-7 a.m.) and evening (8-10 p.m.). Interference was highest between 11 a.m. and 3 p.m., implying more broad frequency spikes occurred later in the day regardless of LULC type.

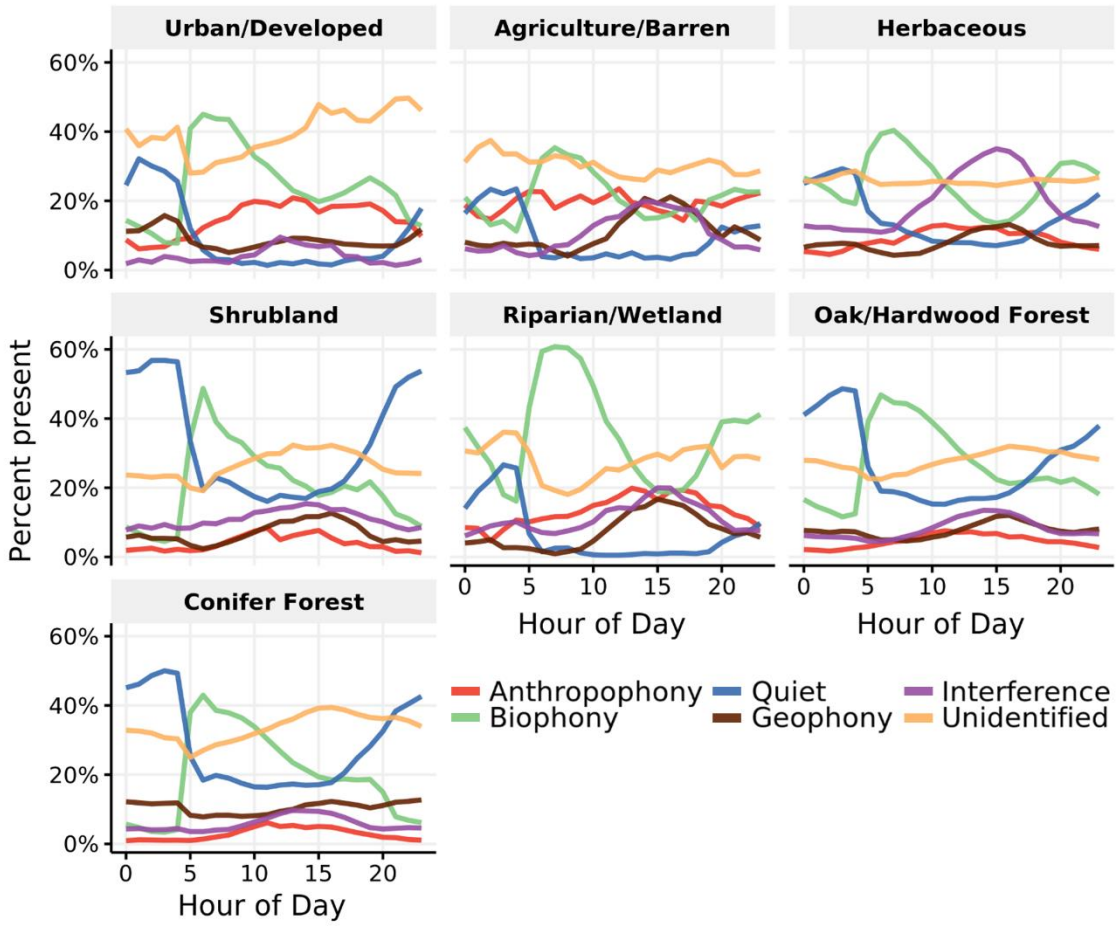


Figure 2-6. Soundscape component classifications aggregated by the hour for all 746 recording sites. Sites were stratified by LULC type, and lines represent the average percent of predicted ABGQI and Unidentified over each hour interval.

2.4.2.2. Annual and deployment date differences

Kruskal-Wallis and Dunn tests revealed significant differences (p-adj < 0.05) in Anthropophony ($\chi^2 = 35.049$, $p < 0.001$), Biophony ($\chi^2 = 29.582$, $p < 0.001$), and Quiet ($\chi^2 = 75.744$, $p < 0.001$) among years when tested on data from only the overlapping range of annual deployment dates (May-01 to July-05; Appendix 2A). These differences reveal year effects in our results. Limiting data to overlapping dates of deployment resulted in significant loss of observations when data were stratified by LULC or road distance (annual data decreased by:

2017 = 46.7%, 2018 = 48.3%, 2019 = 44.0%, 2020 = 17.9%; total sites = 460 / 746). We note this inter-annual variance and acknowledge that differences in start and end dates of survey campaigns among years may add to this variance. Therefore, we use the entire datasets only for analyses that do not require accounting for a year effect (i.e., sections 3.2.3 - 3.2.5 and 3.3).

2.4.2.3. Daytime LULC stratification

Kruskal-Wallis tests using daytime recordings (5 a.m. to 8 p.m.) revealed significant differences among LULC types for Anthropophony ($\chi^2 = 97.798$, $p < 0.001$), Biophony ($\chi^2 = 18.891$, $p = 0.004$), and Quiet ($\chi^2 = 109.92$, $p < 0.001$), but not Geophony (Table S.15.6). Dunn test results showed which LULC types were significantly different within soundscape components (Fig. 7; Appendix 2A).

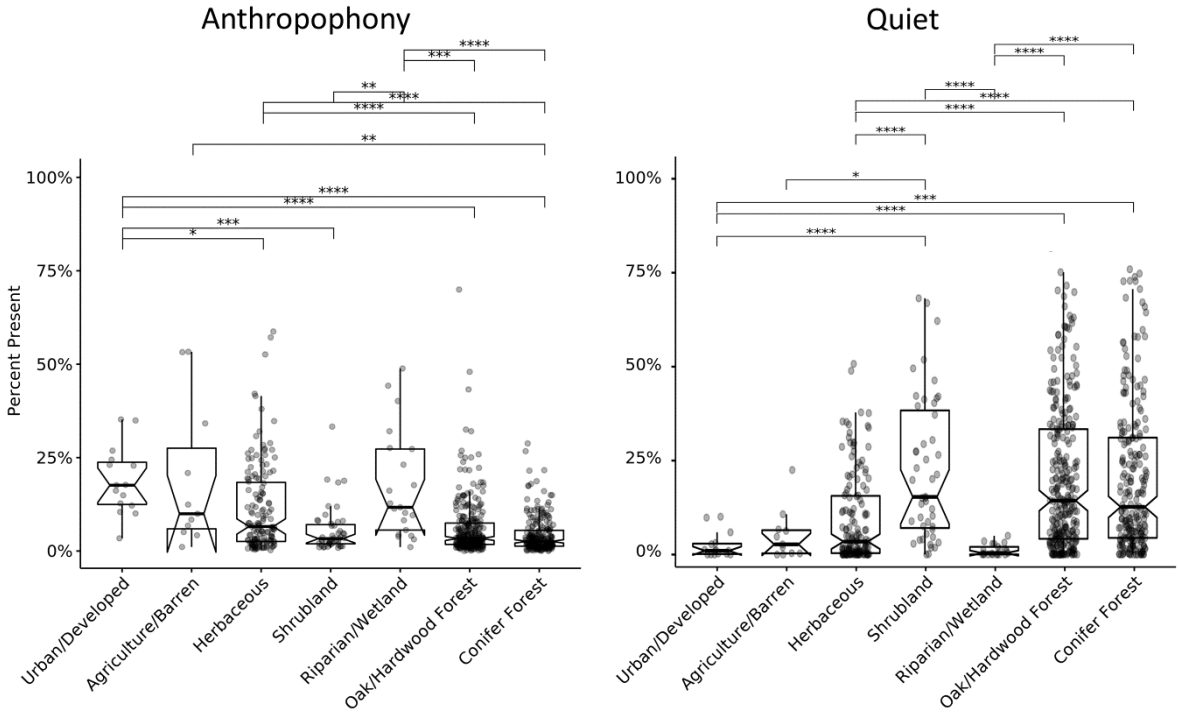


Figure 2-7. Acoustic dissimilarity patterns of sites grouped by LULC for daytime recordings (5 a.m. to 8 p.m.). Significance notation of pairwise Dunn test results on the overhanging brackets, non-significant pairs not annotated, as follows: * p-adj < 0.05; ** p-adj < 0.005; *** p-adj < 0.0005; **** p-adj < 0.00005.

There was a weak but significant difference in Biophony between herbaceous and oak/hardwood forest sites (z value = 30.745, p-adj = 0.044). Overall, the most Biophony occurred at riparian/wetland sites and the least at agriculture/barren sites. Other LULC types had comparable Biophony to each other.

The amount of Anthropophony in urban/developed sites was greater and differed significantly from herbaceous (z value = -3.163, p-adj = 0.0328), shrubland (z value = -4.506, p-adj < 0.001), oak/hardwood forest (z value = -4.993, p-adj < 0.001), and conifer forest (z value = -5.931, p-adj < 0.001) Anthropophony. We observed more Anthropophony in herbaceous and riparian/wetland sites compared to both oak/hardwood (z value = -4.579, p-adj < 0.001 and z value = -4.380, p-adj < 0.001, respectively) and conifer forests (z value = -6.818, p-adj < 0.001 and z value = -5.475, p-adj < 0.001, respectively). Agriculture/barren Anthropophony was higher than conifer forests as well (z value = -3.783, p-adj = 0.003). Hourly patterns in Anthropophony showed the most magnitude change (i.e., the largest difference between the minimum and maximum amount present) in urban/developed sites while riparian sites had a similar, but less pronounced pattern (Fig. 6). Other less human-impacted LULC types, such as oak/hardwood forest, conifer forest, shrubland, and herbaceous, showed similar patterns, but with lower overall amounts of Anthropophony, while agriculture/barren sites showed no discernable diel patterns.

Quiet was lowest and co-occurred with Anthropophony in more human-impacted urban/developed, agriculture/barren, and riparian/wetland sites compared to overall higher amounts of Quiet in shrubland, oak/hardwood forest, and conifer forest sites, while herbaceous sites fell between these two groups (Fig. 7). When spikes of Biophony co-occurred with persistent anthropogenic noise, Quiet correspondingly decreased to near-zero (urban/developed,

agriculture/barren, and riparian/wetland). Quiet never rose above 35% presence at these more human-impacted sites, whereas at less human-impacted sites, Quiet reached over 50% in the evenings and rarely dropped below 15-20% during the day.

2.4.2.4. Road Distance

Kruskal-Wallis tests using daytime recordings (5 a.m. to 8 p.m.) were significant among road-distance classes for Anthropophony ($\chi^2 = 48.543$, $p < 0.001$) and Quiet ($\chi^2 = 46.521$, $p < 0.001$) and not significant for Geophony and Biophony (Fig. 8; sample size and test results in Tables S.15.8 and S.15.9). Anthropophony was highest at sites closest to roads than sites farther from roads (e.g., 0-99m and >1000m [z value = -5.064, p-adj < 0.001]; 100-199m and >1000m [z value = -3.707, p-adj = 0.012]). Similar trends existed for Quiet where sites closer to roads had less Quiet than sites farther from roads (e.g., 0-99m and >900-999m [z value = 3.725, p-adj = 0.011]; 0-99m and >1000m [z value = 4.757, p-adj < 0.001]; Appendix 2A). We observed a higher ratio of Biophony to Anthropophony with increased road distance (Fig. 8).

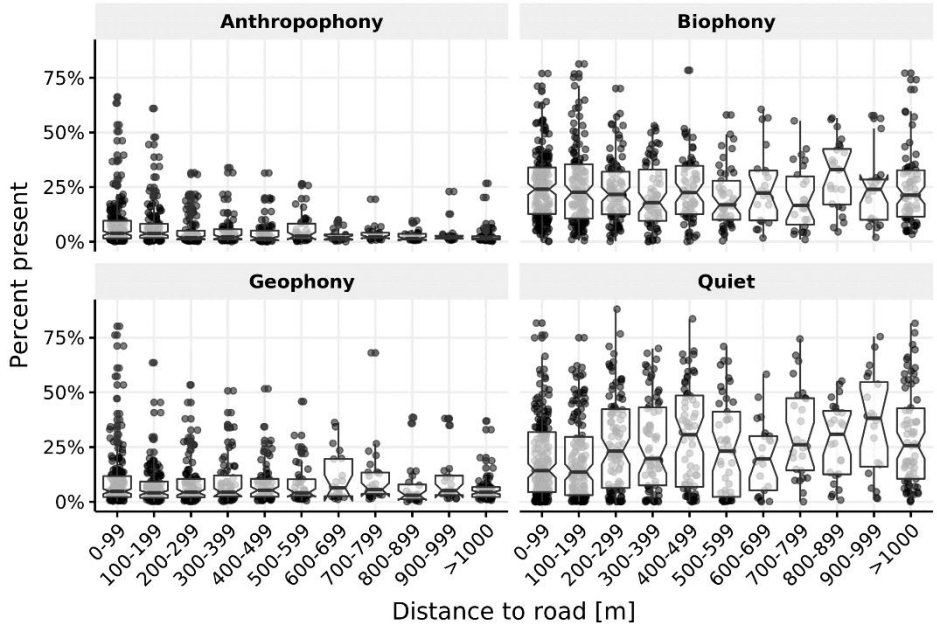


Figure 2-8. Soundscape components grouped by distance from roads for daytime recordings (5 a.m. to 8 p.m.). Notches reflect confidence intervals ($\pm \mathbf{1.58} * \boldsymbol{IQR}/\sqrt{\boldsymbol{n}}$) around the median.

2.4.2.5. Effect of wind speed on soundscapes

Wind speed was positively related to the amount of Geophony ($R^2 = 0.032$, t-value = 6.32, p = 0.072), Anthropophony ($R^2 = 0.007$, t-value = 2.677, p = 0.008), and Interference ($R^2 = 0.020$, t-value = 5.065, $p < 0.001$). Quiet was negatively related to wind speed ($R^2 = 0.007$, t-value = -3.093, p = 0.002) while Biophony did not have a significant relationship (p = 0.2336). Geophony rose throughout the afternoon, which reflects wind activity that also peaks, on average, later in the day. However, Interference also follows this pattern of rising throughout the day until an afternoon peak, and in some LULCs exceeds Geophony.

2.4.3. Factors affecting the amount of soundscape components

Stepwise variable selection and likelihood ratio tests resulted in a regression model containing recording year, the number of 2-s Mel spectrograms per site, soundscape component (ABGQI), LULC, road distance, and the interaction between soundscape class and road distance (Table 3; adjusted $R^2 = 0.25$, $F(19, 3691) = 66.36$, $p < 0.001$). The amount of sound increased with each progressive field season and was weakly related to the number of recordings. Anthropophony decreased the most with distance from roads, Biophony and Geophony decreased to a lesser degree, and Interference and Quiet increased (Fig. 8). LULC was not significant in determining the amount of sound, but including LULC resulted in a model with an AIC score decreased by 15.

Table 2-3. Estimated variable coefficients (Estimate), standard errors (SE), t-values, and p-values for all parameters in the final regression model for the

amount of sound. The amount of sound is the logit rate of 2-s files predicted present for each soundscape component. The reference level is the Year 2017, Anthropophony, and urban/developed. Statistically significant factors in bold (p-value < 0.05).

Variable	Estimate	SE	t-value	p-value
(Intercept)	**-3.646**	**0.684**	**-5.332**	**< 0.001**
Year18	0.049	0.098	0.502	0.615
Year19	**0.246**	**0.074**	**3.342**	**< 0.001**
Year20	**0.308**	**0.080**	**3.844**	**< 0.001**
log(number Mel spectrograms)	**0.123**	**0.062**	**1.967**	**0.049**
Biophony	**1.039**	**0.384**	**2.707**	**0.007**
Geophony	**-0.886**	**0.384**	**-2.309**	**0.021**
Interference	**-1.624**	**0.384**	**-4.232**	**< 0.001**
Quiet	**-1.104**	**0.386**	**-2.855**	**0.004**
LULC-Agric./Barren	0.090	0.273	0.329	0.742
Herbaceous	0.32	0.190	1.695	0.090
Shrubland	0.2612	0.207	1.259	0.208
Riparian/Wetland	0.203	0.239	0.846	0.398
Oak/Hardwood Forest	0.166	0.187	0.889	0.374
Conifer Forest	-0.036	0.187	-0.190	0.849
log(Road Distance)	**-0.265**	**0.049**	**-5.392**	**< 0.001**
Biophony:log(Road Distance)	**0.202**	**0.069**	**2.940**	**0.003**
Geophony:log(Road Distance)	**0.266**	**0.069**	**3.867**	**< 0.001**
Interference:log(Road Distance)	**0.350**	**0.069**	**5.090**	**< 0.001**
Quiet:log(Road Distance)	**0.528**	**0.069**	**7.624**	**< 0.001**

2.5. DISCUSSION

We optimized the development of a soundscape classifier with comparable accuracy to other ecoacoustic, deep learning tasks (Appendix 2A; Fairbrass et al., 2019; Ruff et al., 2021) and expanded these previous efforts to include novel soundscape classes (i.e., Geophony, Quiet, ARU Interference). Additionally, we demonstrated that broadly classified soundscape components reveal systematic acoustic patterns about time (e.g., hourly and annual) and geographic (i.e., road distance, LULC, and wind) properties. We chose to first focus on understanding and communicating model assumptions to contribute to the responsible use of deep learning methods in ecoacoustics (Wearn et al., 2019). Our model can be re-trained, and methods readily extended to other environments (doi.org/10.5281/zenodo.6112713).

2.5.1. Deep learning model implementation

Creating a high-performance classification model with a limited dataset is a vital step in expanding the application of CNNs for ecology and conservation where large labeled datasets or computing resources are not available (Salamon et al., 2017; Wearn et al., 2019). We provide evidence that an existing CNN (MobileNetV2) can be pre-trained and fine-tuned with soundscape data from our study site and applied to ecoacoustic problems using a small training dataset (e.g., target classes with approximately 1,000 samples (Çoban et al., 2020)). Other studies that have classified similar soundscape components achieved similar precision and recall for Biophony and Anthropophony (Appendix 2A) but have lower accuracy or do not classify Geophony, Interference, or Quiet (e.g., Fairbrass et al., 2019; Mullet et al., 2016). We used a custom classification and threshold optimization approach to understand how predictions differed based on the choice of F-score threshold values. Using F(β)-scores with $\beta < 1.00$ allowed us to prioritize precision, which can prevent underfitting the model (Abdi and Hashemi, 2016), a known issue with small datasets and is a preferred approach in soundscape classification tasks (LeBien et al., 2020). Elusive or rare species classification would comparatively benefit by maximizing recall to avoid missing infrequent vocalizations (MacLaren et al., 2018; Shiu et al., 2020).

Ecoacoustic classification may be subject to spatial and temporal autocorrelation sources in model assessment (Ploton et al., 2020), which has not been thoroughly assessed. Recent work demonstrated spatial (Holgate et al., 2021; Shaw et al., 2021) and temporal (Scarpelli et al., 2021) autocorrelation patterns in acoustic recordings and acoustic indices. However, no work to our knowledge has investigated sources of autocorrelation in a deep learning model bias context.

Future work would benefit from an investigation into sources of autocorrelation related to optimistically biased accuracy metrics in model assessment.

2.5.1.1. Class imbalance in training deep learning models

Small training datasets can result in bias when detecting underrepresented classes (Christin et al., 2019; Wearn et al., 2019). A class with low training samples results in poorly constrained model parameters and, therefore, poor class generalization (Wearn et al., 2019). In our case, we had low membership for Interference: n = 430 or 12.9% of the majority class size (Biophony), meaning our ABGQI-CNN may not have fully captured Interference features with the same confidence as other classes. We attempted to address the issue using a minority class oversampling method to bring the class membership of smaller classes up to the maximum membership class (Mohammed et al., 2020). This method resulted in higher performance for Interference at the cost of lower precision for larger classes, such as Biophony and Anthropophony. Because a priority of this classification task was to understand ecologically meaningful soundscape classes, we chose not to use balanced classes in training. We recommend that similar studies using CNN classification be aware of how a class imbalance in small datasets influences results, even after observing high evaluation metrics. Comparatively, if classes of interest have significantly low membership and are important (i.e., rare or cryptic species), techniques to augment or increase class membership are highly advised to reduce model bias.

2.5.1.2. Accounting for instrument error and sound detection

We observed non-trivial amounts of Interference (approximately 10% of recordings) and suggest acoustic studies be aware of ARU deployment techniques and instrument sensitivity to these events (e.g., strapping the unit to a tree to prevent banging in the wind). We believe rapid

broad frequency interference should be appropriately accounted for and can negatively influence automated methods, for instance, that compare background noise to changes in acoustic energy (e.g., the acoustic complexity index) (Pieretti et al., 2011).

In terms of Unidentified sounds, our model performed worse in urban/developed sites with an increase late in the day and when Biophony decreased (Fig. 6). The inability of our CNN to detect all forms of sounds is partly due to our training dataset containing single-label data, not multi-label mixed-signal spectrograms, and limited samples of some Anthropophony sounds (i.e., composed of 74% vehicle traffic and machinery; Appendix 2A). Further, Anthropophony had the most Freesound samples relative to S2L samples in our training set (i.e., Freesound = 883 ROIs: S2L = 912 ROIs), possibly resulting in underfitting. The use of data from multiple ARUs and training with auxiliary Freesound data may reflect the lower performance observed in our soundscape validation (Appendix 2A) and more Unidentified samples in the S2L dataset (29.3%) relative to the test dataset (9.2%). Increasing training data specifically from the S2L dataset would decrease potential generalizability to other datasets but increase our performance. Without these data, we can lower the number of Unidentified samples and generate less conservative predictions using a more lenient threshold optimization metric like F1-score.

Variation from ARU models was considered and found insignificant in the multivariate regression of the amount of sound when we included the effects of other site characteristics. However, we observed a consistent pattern of less pixel variation in spectrograms in LG versus AudioMoth ARUs, resulting in more Quiet in 2017 and 2018, years with LG deployments (note: the CNN was trained with ROIs from both devices). We did not apply a correction but recognized that some variation in soundscape components is due to underlying variation in ARU

sound detectability. To properly investigate where the effects of ARU variation originate (e.g., signal to noise ratio, sensitivity, sampling rate), a paired ARU deployment would control for landscape characteristics and possibly allow for a future correction to acoustic data from varying ARUs. We could not investigate these differences with the current dataset as ARUs were deployed at different sites with different recording lengths. Future research on soundscape differences among LULC types with variable structure and topography may provide insights into differences in sound detection among ARUs (Rappaport et al., 2020).

2.5.2. Soundscape patterns

2.5.2.1. Annual variation and recording length

The lowest Anthropophony was observed in 2020, which may reflect the reduction in human activity at a landscape scale during the onset of social lockdowns due to COVID-19, agreeing with reduced urban noise levels at a local scale (Aletta et al., 2020). Even though Biophony was highest in 2020, trends based on prior years make it difficult to say if this is related to the COVID-19 impact. Sampling effort was positively related to the amount of sound detected (Table 3), implying that extended recording times result in a higher rate of detections. This finding aligns with recent work that recommends extended ARU deployment to compute acoustic indices (Bradfer-Lawrence et al., 2019).

2.5.2.2. Anthropophony

Increases in Anthropophony throughout the day coincided with higher human activity (e.g., car and airplane traffic) and decreased activity in the evenings (Francis et al., 2017), reflecting our expectation that urban areas have the most human activity in Sonoma County. However, Anthropophony decreased in urban areas at nighttime to slightly above other, less human-

impacted areas (e.g., riparian/wetlands and herbaceous). Decreased nighttime Anthropophony is likely because Sonoma County urban areas are mostly suburban, surrounded by rural areas, and is peripheral to more densely populated areas in the San Francisco Bay Area. Higher Anthropophony in riparian/wetland areas may be due to the location of these sites in the Laguna de Santa Rosa area of the county, which is crossed by east-west road corridors. Agriculture/barren sites, primarily vineyards in Sonoma, most likely have a large amount of anthropogenic noise associated with machinery (Lie et al., 2016), little impact from diurnal human activity patterns, and less vegetation structural and compositional complexity, which could negatively affect vocalizing animal species resulting in less Biophony (Burns et al., 2020; Dröge et al., 2021). For example, tropical ecoacoustic studies found that more structurally-complex vegetation can maintain soundscape diversity in agricultural landscapes (Dröge et al., 2021; Gage et al., 2015; Villanueva-Rivera et al., 2011). We observed persistent anthropogenic noise in agriculture/barren areas, which may be why we observe the least Biophony, particularly in the dawn chorus; however, urban areas still experience a large Biophony dawn chorus when Anthropophony is lower. We can increase Anthropophony's recall to capture more anthropogenic noise to improve our understanding of these patterns.

2.5.2.3. Biophony

The highest Biophony reflects the dawn chorus when increased avian and mammal activity occurs (Krause and Farina, 2016) and the dusk chorus when insect and amphibian activity occurs (Gasc et al., 2018; Krause and Farina, 2016; Naguib and Riebel, 2014). Dusk chorus Biophony is most pronounced in herbaceous and riparian/wetland sites, indicating increased insect and amphibian activity at these LULCs. Even though Anthropophony was lower in conifer forest,

oak/hardwood forest, and shrubland sites, we did not observe more Biophony relative to sites with high Anthropophony, an inconsistent finding compared with prior ecoacoustic studies (Doser et al., 2019; Francis et al., 2017). However, lower Biophony in forests may be a product of detectability in denser structural landscapes (Rappaport et al., 2020).

Although Biophony did not increase with decreased human impact, we observed a shift in the relative amounts of Biophony to Anthropophony with less human impact. Most of the day, agriculture/barren and urban/developed sites were the only LULC types with comparable or higher amounts of Anthropophony relative to Biophony. Lack of variation in Biophony across LULCs indicates it is not as sensitive to LULC type as Anthropophony but can still capture the systematic dawn and dusk choruses and reflects high activity in riparian/wetland areas. Not capturing a decrease in Biophony in human-impacted areas may be due to our data not including many sites from urban centers in the county, being more oriented towards forested and lower-impacted landscapes. Notably, Biophony does not differentiate activity between, for example, a repeatedly vocalizing frog compared and a complex dawn chorus with multiple animal species vocalizing. Future work can utilize Biophony as a pre-filter for more refined taxonomic group classification or species with sufficiently labeled data.

2.5.2.4. Quiet

Naturally quiet landscapes are spaces where anthropogenic noise is absent, and invasive biophonic species are minimal (Dumyahn and Pijanowski, 2011; Pavan, 2017). Quiet places have a range of benefits, such as increased wildlife reproductive success and human well-being, highlighting the need to conserve and maintain naturally quiet landscapes (Buxton et al., 2019; Dumyahn and Pijanowski, 2011). We are limited in making inferences regarding naturally quiet

landscapes because we chose to set the upper frequency of analyses to 11 kHz. This choice was made to extend the ABGQI-CNN to other ecoacoustic work, such as examining the relationships between soundscape components and acoustic indices. Many acoustic indices are limited in their upper-frequency range from 8–11 kHz (e.g., Boelman et al., 2007; Kasten et al., 2012). This frequency range allows indices to capture numerous wildlife vocalizations (e.g., frogs, crickets, and birds [0.2-8 kHz]; Villanueva-Rivera et al., 2011) and anthropogenic noise (typically 0-2 kHz; Joo et al., 2011) while prior work has shown a significant amount of environmental acoustic activity occurs below 9-12 kHz (Metcalf et al., 2021; Pavan, 2017; Towsey, 2013). There may be other signals above 11 kHz (i.e., insects and bats) that our modeling efforts did not reflect, which would result in lower amounts of Quiet and higher Biophony. However, Anthropophony, Geophony, Interference, and a significant amount of Biophony occurred below this frequency threshold; therefore, we interpret times of Quiet to reflect periods with generally lower acoustic activity.

Quiet is most usefully interpreted alongside Biophony and Anthropophony, the former contributing to naturally quiet landscapes while the latter deteriorates these landscapes. Quiet was highest at night when human, weather, and most biotic activity were lowest (Mullet et al., 2017b). Conversely, the loudest time of day coincided with the dawn bird chorus and times of persistent anthropogenic activity (e.g., commuter and air traffic). Nevertheless, even though Biophony is negatively related to Quiet, biotic noises tend to be more culturally valuable and have positive effects on human well-being than human noise intrusion (Dumyahn and Pijanowski, 2011; Krause, 2002). Although Biophony was consistent across LULC types, increases in Anthropophony in currently less-impacted LULC types could negatively affect

wildlife communities by increasing fitness costs for species (Francis and Barber, 2013; Mullet et al., 2017a), which could have cascading adverse effects on ecosystems. Deep learning classifiers with finer taxonomic groups or species could measure animal community composition not captured by our general Biophony class and potentially reveal differences between naturally quiet and noisier landscapes impacted by human activity and highlight species that may be robust against noise pollution (Slabbekoorn and Ripmeester, 2008). Future data collection can be aided by identifying times or spaces with increased Quiet to ensure ARUs are deployed with a higher likelihood to record ecologically meaningful signals (e.g., after 5 a.m. in conifer forests).

2.5.2.5. Geophony and Interference

We observed systematic afternoon increases in Geophony in agriculture/barren and riparian/wetlands with less pronounced patterns in other LULC types. In general, we observed relatively low Geophony, and our data may not have captured rain events as recording seasons begin at the onset of Sonoma County's dry season in May. Geophony predominantly reflects wind patterns, with low confidence classifying streams or rain. We believe Interference may serve as a proxy for wind-related Geophony activity based on (1) a weak but correlated pattern with Geophony from the Pepperwood Preserve analysis and (2) evidence from training set creation where we frequently observed Interference events co-occurring with gusts of wind. If we follow this hypothesis, combining Geophony and Interference signals reflects afternoon peaks in sound activity related to meteorological patterns in the wind (Fig. 6). This additive pattern is evident in less structurally-complex LULCs (agriculture/barren, herbaceous, shrubland, riparian/wetland), most notably in herbaceous sites. In herbaceous sites, microphones were attached to small temporary poles and were subject to wind more than other LULC types, where

ARUs were generally mounted on larger tree stems. At sites where the deployment of ARUs results in minimal physical shielding, wind-based Interference can be non-trivial and lead to numerous occurrences of Interference noise events. Interference requires proper identification and consideration in future acoustic analyses (e.g., acoustic indices), so events are not treated as ecologically meaningful signals and can be used to confirm the occurrence of wind. The ABGQI-CNN can reliably identify these recording errors.

2.5.3. Soundscape patterns related to roadway proximity

Roads are fixed sources of anthropogenic noise, and extensive research has documented the deleterious effects of traffic on animal communities (Barber et al., 2011; Buxton et al., 2019; Doser et al., 2019; Ware et al., 2015). Ecoacoustic work has examined patterns in sound relative to traffic noise and found that biotic sound activity is inversely correlated with traffic intensity and distance to roads (Doser et al., 2019; Pieretti and Farina, 2013). However, these studies used either acoustic indices (Pieretti and Farina, 2013), approximated traffic noise from geographic road use data (Barber et al., 2011; Doser et al., 2019), or required intensive manual categorization of spectrograms (Buxton et al., 2019) to quantify traffic levels. Instead, we were able to quantify distinct patterns in anthropogenic noise related to distance from roads among LULCs, demonstrating the benefits of deep learning classification coupled with geospatial analysis.

Our linear regression modeling results support general findings that anthropogenic noise decreases with distance from roads (Buxton et al., 2019; Mullet et al., 2016), but we do not find there is a significant inverse pattern with Biophony as other studies note (Doser et al., 2019), except when we look at LULC types individually (shrubland and oak/hardwood forest areas).

Sites are rarely more than 1,500m from major roads (n = 26, 3.5% of all sites), and there appears to be a heavy influence on patterns from these distant sites on analytical trends. At these distances, topographic position (e.g., on a hilltop or in a valley) may have a more significant effect on soundscapes than sites closer to roads that do not experience high levels of sound attenuation (Lyon, 1973; Yip et al., 2017). Anthropophony in riparian/wetland sites was similar to urban/developed and agriculture/barren sites; however, site distance from roads is significantly higher for riparian (492 ± 392m, n = 21) than urban/developed (40 ± 20m, n = 15) or any other LULC type. Higher Anthropophony in these riparian/wetland sites may be due to proximity to principal road corridors in the Laguna de Santa Rosa area of the county where sound can travel farther along less vegetative and morphologically complex riparian/wetland corridors (Wiley and Richards, 1978).

These findings indicate that we can detect general soundscape patterns relative to distance from roads. However, to better explain the causation of these patterns and relate soundscape classes to biodiversity, we need to incorporate other spatial characteristics of the landscape, such as topography, forest structure, climate, and other human impact layers in a more holistic modeling framework. Urban and environmental planning related to road construction can benefit from quantifying meaningful soundscape components (e.g., ABQ) with these or similar methods prior to, during, and following development to measure the effects of increased human activity on naturally quiet landscapes.

2.6. CONCLUSION

Autonomous sound recording is becoming a more prominent tool for ecological monitoring. However, the abundance of acoustic data requires analytical approaches to account for error and

non-biotic sounds. The ability to rapidly train and deploy a classifier to accurately identify almost 70% of a large acoustic dataset with 93% precision, as we found here, enables ecoacoustic researchers to study broad patterns and interactions of sounds within a soundscape. Identified soundscape components can serve as promising ecoacoustic indicators that: (1) reflect temporal and environmental factors that can be used to limit noisy human activities to times when the impact on wildlife is minimized**,** (2) aid in conservation and management efforts to prioritize at-risk landscapes, and (3) optimize recorder deployment to capture ecologically-meaningful acoustic signals. Furthermore, the ABGQI-CNN and these data can help filter wanted or unwanted sounds to optimize sound monitoring and discriminate meaningful acoustic events when applying acoustic indices, which can be impacted by the presence of Anthropophony, Geophony, or Interference, reducing their measurement value. In summary, we have shown that it is possible to identify ARU error with Interference, quantify areas rich in vocalizing animal activity with Biophony, understand variations in human noise using Anthropophony, and identify quieter landscapes with data products generated from our ABGQI-CNN modeling approach.

2.7. APPENDIX 2A

2.7.1. Sonoma County geography information and Soundscapes 2 Landscapes site stratification

Seasonal average precipitation in Sonoma County is 247 mm in spring [March – May], 9 mm in summer [June – August], 185 mm in fall [September – November], and 600 mm in winter [December – February] (NOAA, 1981–2010 U.S. Climate Normals: Ground stations in the county). Natural vegetation in Sonoma County includes evergreen conifer (Coastal redwoods [Sequoia sempervirens], Douglas fir [Pseudotsuga menziesii]) and broadleaf forests (Tan oak [Notholithocarpus densiflorus], Coast live oak [Quercus agrifolia]) in western coastal mountains, as well as mixed forests of deciduous (Black oak [Quercus kelloggii], Oregon white oak [Quercus garryana], California buckeye [Aesculus californica]) and evergreen broadleaf trees (Coast live oak), conifers (Douglas fir) and shrubs.

We used stratified random sampling to identify locations for ARUs across Sonoma County. The geographic information systems (GIS) software ArcGIS Pro software was used to create strata based on county-wide GIS data (i.e., terrain, streams, land cover), canopy chemistry (i.e., chlorophyll, nitrogen, lignin, water) metrics from summer 2017 airborne hyperspectral imagery (Clark & Kilham, 2016), and forest structure metrics (derived using LAStools) from 2013 airborne lidar. We first stratified the county into upland and lowland zones using a digital elevation model. A county land-cover map (http://sonomavegmap.org/data-downloads/) was used to separate annual croplands, developed areas, grasslands, native forests, orchards, shrublands, vineyards, urban-wildland, and other areas. Further, riparian corridors were delineated as 25-m from lidar-derived streams. Forests were separated into six chemical and

structural variation levels based on principal component analysis applied to multi-seasonal hyperspectral and lidar metrics, respectively. Many of these random sample points fell in inaccessible terrain (e.g., in a ravine or atop a steep hill); therefore, when deploying ARUs, citizen scientists chose a subset of the random sample points on each property based on navigation feasibility. When no stratified sample points fell on the property (usually due to small property size), or when none of the sample points were accessible, we defined a set of parameters for the citizen scientist to use to select a site on the property: (1) away from the road and house, (2) >50 meters from any bird feeders on the property, (3) feasible to navigate to, and (4) use no a priori knowledge regarding bird activity.

2.7.2. Sampling effort – LULC by-year count of S2L sites

Table 2A-1. Count of LULC sites by year

Year	Urban/ Developed	Agriculture/ Barren	Herbaceous	Shrubland	Riparian/ Wetland	Oak/Hardwood Forest	Conifer Forest	Total
2017	2	4	15	8	0	63	30	122
2018	0	0	20	6	6	30	27	89
2019	10	3	78	25	15	111	103	345
2020	3	4	34	8	0	79	62	190
Total	15	11	147	47	21	283	222	746

2.7.3. ROI recording file sampling distribution by the hour of the day

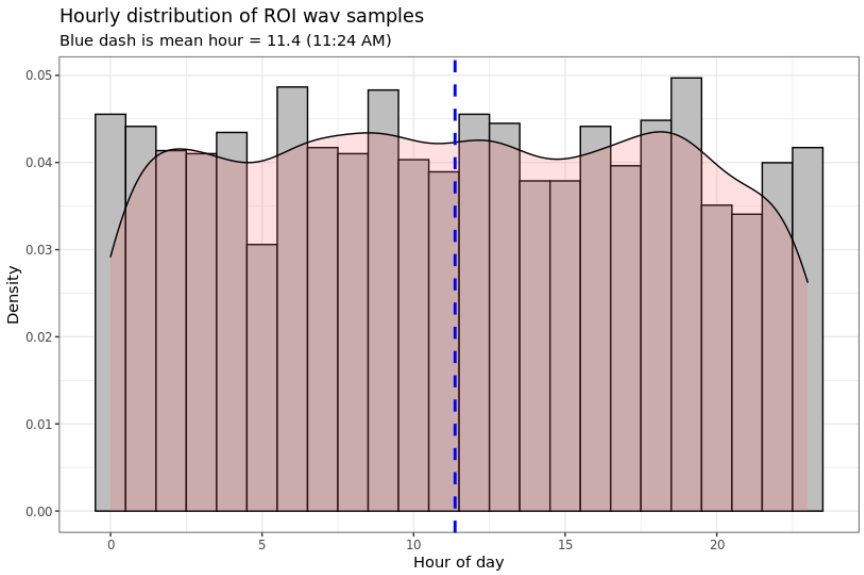


Figure 2A-1. Density of recordings grouped by hour of day.

2.7.4. Regions of interest sound class composition

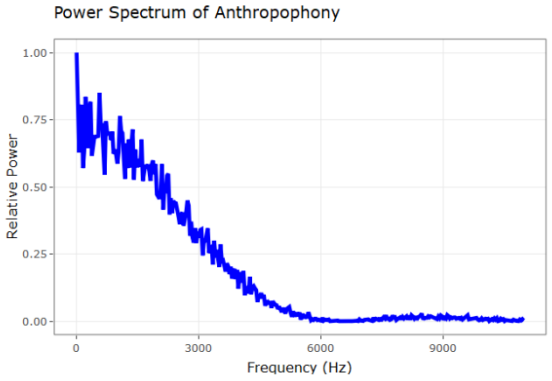
Table 2A-2. Target soundscape classes and specific sounds that compose each class are used to annotate regions of interest. Total spectrogram samples, including Freesound, are noted in parentheses.

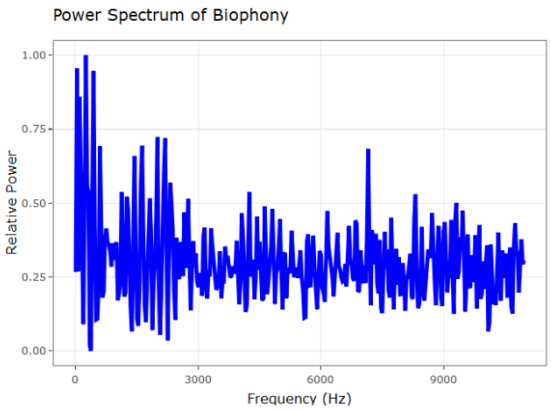
Sound Class	Included sounds
Anthropophony	Air traffic (n = 398), vehicle horn (n = 53), vehicle traffic (n = 828), rail traffic (n = 23), siren (n = 17), machinery (e.g., generators, chain saws; n = 781), human voice (n = 70)
Biophony	Birds (n = 1,745), insects (n = 667), amphibians (n = 778), mammals (n = 146)
Geophony	Rain (n = 415), wind – constant (n = 484), wind – gust (n = 432), stream (n = 589), ocean (n = 35)
Quiet	Little to no perceptible noise, i.e., background (n = 1,023)
Interference	Broad frequency physical or electronic recording disturbance (n = 430)

2.7.5. Sound class normalized spectral power characteristics – power and frequency

Power spectral density plots display power along the frequency domain used for the CNN (0-11 kHz). They were calculated using the discrete fast Fourier transform (Dfft) from the package

fftw (version 1.0-5) and a 32-ms window from which power was derived using $\frac{|Dfft|^2}{n\ samples}$. This calculation was done on a single, concatenated .wav file containing all 2-s ROI segments for each model class. These plots demonstrate the frequency regions with high and low levels of power. To note, quiet contains a majority of its signal power in low-frequency regions (<500 Hz) primarily due to the inherent noise from the recording device. Comparatively, other sound classes vary in power along with the frequency range. Anthropophony has high power at low frequencies, fading to minimal power above 6 kHz. Biophony shows the recorder activity at low frequency and then a range of activity through higher frequencies. Geophony displays most of its power below 1 kHz with a 7-9 kHz spike. Interference has a consistent amount of power long the frequency range reflecting its characteristic broadband frequency signature.





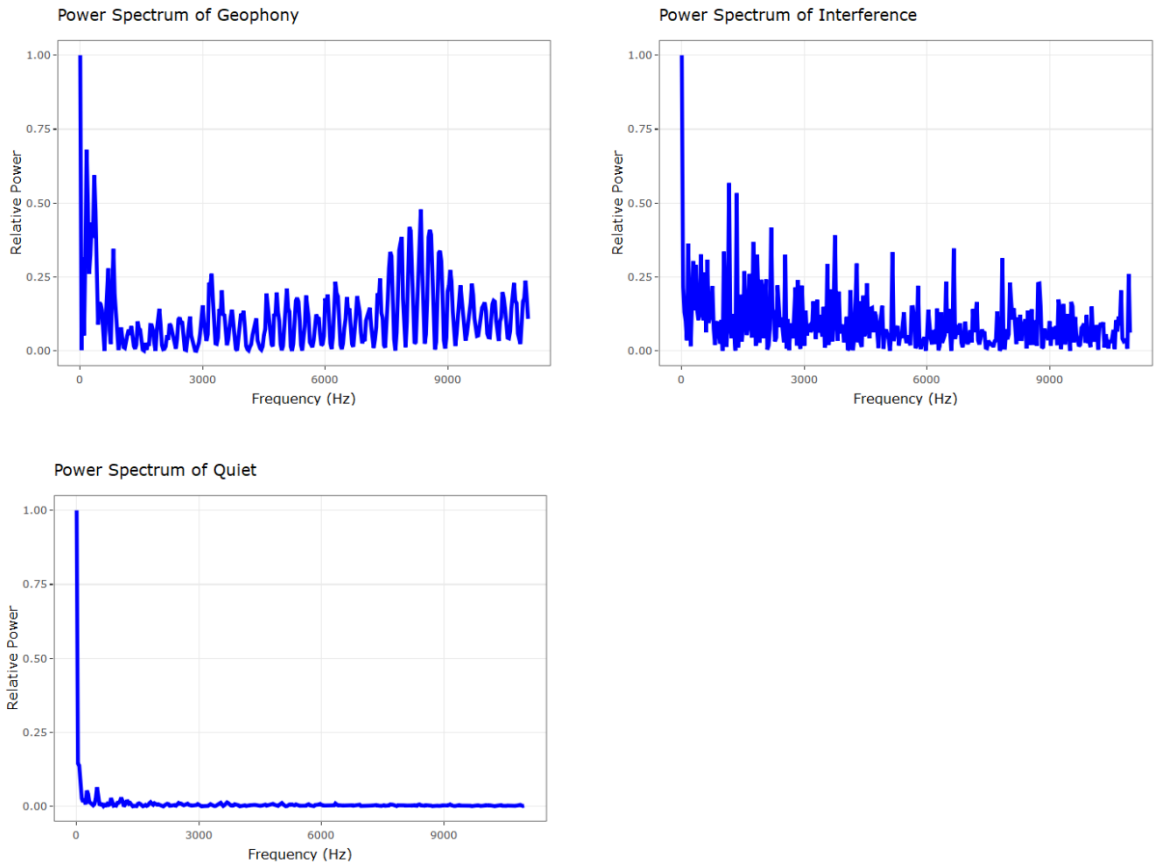


Figure 2A-2. Power spectra for soundscape component training data.

2.7.6. ROI identification and Mel Spectrogram generation methods

2.7.6.1. Mel Spectrogram creation

The 2-s Mel spectrograms were created using Python's librosa library, version 0.6.3, using an HTK implementation with the following parameters: max frequency = 11,025 Hz, sampling rate = 22,050, fast Fourier transfer window = 728, 32 samples per frame (a temporal resolution control), and 128 Mels displayed on the linear frequency scale (converting the power spectrogram to decibel (dB) units).

2.7.6.2. Freesound event detection and segmentation

Freesound filenames are provided in the project Zenodo repositories. Each Freesound file was listened to for quality by C.Q. (i.e., was the intended sound present) and relative clarity (i.e., was the intended signal identifiable or mixed significantly with other signals). We identified potential ROIs with R version 4.0.2 through RStudio version 1.3.1093 using autodetec function in warbleR package (Araya-Salas ad Smith-Vidaurre, 2017), which provided start and stop times of possible ROIs. We used start and end times of events to generate Mel spectrograms in Python, following the same methods as S2L ROI segmentation.

2.7.7. Augmentation experiment

We applied a single augmentation, shifting the pitch of samples using librosa.effects.pitch_shift. The pitch was offset using a random value from a Gaussian normal distribution (mean = 0, sd = 0.25) (Lasseck 2019). These data were included in only the training data split – increasing the dataset from 7,814 to 15,628 samples (augmented files were not included from the validation or test datasets). Accuracy was higher with augmentation using the same CNN architecture, S2L bird vocalization pretraining, and Freeesound data (F0.75 score increase of 0.070).

Table 2A-3. Evaluation metrics from test data following recording augmentation

	Precision	**Recall**	**F0.75-score**
Anthropophony	1.000	0.880	0.961
Biophony	0.939	0.920	0.929
Geophony	0.978	0.880	0.926
Quiet	0.960	0.960	0.960
Interference	0.980	0.980	0.980
Avg.	0.970	0.924	0.953

However, because our spectrogram generation and test dataset code were not adapted for augmentation of the dataset following training, the training dataset includes augmented spectrograms of all the non-augmented testing data. This may positively bias our testing metrics. To provide an independent assessment of how augmentation influences model performance, we used the soundscape validation dataset (S.12) to compare with the non-augmented performance (ABGQI-CNN). We did not utilize augmentation at this time because soundscape validation performance was only marginally improved at the cost of a doubling in training computation time. However, other augmentation, especially mixed signals or multilabel data, could improve performance and is worth further investigation.

Table 2A-4. Evaluation metrics derived from the soundscape-validation dataset following recording augmentation

	Precision	**Recall**	**F1-score**
Anthropophony	0.564	0.562	0.563
Biophony	0.937	0.827	0.879
Geophony	0.670	0.691	0.680
Interference	0.682	0.905	0.778

2.7.8. S2L bird ROI data for MobileNetv2 pre-training

The Soundscapes to Landscapes project has amassed ~750,000 minutes of data (4.3 Terabytes), from which we needed citizen scientists (CS) to help collect sample clips of bird calls and songs for training deep learning bird classification models. We used the web-based Arbimon platform to facilitate this work, allowing users to sort, visualize, listen to, and identify bird vocalizations from sound recordings (Aide et al., 2013). We partnered with Sieve Analytics (the platform's creator) to design a custom C.S. interface for bird call/song reference data collection.

In the Arbimon CS interface, a volunteer with bird-call identification knowledge (i.e., "expert") delineates a bounding box representing a single bird call or song within a 1-minute spectrogram. The system then uses this template region of interest (ROI) and a pattern matching algorithm (LeBien et al. 2020) to find similar ROIs above a specified correlation threshold in our extensive collection of sound recordings. This results in hundreds to thousands of potential matches. The expert experimented with the template used and the correlation threshold to provide as many present ROI matches as possible for C.S. review while minimizing the number of clips where the bird was absent. Experts generally varied the threshold by species (minimum=0.075, maximum=0.45, average=0.29) with up to 10 matches per recording and up to 500 matches per site between 5 a.m. to 7 p.m. for diurnal species and 8 p.m. to 6 a.m. for nocturnal species.

The Arbimon CS interface allows user- and expert-level C.S. to quickly validate if matched ROIs include the bird call or song found in the template ROI. Using the template ROI as a reference, the user-level C.S. can validate a given matched ROI by assessing the visual spectrogram pattern or by listening to the sound clip, a process that needs minimal training and requires no a priori knowledge of bird calls. The system records validations of the same ROI among multiple user-level C.S. and provides a consensus vote. We chose a threshold of three user-level C.S. votes for either present or absent to reach a consensus, and matched ROIs had a maximum of five votes before a given ROI entered a validated set and was removed from further review. The expert C.S. can view statistics for all C.S. to compare their relative accuracy and make adjustments, such as improving training to reduce errors. Additionally, expert C.S. can provide sole votes on ROIs to quickly boost the number of present ROIs in the reference data,

and they can also review consensus validations to reduce false presences. This approach allowed our C.S. to develop large numbers of reference ROI data for 54 species of birds with 178,160 absences and 51,906 presences.

2.7.9. ABGQI-CNN training

The ABGQI-CNN was trained in two stages: (1) model training on the classifier only (i.e., dense, fully-connected layer), freezing all other trainable parameters for 10 epochs (learning rate = 0.0001), and (2) fine-tuning the network by freezing layers 1-50 and adjusting the final 105 layers for 10 more epochs. See GitHub code repository for code used to fine-tune the pre-trained MobileNetV2 CNN (1_fine-tune_ABGQI-CNN.ipynb). Model weights were saved only when the internal validation loss decreased.

2.7.10. Soundscape modeling, multivariate regression methods

We began with a full model that included all candidate variables, quadratic and cubic distance to road, interactions between LULC with road distance, and interaction between sound type with road distance. We also generated a model containing only soundType acting as the simplest model we explored. Below are our calls to the linear modeling command in R and a table explaining each covariate and the response. If a description includes a count (n), this represents a categorical variable. We included nonlinear terms represented with a caret (^) and interactions between two covariates represented with (*). Model fit was evaluated using MASS::stepAIC, with forward and backward selection starting with the full model. At each step, model fit was

evaluated, and the covariate with the lowest AIC, which may give less information to the model, is eliminated, and the model is refit and reevaluated iteratively.

Table 2A-5. Description of response and covariates for the linear regression modeling

Response	**Description**
logitRate	Logit value of the rate of positively predicted 2-s samples at the site. Positively predicted 2-s samples are the count of samples predicted positive. The rate of positive predicted samples is calculated as $\frac{positive}{total\ samples}$. Each site (n =746) then had five observations/rows, one for each soundType.

Covariate	**Description**
LULC	Land use land cover classes (n = 7).
logRoadDist	The natural logarithm of the distance to the nearest road from the observation.
Year	The year recordings took place (n = 4).
Log(Number_of_spectrograms)	The log of the number of total 2-s spectrograms at the site. A measure of sampling effort.
soundType	The class of sound (n = 5). Each sound was modeled as an offset in this model.
Recorder	ARU model (n = 2).
DOY	The day of initial field deployment. A measure of seasonality.

Full model (AIC = 13702.69):
lm(formula = logitRate ~ LULC + I(logRoadDist^2) + I(logRoadDist^3) + Year + log(number_mel_spectrograms) + soundType + Recorder + DOY + LULC*logRoadDist + soundType:logRoadDist)

Simple model (AIC = 13769.80):
lm(formula = logitRate ~ soundType)

Final model (AIC = 13687.75):
lm(formula = logitRate ~ LULC + Year + log(number_mel_spectrograms) + soundType + logRoadDist + soundType:logRoadDist)

2.7.11. Cross-validation F(β) threshold results

In the optimal cross-validation iteration F0.75-score model, there were 15.6% unidentified test ROIs, which decreased to 9.2% with the optimal F1-score model. However, this improvement in the F1-score model reducing unidentified samples came at the cost of 0.033 overall lower precision. Change in average precision was similar from the F0.50-score model (0.968 ± 0.032) to the F0.75-score model (0.936 ± 0.038); however, average recall was low for the F0.50-score model (0.760 ± 0.100) compared to the F0.75-score model (0.804 ± 0.055) coupled with 21.6% unidentified data in the optimal F0.50-score model. F0.75-score decreased the amount of unidentified sound relative to F0.50-score while decreasing false positives compared to the F1-score thresholds. Increasing the F(β)-score in threshold optimization resulted in lower threshold values (i.e., more liberal classification of positive predictions) and higher false-positive rates.

Table 2A-6. Cross-validation model performance for F(β) threshold values based on withheld test data for each sound class.

F(β) threshold	Precision	Recall	F0.75-score	F1-score
0.50	0.940 ± 0.02	0.698 ± 0.04	0.829 ± 0.03	0.795 ± 0.03
0.75	0.902 ± 0.03	0.754 ± 0.03	0.838 ± 0.03	0.817 ± 0.03
1.00	0.844 ± 0.04	0.815 ± 0.04	0.829 ± 0.03	0.825 ± 0.03

2.7.12. Soundscape-validation: an independent assessment of accuracy

For soundscape validation, we randomly sampled recordings from the entire S2L dataset based on a stratification determined by time and acoustic indices (Salas 2020). First, a suite of acoustic indices was calculated: Acoustic Complexity Index, Acoustic Evenness Index, Acoustic Diversity Index, Normalised Difference Soundscape Index, Acoustic Entropy Index, Temporal Entropy (Sueur et al. 2014), and Bioacoustic Index (Boelman et al. 2007). We first split recordings by time of day: dawn (6 a.m. to 9 a.m.), mid-day (11 a.m. to 2 p.m.), and late

afternoon (4 p.m. to 7 p.m.). We then applied a K-mean clustering to acoustic indices to produce clusters within each period. We chose the number of clusters by examining the drop of the within-cluster mean sum of squared errors (MSE) and selected a number that represented the point where the MSE reached an asymptotic value from which adding clusters minimally decreased the within-cluster MSE. In all three time periods, the number of clusters chosen was 6. We sampled 50 recordings per cluster, yielding 900 GV samples for review, of which 710 were reviewed.

The 710 recordings were listened to for the presence and absence of anthropophony, biophony, geophony, and interference at the 1-min level. At the review, quiet was not established as a modeling class and is omitted here. We compared reviewed recordings to 2-s spectrogram ABGQI-CNN predictions aggregated to the 1-min level (presence and absence) and generated confusion matrix classification evaluation metrics. We expect these evaluation metrics to be lower than the test dataset metrics as this test reflects the accuracy of our applied analyses while test metrics reflect model accuracy. We observe high accuracy for biophony and interference and lower accuracy for anthropophony and geophony.

Table 2A-7. Soundscape validation accuracy at the 1-min recording level.

	True Pos.	False Pos.	False Neg.	True Neg.	Precision	Recall	F1-score
Anthropophony	131	108	72	396	0.548	0.645	0.593
Biophony	437	35	101	134	0.926	0.812	0.865
Geophony	204	118	116	269	0.634	0.638	0.636
Interference	287	110	30	280	0.723	0.905	0.804

2.7.13. Summary of soundscape classes in a random sample of unidentified samples

We randomly selected 60 recordings with significant unidentified predictions (1-min recordings with greater than the median number of 2-s samples with unidentified) and included a summary of what soundscape classes author C.Q. heard during each 2-s unidentified sample (n = 832).

The most significant samples remained unidentifiable due to static, distant ambiguous low-frequency noise or masking by other sounds. Anthropophony was the next most commonly unidentified and included sounds we did not have in the ROI training set (i.e., talking, digital beeping) or were faint that may not display in the Mel spectrograms (distant traffic or plane). This latter point was also the case with biophony and geophony. A significant number of samples with biophony were extremely faint and may not have resolved in the 2-s Mel spectrograms. Other prominent issues included a persistent electronic buzz in recordings between 7-10 kHz that may have masked other sounds. Co-occurring, mixed classes appear to have been missed in some cases, though at a lower rate than the faint signals or ambiguous noises described above.

Table 2A-8. Percent of sound types that co-occur during 2-s unidentified samples from CNN

Sounds	Count	Percent
Unknown noise	129	15.5%
Anthropophony (A)	113	13.6%
Quiet (Q)	81	9.75%
Biophony (B)	76	9.15%
Geophony (G)	72	8.7%
B + I	70	8.4%
B + G	70	8.4%
Interference (I)	63	7.6%
A + B	60	7.2%

A + I	28	3.4%
A + B + I	20	2.4%
G + I	22	2.7%
B + G + I	19	2.3%
A + G	6	0.7%
A + G + I	1	0.1%
A + B + G	1	0.1%

2.7.14. Amount of sound stratified by year

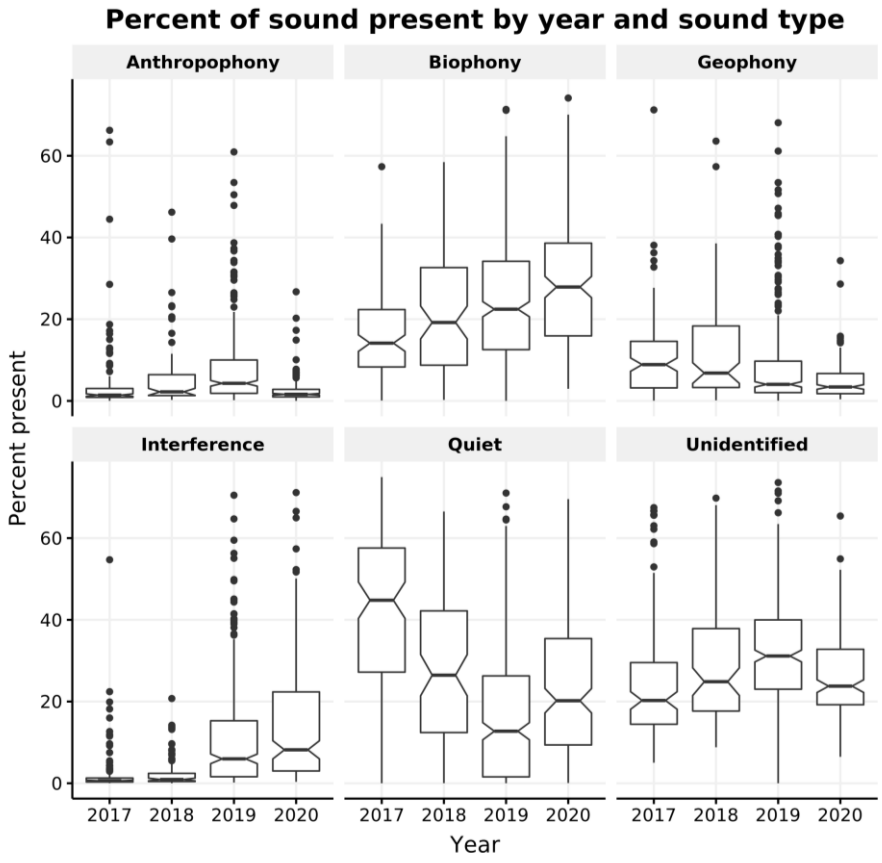


Figure 2A-3. The amount of sound for each sound type is shown below, grouped by site.

2.7.15. Pairwise test statistic summary tables

Table 2A-9. Sample counts for paired test groups—number of sites per LULC.

LULC	n
Urban/Developed	15
Agriculture/Barren	11
Herbaceous	147
Shrubland	47
Riparian/Wetland	21
Oak/Hardwood Forest	283
Conifer Forest	221

2.7.15.1. Diurnal LULC patterns

Table 2A-10. Mann-Whitney's U tests reveal significant differences in the amount of day and night sound. Significant tests in bold ($p < 0.05$).

Soundscape class	U	p-value
Anthropophony	**2.364×10^7**	**0.0000**
Biophony	**1.917×10^7**	**0.0000**
Geophony	**3.464×10^7**	**0.0000**
Quiet	**5.032×10^7**	**0.0000**
Interference	**3.161×10^7**	**0.0000**
Unidentified	**3.556×10^7**	**0.0000**

2.7.15.2. Annual and date of deployment differences

Table 2A-11. The number of sites (sample counts) for non-overlapping and overlapping deployment dates by year. Overlapping date range: May-01 to July-05

Year	Non-overlapping n	Overlapping n
2017	122	65
2018	89	46
2019	345	193
2020	190	156

Table 2A-12. Kruskal-Wallis analysis testing significant differences in the amount of sound among years when subset to overlapping deployment dates (May-01 to July-05). Significant tests in bold ($p < 0.05$).

Soundscape class	Chi-square (χ^2)	p-value
Anthropophony	**35.049**	**1.19×10^{7}**
Biophony	**29.582**	**1.69×10^{6}**
Quiet	**75.744**	**2.51×10^{16}**

Table 2A-13. Dunn analyses testing significant differences of annual sound for overlapping deployment dates (May-01 to July-05). Significant pairs in bold (p-adj < 0.05).

Soundscape class	Year pair	z value	p-adj value
Anthropophony	2017-2018	0.4824	1.0000
	2017-2019	**3.4100**	**0.003898**
	2017-2020	-0.8233	1.0000
	2018-2019	2.4139	0.09469
	2018-2020	-1.2785	1.0000
	2019-2020	**-5.6710**	**8.516×10^{8}**
Biophony	2017-2018	1.9850	0.2829
	2017-2019	**4.6692**	**1.814×10^{5}**

	2017-2020	**5.0657**	**$2.442x10^{6}$**
	2018-2019	1.7501	0.4806
	2018-2020	2.1778	0.1765
	2019-2020	0.7267	1.0000
Quiet	2017-2018	-1.0886	1.0000
	2017-2019	**-7.7693**	**$4.737x10^{14}$**
	2017-2020	**-5.3004**	**$6.932x10^{7}$**
	2018-2019	**-5.5123**	**$2.124x10^{7}$**
	2018-2020	**-3.4138**	**0.003843**
	2019-2020	**3.0806**	**0.01240**

2.7.15.3. Daytime LULC stratification

Table 2A-14. Kruskal-Wallis analysis testing significant differences of LULC daytime sound. Significant tests in bold ($p < 0.05$).

Soundscape class	Chi-square (χ^2)	p-value
Anthropophony	**97.798**	**0.0000**
Biophony	**18.891**	**0.004351**
Geophony	4.8538	0.5627
Quiet	**109.92**	**0.0000**

Table 2A-15. Dunn analyses testing significant differences of LULC daytime sound for significant Kruskal-Wallis tests. Significant pairs in bold (p-adj < 0.05).

Soundscape class	LULC pair	z value	p-adj value
Anthropophony	Urban/Developed - Agriculture/Barren	-1.0425	1.0000
	Urban/Developed - Herbaceous	**-3.1630**	**0.03279**
	Urban/Developed - Shrubland	**-4.5056**	**0.0001390**

	Urban/Developed - Riparian/Wetland	-0.9831	1.0000
	Urban/Developed - Oak/Hardwood Forest	**-4.9930**	**1.249×10^{5}**
	Urban/Developed - Conifer Forest	**-5.9314**	**6.310×10^{8}**
	Agriculture/Barren - Herbaceous	-1.4188	1.0000
	Agriculture/Barren - Shrubland	-2.7537	0.1238
	Agriculture/Barren - Riparian/Wetland	0.2190	1.0000
	Agriculture/Barren - Oak/Hardwood Forest	-2.9581	0.06501
	Agriculture/Barren - Conifer Forest	**-3.7830**	**0.003255**
	Herbaceous - Shrubland	-2.8575	0.08968
	Herbaceous - Riparian/Wetland	2.2504	0.5129
	Herbaceous - Oak/Hardwood Forest	**-4.5794**	**9.793×10^{5}**
	Herbaceous - Conifer Forest	**-6.8184**	**1.934×10^{10}**
	Shrubland - Riparian/Wetland	**3.8243**	**0.002754**
	Shrubland - Oak/Hardwood Forest	0.08411	1.0000
	Shrubland - Conifer Forest	-1.5333	1.0000
	Riparian/Wetland - Oak/Hardwood Forest	**-4.3797**	**0.0002495**
	Riparian/Wetland - Conifer Forest	**-5.4752**	**9.177×10^{7}**
	Oak/Hardwood Forest - Conifer Forest	-2.8939	0.07989
Biophony	Urban/Developed - Agriculture/Barren	-1.7355	1.0000
	Urban/Developed - Herbaceous	-1.1698	1.0000
	Urban/Developed - Shrubland	-0.8144	1.0000
	Urban/Developed - Riparian/Wetland	0.3717	1.0000
	Urban/Developed - Oak/Hardwood Forest	-0.0170	1.0000
	Urban/Developed - Conifer Forest	-1.0343	1.0000
	Agriculture/Barren - Herbaceous	1.1896	1.0000
	Agriculture/Barren - Shrubland	1.3358	1.0000

	Agriculture/Barren - Riparian/Wetland	2.1886	0.6012
	Agriculture/Barren - Oak/Hardwood Forest	2.2271	0.5447
	Agriculture/Barren - Conifer Forest	1.3370	1.0000
	Herbaceous - Shrubland	0.4510	1.0000
	Herbaceous - Riparian/Wetland	1.8978	1.0000
	Herbaceous - Oak/Hardwood Forest	**3.0745**	**0.04428**
	Herbaceous - Conifer Forest	0.3869	1.0000
	Shrubland - Riparian/Wetland	1.3988	1.0000
	Shrubland - Oak/Hardwood Forest	3.0745	1.0000
	Shrubland - Conifer Forest	-0.2145	1.0000
	Riparian/Wetland - Oak/Hardwood Forest	-0.5754	1.0000
	Riparian/Wetland - Conifer Forest	-1.7590	1.0000
	Oak/Hardwood Forest - Conifer Forest	-3.0275	0.05178
Quiet	Urban/Developed - Agriculture/Barren	0.7280	1.0000
	Urban/Developed - Herbaceous	2.0062	0.9414
	Urban/Developed - Shrubland	**4.6819**	**5.968×10^{5}**
	Urban/Developed - Riparian/Wetland	-0.4612	1.0000
	Urban/Developed - Oak/Hardwood Forest	**4.5876**	**9.414×10^{5}**
	Urban/Developed - Conifer Forest	**4.5724**	**0.0001013**
	Agriculture/Barren - Herbaceous	0.8152	1.0000
	Agriculture/Barren - Shrubland	**3.2825**	**0.02161**
	Agriculture/Barren - Riparian/Wetland	-1.1953	1.0000
	Agriculture/Barren - Oak/Hardwood Forest	3.0149	0.05398
	Agriculture/Barren - Conifer Forest	3.0135	0.05424
	Herbaceous - Shrubland	**5.0405**	**9.751×10^{6}**
	Herbaceous - Riparian/Wetland	-2.9994	0.05681

	Herbaceous - Oak/Hardwood Forest	**6.6068**	**$8.244x10^{10}$**
	Herbaceous - Conifer Forest	**6.3573**	**$4.311x10^{9}$**
	Shrubland - Riparian/Wetland	**-5.8836**	**$8.428x10^{8}$**
	Shrubland - Oak/Hardwood Forest	-1.0979	1.0000
	Shrubland - Conifer Forest	-1.0502	1.0000
	Riparian/Wetland - Oak/Hardwood Forest	**6.0636**	**$2.795x10^{8}$**
	Riparian/Wetland - Conifer Forest	**6.0258**	**$3.534x10^{8}$**
	Oak/Hardwood Forest - Conifer Forest	0.0480	1.0000

2.7.15.4. Distance to roads

Table 2A-16. Sample counts for paired test groups—the number of sites in road distance classes.

Road distance classes	n
0-99m	184
100-199m	131
200-299m	94
300-399m	68
400-499m	64
500-599m	44
600-699m	21
700-799m	25
800-899m	25
900-999m	20
> 1000m	70

Table 2A-17. Kruskal-Wallis analysis tested significant distance differences to road groups (100m) daytime sound. Significant tests in bold ($p < 0.05$).

Soundscape class	Chi-square (χ^2)	p-value
Anthropophony	**48.543**	**4.939×10^{7}**
Biophony	16.5	0.08618
Geophony	18.139	0.05267
Quiet	**46.521**	**1.154×10^{6}**

Table 2A-18. Dunn analyses testing significant differences of annual sound for significant Kruskal Wallis tests. Only significant pairs are shown and are bold (p-adj < 0.05). All values can be derived from 1_KWDunn_tests-LULC_Roads_DateofDeploy.R in the code repository.

Soundscape class	Road pair	z value	p-adj value
Anthropophony	**0-99m & 400-499m**	**-3.9048**	**0.005186**
	0-99m & >1000m	**-5.0636**	**2.263×10^{5}**
	100-199m & 400-499m	**-3.7069**	**0.01154**
	100-199m & >1000m	**-4.7939**	**8.997×10^{5}**
Quiet	**0-99m & 900-999m**	**3.7253**	**0.01073**
	0-99m & >1000m	**4.7566**	**0.0001083**
	100-199m & >1000m	**3.8575**	**0.006300**

2.7.16. Computational efficiency

The ABGQI-CNN took 4 hours to train on a laptop with an NVIDIA Quadro M1200 GPU and 16 G.B. of RAM, and approximately 100 collective hours of ROI dataset generation. Inference was a function of the S2L dataset size and took approximately 250 hours of computing time, but was calculated on multiple

computing nodes taking only 6 hours to complete. The model can be trained and deployed on a dataset similar to the S2L dataset in a single day with a small, labeled dataset. We estimate inference can be done on a single CPU with 2GB RAM at a rate of 200, 1-min recordings (each represented by 30, 2-s spectrograms) per hour.

2.7.17. ABGQI-CNN relative performance

Table 2A-19. Comparison of soundscape-component classification across select studies implementing CNN image classification or other methods. These include select studies from the environmental sound classification (ESC) field. We have used precision (Prec.) and recall (Rec.) as these metrics were most common in studies; other metrics reported when precision and recall were unavailable. nRMSE = normalized root mean square error; Acc. = n correct predictions / n incorrect predictions.

Study	Classification Method	Anthropophony	Biophony	Geophony
Quinn et al. (2022) This study.	ABGQI-CNN with transfer learning and pre-training (5 classes)	Prec.=0.975 Rec.=0.780 Acc.=0.952	Prec.=0.913 Rec.=0.840 Acc.=0.952	Prec.=0.923 Rec.=0.720 Acc.=0.932
Fairbrass et al. (2019)	Custom CNNs for Anthropo- and Biophony (2, 1 class CNNs)	Prec.=0.977 Rec.=0.858	Prec.=0.934 Rec.=0.710	NA NA
Mullet et al. (2016)	Stochastic gradient boosting regression and spatial modeling (3 classes)	nRMSE=21%	nRMSE=20%	nRMSE=20%
Salamon and Bello (2017)	Custom CNN with data augmentation (10 classes)	Acc.=0.803±0.125	NA	NA
Piczak (2015)	Custom CNN using multiple soundscape datasets and augmentation	Acc.(ESC.50)=0.645, Acc.(UrbanSound8K) =0.737	NA	NA

Inik and Seker (2020) Table 3 summarizes ESC accuracies	Grid-search of CNN parameters classifying UrbanSounds8k (10 classes)	Acc.=0.8245	NA	NA

2.8. REFERENCES

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Rafal Jozefowicz, Y.J., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Schuster, M., Monga, R., Moore, S., Murray, D., Olah, C., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vijay Vasudevan, F.V., 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.

Abdi, L., Hashemi, S., 2016. To Combat Multi-Class Imbalanced Problems by Means of Over-Sampling Techniques. IEEE Trans. Knowl. Data Eng. 28, 238–251. doi.org/10.1109/TKDE.2015.2458858

Aide TM, Corrada-Bravo C, Campos-Cerqueira M, Milan C, Vega G, Alvarez R. Real-time bioacoustics monitoring and automated species identification. PeerJ. 2013 Jul 16. doi: doi.org/10.7717/peerj.103

Aletta, F., Oberman, T., Mitchell, A., Tong, H., Kang, J., 2020. Assessing the changing urban sound environment during the COVID-19 lockdown period using short-term acoustic measurements. Noise Mapp. 7, 123–134. doi.org/10.1515/noise-2020-0011

Araya-Salas, M. and Smith-Vidaurre, G. (2017), warbleR: an r package to streamline analysis of animal acoustic signals. Methods Ecol Evol. 8, 184-191.

Balantic, C.M., Donovan, T.M., 2020. Statistical learning mitigation of false positives from template-detected data in automated acoustic wildlife monitoring. Bioacoustics 29, 296–

321. doi.org/10.1080/09524622.2019.1605309

Barber, J.R., Burdett, C.L., Reed, S.E., Warner, K.A., Formichella, C., Crooks, K.R., Theobald, D.M., Fristrup, K.M., 2011. Anthropogenic noise exposure in protected natural areas: Estimating the scale of ecological consequences. Landsc. Ecol. 26, 1281–1295. doi.org/10.1007/s10980-011-9646-7

Bedoya, C., Isaza, C., Daza, J.M., López, J.D., 2017. Automatic identification of rainfall in acoustic recordings. Ecol. Indic. 75, 95–100. doi.org/10.1016/j.ecolind.2016.12.018

Boelman, N.T., Asner, G.P., Hart, P.J., Martin, R.E., 2007. Multi-trophic invasion resistance in Hawaii: Bioacoustics, field surveys, and airborne remote sensing. Ecol. Appl. 17, 2137–2144. doi.org/10.1890/07-0004.1

Bradfer-Lawrence, T., Gardner, N., Bunnefeld, L., Bunnefeld, N., Willis, S.G., Dent, D.H., 2019. Guidelines for the use of acoustic indices in environmental research. Methods Ecol. Evol. 00, 1–12. doi.org/10.1111/2041-210X.13254

Burns, P., Clark, M., Salas, L., Hancock, S., Leland, D., Jantz, P., Dubayah, R., Goetz, S.J., 2020. Incorporating canopy structure from simulated GEDI lidar into bird species distribution models. Environ. Res. Lett. 15. doi.org/10.1088/1748-9326/ab80ee

Bush, A., Sollmann, R., Wilting, A., Bohmann, K., Cole, B., Balzter, H., Martius, C., Zlinszky, A., Calvignac-Spencer, S., Cobbold, C.A., Dawson, T.P., Emerson, B.C., Ferrier, S., Gilbert, M., Thomas, P., Herold, M., Jones, L., Leendertz, F.H., Matthews, L., Millington, J.D.A., Olson, J.R., Ovaskainen, O., Raffaelli, D., Reeve, R., Rödel, M.-O., Rodgers, T.W.,

Snape, S., Visseren-Hamakers, I., Vogler, A.P., White, P.C.L., Wooster, M.J., Yu, D.W., 2018. The Promise and Practice of Connecting Earth Observation to Biodiversity and Ecosystem Services. Nat. Ecol. Evol. doi.org/10.1038/s41559-017-0176

Buxton, R.T., McKenna, M.F., Mennitt, D., Brown, E., Fristrup, K., Crooks, K.R., Angeloni, L.M., Wittemyer, G., 2019. Anthropogenic noise in US national parks – sources and spatial extent. Front. Ecol. Environ. 559–564. doi.org/10.1002/fee.2112

Clark, M. L., & Kilham, N. E. (2016). Mapping of land cover in northern California with simulated hyperspectral satellite imagery. ISPRS Journal of Photogrammetry and Remote Sensing, 119, 228-245.

Çoban, E. B., Pir, D., So, R., & Mandel, M. I. (2020, May). Transfer Learning from Youtube Soundtracks to Tag Arctic Ecoacoustic Recordings. In ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 726-730). IEEE.

Christin, S., Hervet, É., Lecomte, N., 2019. Applications for deep learning in ecology. Methods Ecol. Evol. 10, 1632–1644. doi.org/10.1111/2041-210X.13256

Depraetere, M., Pavoine, S., Jiguet, F., Gasc, A., Duvail, S., Sueur, J., 2012. Monitoring animal diversity using acoustic indices: Implementation in a temperate woodland. Ecol. Indic. 13, 46–54. doi.org/10.1016/j.ecolind.2011.05.006

Desjonquères, C., Rybak, F., Castella, E., Llusia, D., Sueur, J., 2018. Acoustic communities reflects lateral hydrological connectivity in riverine floodplain similarly to

macroinvertebrate communities. Sci. Rep. 8, 1–11. doi.org/10.1038/s41598-018-31798-4

Doser, J.W., Hannam, K.M., Finley, A.O., 2019. Characterizing functional relationships between technophony and biophony: A western New York soundscape case study. Landsc. Ecol. 2. doi.org/10.1007/s10980-020-00973-2

Dröge, S., Martin, D.A., Andriafanomezantsoa, R., Burivalova, Z., Fulgence, T.R., Osen, K., Rakotomalala, E., Schwab, D., Wurz, A., Richter, T., Kreft, H., 2021. Listening to a changing landscape: Acoustic indices reflect bird species richness and plot-scale vegetation structure across different land-use types in north-eastern Madagascar. Ecol. Indic. 120. doi.org/10.1016/j.ecolind.2020.106929

Duchac, L.S., Lesmeister, D.B., Dugger, K.M., Ruff, Z.J., Davis, R.J., 2020. Passive acoustic monitoring effectively detects Northern Spotted Owls and Barred Owls over a range of forest conditions. Condor 122, 1–22. doi.org/10.1093/condor/duaa017

Dumyahn, S.L., Pijanowski, B.C., 2011. Soundscape conservation. Landsc. Ecol. 26, 1327–1344. doi.org/10.1007/s10980-011-9635-x

Eldridge, A., Guyot, P., Moscoso, P., Johnston, A., Eyre-Walker, Y., Peck, M., 2018. Sounding out ecoacoustic metrics: Avian species richness is predicted by acoustic indices in temperate but not tropical habitats. Ecol. Indic. 95, 939–952. doi.org/10.1016/j.ecolind.2018.06.012

Fairbrass, A.J., Firman, M., Williams, C., Brostow, G.J., Titheridge, H., Jones, K.E., 2019. CityNet—Deep learning tools for urban ecoacoustic assessment. Methods Ecol. Evol. 10, 186–197. doi.org/10.1111/2041-210X.13114

Fairbrass, A.J., Rennett, P., Williams, C., Titheridge, H., Jones, K.E., 2017. Biases of acoustic indices measuring biodiversity in urban areas. Ecol. Indic. 83, 169–177. doi.org/10.1016/j.ecolind.2017.07.064

Farina, A., Pieretti, N., 2014. Sonic environment and vegetation structure: A methodological approach for a soundscape analysis of a Mediterranean maqui. Ecol. Inform. 21, 120–132. doi.org/10.1016/j.ecoinf.2013.10.008

Ferrell, R.M., Comendant, T., Micheli, E., Dodge, C., Stern, M., Flint, L., Flint, A., Neville., J.A., 2021a. Pepperwood Long-Term Soil and MET Data - Oak and Grass Stations. Environmental Data Initiative. Retrieved from pasta.lternet.edu/package/eml/edi/943/1

Ferrell, R.M., Comendant, T., Micheli, E., Neville, J.A., 2021b. Pepperwood MET soil moisture sites 2019 - 2021. Environmental Data Initiative. Retrieved from pasta.lternet.edu/package/eml/edi/865/1

Francis, C.D., Barber, J.R., 2013. A framework for understanding noise impacts on wildlife: An urgent conservation priority. Front. Ecol. Environ. 11, 305–313. doi.org/10.1890/120183

Francis, C.D., Newman, P., Taff, B.D., White, C., Monz, C.A., Levenhagen, M., Petrelli, A.R., Abbott, L.C., Newton, J., Burson, S., Cooper, C.B., Fristrup, K.M., McClure, C.J.W., Mennitt, D., Giamellaro, M., Barber, J.R., 2017. Acoustic environments matter: Synergistic benefits to humans and ecological communities. J. Environ. Manage. 203, 245–254. doi.org/10.1016/j.jenvman.2017.07.041

Furumo, P.R., Aide, M.T., 2019. Using soundscapes to assess biodiversity in Neotropical oil

palm landscapes. Landsc. Ecol. 34, 911–923. doi.org/10.1007/s10980-019-00815-w

Gage, S.H., Joo, W., Kasten, E.P., Fox, J., Biswas, S., 2015. Acoustic observations in agricultural landscapes. Ecol. Agric. landscapes long-term Res. path to Sustain. 360–377.

Gasc, A., Gottesman, B.L., Francomano, D., Jung, J., Durham, M., Mateljak, J., Pijanowski, B.C., 2018. Soundscapes reveal disturbance impacts: biophonic response to wildfire in the Sonoran Desert Sky Islands. Landsc. Ecol. 33, 1399–1415. doi.org/10.1007/s10980-018-0675-3

Gordon, T.A.C., Harding, H.R., Wong, K.E., Merchant, N.D., Meekan, M.G., McCormick, M.I., Radford, A.N., Simpson, S.D., 2018. Habitat degradation negatively affects auditory settlement behavior of coral reef fishes. Proc. Natl. Acad. Sci. U. S. A. 115, 5193–5198. doi.org/10.1073/pnas.1719291115

Grant, P.B.C., Samways, M.J., 2016. Use of ecoacoustics to determine biodiversity patterns across ecological gradients. Conserv. Biol. 30, 1320–1329. doi.org/10.1111/cobi.12748

Hill, A.P., Prince, P., Piña Covarrubias, E., Doncaster, C.P., Snaddon, J.L., Rogers, A., 2018. AudioMoth: Evaluation of a smart open acoustic device for monitoring biodiversity and the environment. Methods Ecol. Evol. 9, 1199–1211. doi.org/10.1111/2041-210X.12955

Holgate, B., Maggini, R., Fuller, S., 2021. Mapping ecoacoustic hot spots and moments of biodiversity to inform conservation and urban planning. Ecol. Indic. 126. doi.org/10.1016/j.ecolind.2021.107627

Joo, W. Gage, S.H., Kasten E.P., 2011. Analysis and interpretation of variability in soundscapes

along an urban-rural gradient. Landsc. Urban Plan. 103, 259-276. doi.org/10.1016/j.landurbplan.2011.08.001

Kahl, S., 2020. Identifying Birds by Sound: Large-scale Acoustic Event Recognition for Avian Activity Monitoring. Technische Universitat Chemnitz.

Kahl, S., Wilhelm-Stein, T., Klinck, H., Kowerko, D., Eibl, M., 2018. Recognizing Birds from Sound - The 2018 BirdCLEF Baseline System. arXiv preprint arXiv:1804.07177

Kahl, S., Wood, C.M., Eibl, M., Klinck, H., 2021. BirdNET: A deep learning solution for avian diversity monitoring. Ecol. Inform. 61, 101236. doi.org/10.1016/j.ecoinf.2021.101236

Kasten, E.P., Gage, S.H., Fox, J., Joo, W., 2012. The remote environmental assessment laboratory's acoustic library: An archive for studying soundscape ecology. Ecol. Inform. 12, 50–67. doi.org/10.1016/j.ecoinf.2012.08.001

Knight, E.C., Hannah, K.C., Foley, G.J., Scott, C.D., Brigham, R.M., Bayne, E., 2017. Recommendations for acoustic recognizer performance assessment with application to five common automated signal recognition programs. Avian Conserv. Ecol. 12, art14. doi.org/10.5751/ACE-01114-120214

Krause, B., 2002. The Loss of Natural Soundscapes. Earth Isl. J. 17, 27–29.

Krause, B., Farina, A., 2016. Using ecoacoustic methods to survey the impacts of climate change on biodiversity. Biol. Conserv. 195, 245–254. doi.org/10.1016/j.biocon.2016.01.013

Lasseck, M., 2019. Bird species identification in soundscapes. CEUR Workshop Proc. 2380.

LAStools, "Efficient LiDAR Processing Software", obtained from

http://rapidlasso.com/LAStools

LeBien, J., Zhong, M., Campos-Cerqueira, M., Velev, J.P., Dodhia, R., Ferres, J.L., Aide, T.M., 2020. A pipeline for identification of bird and frog species in tropical soundscape recordings using a convolutional neural network. Ecol. Inform. 59, 101113. doi.org/10.1016/j.ecoinf.2020.101113

Lecun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436–444. doi.org/10.1038/nature14539

Lellouch, L., Pavoine, S., Jiguet, F., Glotin, H., Sueur, J., 2014. Monitoring temporal change of bird communities with dissimilarity acoustic indices. Methods Ecol. Evol. 5, 495–505. doi.org/10.1111/2041-210X.12178

Lie, A., Skogstad, M., Johannessen, H.A., Tynes, T., 2016. Occupational noise exposure and hearing : a systematic review. Int. Arch. Occup. Environ. Health 89, 351–372. doi.org/10.1007/s00420-015-1083-5

Lin, T.H., Tsao, Y., 2020. Source separation in ecoacoustics: a roadmap towards versatile soundscape information retrieval. Remote Sens. Ecol. Conserv. 6, 236–247. doi.org/10.1002/rse2.141

Lyon, R.H., (1973). Propagation of Environmental Noise: More theoretical and experimental work could permit the prediction and subsequent control of environmental noise. Science, 179(4078), 1083-1090.

MacLaren, A.R., McCracken, S.F., Forstner, M.R.J., 2018. Development and Validation of

Automated Detection Tools for Vocalizations of Rare and Endangered Anurans 9, 144–154. doi.org/10.3996/052017-JFWM-047

McFee, B., Raffel, C., Liang, D., Ellis, D.P.W., McVicar, M., Battenberg, E., Nieto, O., 2015. librosa: Audio and music signal analysis in python. In Proceedings of the 14th python in science conference, 18-25.

Metcalf, O.C., Barlow, J., Devenish, C., Marsden, S., Berenguer, E., Lees, A.C., 2021. Acoustic indices perform better when applied at ecologically meaningful time and frequency scales. Methods Ecol. Evol. 12, 421–431. doi.org/10.1111/2041-210X.13521

Metcalf, O.C., Lees, A.C., Barlow, J., Marsden, S.J., Devenish, C., 2020. hardRain: An R package for quick, automated rainfall detection in ecoacoustic datasets using a threshold-based approach. Ecol. Indic. 109, 105793. doi.org/10.1016/j.ecolind.2019.105793

Mohammed, R., Rawashdeh, J., Abdullah, M., 2020. Machine Learning with Oversampling and Undersampling Techniques: Overview Study and Experimental Results. 2020 11th Int. Conf. Inf. Commun. Syst. ICICS 2020 243–248. doi.org/10.1109/ICICS49469.2020.239556

Mullet, T.C., Farina, A., Gage, S.H., 2017a. The Acoustic Habitat Hypothesis: An Ecoacoustics Perspective on Species Habitat Selection. Biosemiotics 10, 319–336. doi.org/10.1007/s12304-017-9288-5

Mullet, T.C., Gage, S.H., Morton, J.M., Huettmann, F., 2016. Temporal and spatial variation of a winter soundscape in south-central Alaska. Landsc. Ecol. 31, 1117–1137. doi.org/10.1007/s10980-015-0323-0

Mullet, T.C., Morton, J.M., Gage, S.H., Huettmann, F., 2017b. Acoustic Footprint of Snowmobile Noise and Natural Quiet Refugia in an Alaskan Wilderness. Nat. Areas J. 37, 332–349. doi.org/10.3375/043.037.0308

Naguib, M., Riebel, K., 2014. Singing in space and time: the biology of birdsong. doi.org/10.1007/978-94-007-7414-8

Newport, J., Shorthouse, D.J., Manning, A.D., 2014. The effects of light and noise from urban development on biodiversity: Implications for protected areas in Australia. Ecol. Manag. Restor. 15, 204–214. doi.org/10.1111/emr.12120

Pavan, G., 2017. Fundamentals of Soundscape Conservation. Ecoacoustics Ecol. Role Sounds 235–258. doi.org/10.1002/9781119230724.ch14

Pérez-Granados, C., Traba, J., 2021. Estimating bird density using passive acoustic monitoring: a review of methods and suggestions for further research. Ibis (Lond. 1859). 1–19. doi.org/10.1111/ibi.12944

Piczak, K.J., 2015. Environmental sound classification with convolutional neural networks. IEEE Int. Work. Mach. Learn. Signal Process. MLSP 2015-November. doi.org/10.1109/MLSP.2015.7324337

Pieretti, N., Farina, A., 2013. Application of a recently introduced index for acoustic complexity to an avian soundscape with traffic noise. J. Acoust. Soc. Am. 134, 891–900. doi.org/10.1121/1.4807812

Pieretti, N., Farina, A., Morri, D., 2011. A new methodology to infer the singing activity of an

avian community: The Acoustic Complexity Index (ACI). Ecol. Indic. 11, 868–873. doi.org/10.1016/j.ecolind.2010.11.005

Pijanowski, B.C., Farina, A., Gage, S.H., Dumyahn, S.L., Krause, B.L., 2011. What is soundscape ecology? An introduction and overview of an emerging new science. Landsc. Ecol. 26, 1213–1232. doi.org/10.1007/s10980-011-9600-8

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. Nat. Commun. 11, 1–11. doi.org/10.1038/s41467-020-18321-y

Python Software Foundation. (2016). Python Language Reference. Retrieved from http://www.python.org

Rappaport, D.I., Royle, J.A., Morton, D.C., 2020. Acoustic space occupancy: Combining ecoacoustics and lidar to model biodiversity variation and detection bias across heterogeneous landscapes. Ecol. Indic. 113, 106172. doi.org/10.1016/j.ecolind.2020.106172

R Core Team, 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from www.R-project.org/.

Rice, W.L., Newman, P., Miller, Z.D., Taff, B.D., 2020. Protected areas and noise abatement: A spatial approach. Landsc. Urban Plan. 194, 103701. doi.org/10.1016/j.landurbplan.2019.103701

Rose, S.J., Allen, D., Noble, D., Clarke, J.A., 2018. Quantitative analysis of vocalizations of

captive Sumatran tigers (Panthera tigris sumatrae). Bioacoustics 27, 13–26. doi.org/10.1080/09524622.2016.1272003

Ruff, Z.J., Lesmeister, D.B., Appel, C.L., Sullivan, C.M., 2021. Workflow and convolutional neural network for automated identification of animal sounds. Ecol. Indic. 124, 107419. doi.org/10.1016/j.ecolind.2021.107419

Salamon, J., Bello, J.P., 2017. Deep Convolutional Neural Networks and Data Augmentation for Environmental Sound Classification. IEEE Signal Process. Lett. 24, 279–283. doi.org/10.1109/LSP.2017.2657381

Salamon, J.J., Bello, J.P., Farnsworth, A., Kelling, S., 2017. Fusing shallow and deep learning for bioacoustic bird species classification, in: 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). pp. 141–145.

Salas, L., 2020. soundscapeIndicesOrdination.R. Available at *GitHub Repository*, github.com/pointblue/s2lcitsci

Sánchez-Giraldo, C., Bedoya, C.L., Morán-Vásquez, R.A., Isaza, C. V., Daza, J.M., 2020. Ecoacoustics in the rain: understanding acoustic indices under the most common geophonic source in tropical rainforests. Remote Sens. Ecol. Conserv. 6, 248–261. doi.org/10.1002/rse2.162

Scarpelli, M.D.A., Liquet, B., Tucker, D., Fuller, S., Roe, P., 2021. Multi-Index Ecoacoustics Analysis for Terrestrial Soundscapes: A New Semi-Automated Approach Using Time-Series Motif Discovery and Random Forest Classification. Front. Ecol. Evol. 9, 1–14.

doi.org/10.3389/fevo.2021.738537

Schafer, R.M., 1993. The soundscape: our sonic environment and the tuning of the world. Simon and Schuster.

Sethi, S.S., Jones, N.S., Fulcher, B.D., Picinali, L., Clink, D.J., Klinck, H., Orme, C.D.L., Wrege, P.H., Ewers, R.M., 2020. Characterizing soundscapes across diverse ecosystems using a universal acoustic feature set. Proc. Natl. Acad. Sci. U. S. A. doi.org/10.1073/pnas.2004702117

Shaw, T., Hedes, R., Sandstrom, A., Ruete, A., Hiron, M., Hedblom, M., Eggers, S., Mikusiński, G., 2021. Hybrid bioacoustic and ecoacoustic analyses provide new links between bird assemblages and habitat quality in a winter boreal forest. Environ. Sustain. Indic. 11. doi.org/10.1016/j.indic.2021.100141

Shiu, Y., Palmer, K.J., Roch, M.A., Fleishman, E., Liu, X., Nosal, E., 2020. Use of deep neural networks for automated detection of marine mammal species 1–29. doi.org/10.1038/s41598-020-57549-y

Shonfield, J., Bayne, E.M., 2017. Autonomous recording units in avian ecological research: current use and future applications. Avian Conserv. Ecol. 12. doi.org/10.5751/ace-00974-120114

Slabbekoorn, H., Ripmeester, E. A. P., 2008. Birdsong and anthropogenic noise: implications and applications for conservation. Mol. Ecol., 17, 72-83. doi.org/10.1111/j.1365-294X.2007.03487.x

Southworth, M., 1969. The Sonic Environment of Cities. Environment Behav. 1, 22.

Sueur, J., Farina, A., Gasc, A., Pieretti, N., Pavoine, S., 2014. Acoustic indices for biodiversity assessment and landscape investigation. Acta Acust. united with Acust. 100, 772–781. doi.org/10.3813/AAA.918757

Sueur, J., Pavoine, S., Hamerlynck, O., Duvail, S., 2008. Rapid acoustic survey for biodiversity appraisal. PLoS One 3, 1–10. doi.org/10.1371/journal.pone.0004065

Towsey, M., 2013. Noise removal from waveforms and spectrograms derived from natural recordings of the environment.

Venables, W. N. & Ripley, B. D., 2002. Modern Applied Statistics with S. Fourth Edition. Springer, New York. ISBN 0-387-95457-0

Villanueva-Rivera, L.J., Pijanowski, B.C., Doucette, J., Pekin, B., 2011. A primer of acoustic analysis for landscape ecologists. Landsc. Ecol. 26, 1233–1246. doi.org/10.1007/s10980-011-9636-9

Ware, H.E., McClure, C.J.W., Carlisle, J.D., Barber, J.R., Daily, G.C., 2015. A phantom road experiment reveals traffic noise is an invisible source of habitat degradation. Proc. Natl. Acad. Sci. U. S. A. 112, 12105–12109. doi.org/10.1073/pnas.1504710112

Wearn, O.R., Freeman, R., Jacoby, D.M.P., 2019. Responsible AI for conservation. Nat. Mach. Intell. 1, 72–73. doi.org/10.1038/s42256-019-0022-7

Wiley, R.H., Richards, D.G., 1978. Physical Constraints on Acoustic Communication in the Atmosphere : Implications for the Evolution of Animal Vocalizations. Behav. Ecol.

Sociobiol. 3, 69–94.

Yip, D.A., Bayne, E.M., Sólymos, P., Campbell, J., Proppe, D., 2017. Sound attenuation in forest and roadside environments: Implications for avian point-count surveys. Condor 119, 73–84. doi.org/10.1650/CONDOR-16-93.1

Yosinski, J., Clune, J., Bengio, Y., Lipson, H., 2014. How transferable are features in deep neural networks? Adv. Neural Inf. Process. Syst. 4, 3320–3328. arXiv preprint arXiv:1411.1792

Zhong, M., LeBien, J., Campos-Cerqueira, M., Dodhia, R., Lavista Ferres, J., Velev, J.P., Aide, T.M., 2020. Multispecies bioacoustic classification using transfer learning of deep convolutional neural networks with pseudo-labeling. Appl. Acoust. 166, 107375. doi.org/10.1016/j.apacoust.2020.107375

CHAPTER 3: SOUNDSCAPE COMPONENTS INFORM ACOUSTIC INDEX PATTERNS AND REFINE ESTIMATES OF BIRD SPECIES RICHNESS

3.1. ABSTRACT

Ecoacoustic monitoring has proliferated as autonomous recording units (ARU) have become more accessible. ARUs provide a non-invasive, passive method to assess ecosystem dynamics related to vocalizing animal behavior and human activity. With the ever-increasing volume of acoustic data, the field has grappled with summarizing ecologically meaningful patterns in recordings. Almost 70 acoustic indices have been developed that offer summarized measurements of bioacoustic activity and ecosystem conditions. However, their systematic relationships to ecologically meaningful patterns in varying sonic conditions are inconsistent and lead to non-trivial interpretations. We used an acoustic dataset of over 725,000 min of recordings across 1,195 sites in Sonoma County, California, to evaluate the relationship between 15 established acoustic indices and sonic conditions summarized using five soundscape components classified using a convolutional neural network: anthropophony (anthropogenic sounds), biophony (biotic sounds), geophony (wind and rain), quiet (lack of emergent sound), and interference (ARU feedback). We used generalized additive models to assess acoustic indices and biophony as ecoacoustic indicators of avian diversity. Models that included soundscape components explained acoustic indices with varying degrees of performance (avg. adj-R^2 = 0.61 ± 0.16; n = 1,195). For example, we found the normalized difference soundscape index was the most sensitive index to biophony while being less influenced by ambient sound. However, all indices were affected by non-biotic sound sources to varying degrees. We found that biophony and acoustic indices combined were highly predictive in modeling bird species richness

(deviance = 65.8%; RMSE = 3.9 species; n = 1,185 sites) for targeted, morning-only recording periods. Our analyses demonstrate the confounding effects of non-biotic soundscape components on acoustic indices, and we recommend that applications be based on anticipated sonic environments. For instance, in the presence of extensive rain and wind, we suggest using an index minimally affected by geophony. Furthermore, we provide evidence that a measure of biodiversity (bird species richness) is related to the aggregate biotic acoustic activity (biophony). This established relationship adds to recent work that identifies biophony as a reliable and generalizable ecoacoustic measure of biodiversity.

3.2. INTRODUCTION

Monitoring animal diversity enables us to understand how species and communities change across time and space in relation to natural and anthropogenic forces, such as climate change (Magurran et al., 2010). Ecoacoustic monitoring provides cost- and time-effective methods to quantify ecosystem changes (Pijanowski et al., 2011; Sueur et al., 2014). Recording of environmental sounds across landscapes can capture animal community dynamics by tracking native and invasive vocalizing species (Wood et al., 2019), as well as the multiple drivers responsible for biodiversity loss, such as agricultural expansion (Dröge et al., 2021) and logging (Burivalova et al., 2019; Rappaport et al., 2022). Studies relating biodiversity to acoustic patterns are promising but require further exploration for operational monitoring and conservation efforts (Sueur and Farina, 2015; Gibb et al., 2018; Burivalova et al., 2019).

Ecoacoustics leverages differences in sound emanating from a landscape, i.e., a soundscape (Pijanowski et al., 2011), to infer patterns and changes in animal and human communities. Over the past 15 years, ecoacoustic studies have focused on establishing methods that attempt to distill

ecologically-meaningful information from recordings (Sueur and Farina, 2015). These efforts have resulted in the creation of approximately 70 acoustic indices (Buxton et al., 2018b) that summarize the time and frequency domain of acoustic recordings to provide simple metrics of acoustic activity (Sueur et al., 2008; Pijanowski et al., 2011; Gasc et al., 2015) and biodiversity-related patterns (Sueur et al., 2008; Ross et al., 2021).

Although such indices allow for more convenient and efficient summarization of acoustic data than the expert knowledge and time required for identifying individual sound events (e.g., bird species; Snyder et al., 2022), explicit links among acoustic indices (BradferLawrence et al., 2020) and biotic acoustic activity ("biophony") or biodiversity are limited (Duarte et al., 2021), and empirical relationships between biophony and established biodiversity indicators remain equivocal. These gaps persist even though biophony is commonly conceptualized in ecoacoustics as a proxy for biodiversity and is treated as the intrinsic soundscape property that drives acoustic index patterns (Pijanowski et al., 2011). This lack of explicit, foundational association between biophony and acoustic indices has led to disparate findings. For example, in one study, acoustic indices have been found to reflect avian species richness in temperate habitats but were weaker in tropical ones (Eldridge et al., 2018), while another tropical study found similar indices were related to bird abundance and biological activity (Retamosa Izaguirre et al., 2018). Other studies have demonstrated that acoustic indices have no relationship to avian species richness (Moreno-Gómez et al., 2019).

Inconsistencies in relationships among acoustic indices and biodiversity measures may be partly due to non-biotic sounds mixing with biotic sounds in recordings; however, these confounding issues are rarely accounted for (Fairbrass et al., 2017; Duarte et al., 2021) and have

only recently been systematically investigated (e.g., Ross et al., 2021). A meta-analysis of acoustic index applications revealed that although acoustic indices were significantly related to biological activity in 74% of studies, they were also related to anthropogenic activity in 88% of studies (Buxton et al., 2018b). The influence of anthropogenic acoustic activity (i.e., "anthropophony"; Fairbrass et al., 2017), ambient weather sounds (i.e., "geophony"; Depraetere et al., 2012; Sánchez-Giraldo et al., 2020), and ambient noise (Gasc et al., 2015) present further difficulty in applying acoustic indices as ecosystem monitoring tools.

The efficacy of acoustic indices and their predictable interpretation across study domains remains unclear, resulting in reduced generalizability. Gibb et al. (2018) discuss current issues in "accuracy, transferability, and limitations of many [ecoacoustic] analytical methods" and further highlight how comparison of acoustic index values becomes non-trivial across study sites and surveys, particularly outside of undisturbed forest regions (i.e., in the presence of variable sonic conditions). These issues are particularly vexing, as affordable autonomous recording units (ARU) have led to increasingly large datasets requiring automated methods for interpretation (Snyder et al., 2022). Indices should be interpretable across new study domains to establish spatial and temporal consistency for tracking change (Bradfer-Lawrence et al., 2020). We believe a significant factor resulting in the lack of transferable interpretation is the effect of non-biotic sounds on index values, which must be better understood and accounted for to facilitate interpretable transferability.

Here we categorize soundscapes into acoustic activity events (Krause, 2002; Pijanowski et al., 2011), which we call soundscape components. Soundscape components include anthropophony (anthropogenic sounds), biophony (animal sounds), geophony (wind and rain),

quiet (lack of emergent sound), and interference (ARU feedback), collectively ABGQI (Quinn et al., 2022). We use these terms to reflect the presence of the events they represent, such as a birdcall, amphibian chorus, engine noise, or wind gust at a 2-s temporal resolution which we then aggregate to percent time present as opposed to abstract representations of acoustic activity in frequency bands or acoustic index values. Grinfeder et al. (2022) developed a model that classified anthropophony, biophony, geophony, and quiet that demonstrated the ability of soundscape components to track soundscape trends over long periods. Additionally, two studies used manual (Mullet et al., 2016) and supervised, deep learning (Fairbrass et al., 2019) methods to classify biophony in large acoustic datasets, with the latter method outperforming the ability to capture biophony when compared to a suite of acoustic indices (Fairbrass et al., 2019). These studies provide evidence that soundscape components may provide a more transferable and flexible method for assessing levels of biodiversity represented by biophony and human impact across landscapes, adding to the potential for soundscape components to be used for ecosystem monitoring.

We leverage a 5-year dataset from low-cost ARUs with over 725,000 min of acoustic recordings across 1,195 sites spanning urban to natural habitats in Sonoma County, California, United States. Our objectives here are to 1) understand how multiple acoustic indices are modeled by the amount of ABGQI events at a site level, 2) explore the relationship between biophony and bird diversity to determine if biophony can be used as an indicator of bird diversity, and 3) offer a solution to the confounding effects of non-biotic soundscape components on acoustic indices. We build upon studies that have related acoustic indices to soundscape components (Fairbrass et al., 2017), measures of biodiversity (Gasc et al., 2015), and compare

patterns here with a study that analyzed the effects of sonic conditions on index interpretation (Ross et al., 2021).

3.3. METHODS

3.3.1. Study region and data collection

As part of the Soundscapes to Landscapes project (Snyder et al., 2022; soundscapes2landscapes.org, hereafter S2L), citizen scientists deployed passive ARUs across Sonoma County, California, United States (4,152 km^2), overlapping with the bird breeding season from late March to June, through years 2017–2021 (Figure 3-1). At each site, ARUs recorded 1 min every 10 min, amounting to 144 min of acoustic recordings per 24 h. S2L employed two different ARU models: AudioMoths (AM) programmed with an upper-frequency range of 24 kHz (sampling rate = 48,000 Hz, digitization depth = 16-bit; flat frequency response ± 2 dB between 100 and 10,000 Hz; sensitivity −18 dB V/ Pa re: 94 dB SPL @ 1 kHz; Hill et al., 2018) and LG phone recorders with an upper-frequency range of 22.05 kHz (sampling rate = 44,100 Hz, digitization depth = 16-bit; flat frequency response between 50 and 20,000 Hz; sensitivity −45 ± 2 dB V/Pa re: 94 dB SPL @ 1 kHz; Campos-Cerqueira and Aide, 2016). We established an approximate minimum distance of 500 m between ARUs deployed at the same time, though 22 intentionally paired AMLG deployments were included in analyses from 2021 (i.e., distance = 0 m). AMs were covered in a protective vinyl pouch, while LG recorders were housed in a hard shell with an external microphone. In both cases, ARUs were affixed to woody vegetation if present or a temporary stake at approximately 1.5 m. We used 1,195 unique sites and 726,801 min of acoustic data for analyses (Table 3-1).

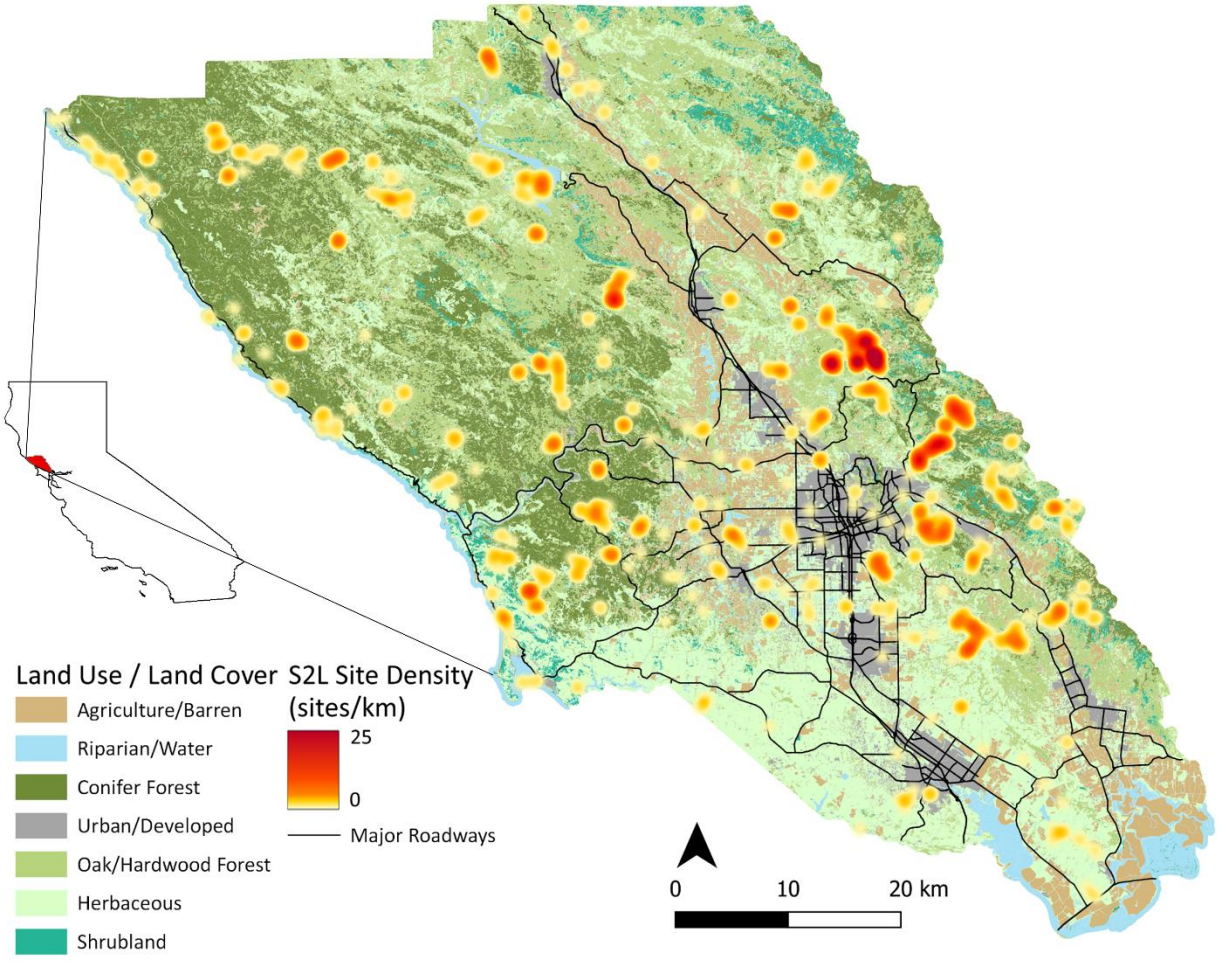


Figure 3-1. Sonoma County, California, USA land use/land cover classes from Sonoma County Fine-scale Vegetation and Habitat Map (sonomavegmap.org) used as a component of site selection stratification from 2017 to 2021 (n = 1,195).

Table 3-1. Summary of acoustic sampling by year and device.

	LG ARU		AM ARU				
Year	**Sites**	**Minutes**	**Sites**	**Minutes**	**Percent AM**	**Total Sites**	**Total Minutes**
2017	119	75,643	0	0	0%	119	75,643
2018	52	38,914	35	30,650	40%	87	69,564
2019	26	16,726	305	179,883	92%	331	196,609
2020	0	0	186	133,869	100%	186	133,869
2021	24	10,672	448	240,444	95%	472	251,116
Total	221	141,955	974	584,846	82%	1,195	726,801

3.3.2. Acoustic features

3.3.2.1. Acoustic indices

We calculated 15 acoustic indices for each 1-min recording (Table 3-2 and Appendix 3A). We chose these indices because they have been frequently applied in terrestrial ecoacoustic and bioacoustic research (Sueur et al., 2014) and have consistently demonstrated varying sensitivity to sources of biophony, geophony, and anthropophony. We used minute-level values of indices to calculate two mean site-level values used in statistical analyses, one for 24-h and a second for morning only (4 a.m.–12 p.m.; used in bird species richness analyses).

Table 3-2. 15 acoustic indices and descriptors for soundscape analysis from ecoacoustic studies.

Index name	Abbreviation	Relationships between index behavior and ecoacoustic source
Acoustic Complexity Index	ACI	Strong positive correlation between the number of bird vocalizations with increasing significance when aggregated over extended recording periods (Pieretti et al., 2011).
Acoustic Diversity Index	ADI (=H')	Higher values coincide with dawn and dusk choruses. Separately, low values were observed at agricultural sites (Villanueva-Rivera et al., 2011).
Acoustic Evenness Index	AEI	Values are higher when evenness is low, reflecting sounds occurring in mostly one frequency band (perfect inequality) (Villanueva-Rivera et al., 2011).
Bioacoustic Index	BI	Values correlate with avian abundance from concurrent point counts (Boelman et al., 2007).
Temporal Entropy	H_t	The modulation of amplitude/sound over time (Sueur et al., 2008). A valuable indicator of species diversity (Towsey et al., 2014).
Spectral Entropy	H_s (also H_f)	The modulation of the frequency spectrum of sounds (Sueur et al., 2008). A valuable indicator of species diversity (Towsey et al., 2014).
Acoustic Entropy	H	Computed as the product of H_t and H_s, it increases from 0 (pure tone) to 1 (fully stochastic, random noise) with species richness. Wind, water, and human activity may decrease reliability (Sueur et al., 2008).
Median of amplitude envelope	M	Estimates the number of animal vocalizations (Depraetere et al., 2012).
Normalized Difference Soundscape Index-Anthropophony	NDSI-α	Focuses on the power spectral density between 1-2 kHz; increasing with more Anthropophony (Kasten et al., 2012).
Normalized Difference Soundscape Index - Biophony	NDSI-β	Focuses on the largest 1 kHz PSD bin between 2-8 kHz; increasing with more Biophony (Kasten et al., 2012).
Normalized Difference Soundscape Index	NDSI	The ratio of (NDSI-α - NDSI-β) / (NDSI-β + NDSI-α) in the range [-1 to +1]. A value of +1 indicates no Anthropophony, while -1 indicates no Biophony (Kasten et al., 2012).
Roughness[†]	RN	Captures temporal changes in sound over all frequency bins; developed for binaural analyses (Rychtáriková & Vermeir, 2013).
Rugosity[†]	RUGO	Continuous signals have lower values, while noisy signals have higher values (Atauri Mezquida & Llorente Martínez, 2009).

Spectral flatness†	SFM	Tends to 1 for noisy signals and 0 for perfect oscillation (Bormpoudakis et al., 2013; Mitrović et al., 2010).
Zero-Crossing Rate†	ZCR	Measures the number of times per second that a signal crosses the instantaneous pressure of 0; trends high for noisy and low for tonal sounds (Bormpoudakis et al., 2013; Eldridge et al., 2018).

† Acoustic descriptors – referred to here as acoustic indices.

3.3.2.2. Soundscape components

We used the percent presence of ABGQI soundscape events to investigate how sonic conditions in a large acoustic dataset relate to ecological indicators. We detected ABGQI events in our recordings using a convolutional neural network (CNN) classifier initially developed by and full details can be found in Quinn et al. (2022). Briefly, we generated probabilities for each ABGQI class for 21,804,030 2-s (2-s) samples (the CNN produces an independent probability for each class for each 2-s sample). The ABGQI probabilities were classified as present or absent using optimized class wise threshold values. The 2-s presences for each ABGQI class were then averaged for each site, resulting in a feature value for each soundscape component ranging from 0% to 100% present at the sitelevel. Two site-averages were calculated, one for 24-h and a second for morning only (4 a.m.–12 p.m.; used in bird species richness analyses). For example, if a site had 1,200 2-s samples from 40 1-min recordings, and anthropophony was predicted present in 300 of these 2-s samples, then 25% of recordings would be predicted to contain anthropophony. Notably, our ABGQI CNN treats each soundscape component independently, so any combination of soundscape components can be predicted for a given 2-s sample and will not necessarily sum to 100%. We interpret the effects of soundscape components as follows.

- Biophony: reflects general vocalizing animal activity. These primarily include bird calls and frog and insect choruses while capturing flying insect activity more uncommonly.

- Anthropophony: primarily reflects anthropogenic activity such as vehicle traffic, various combustion engines, and airplane activity.
- Quiet: reflects the general absence of emergent sound. We classified quiet to explicitly model periods of silence and not assume the absence of other soundscape components implied silence, as the CNN model included classification error of ABGI. A positive effect implies a relationship with more quiet periods, while a negative effect implies stronger effects from other soundscape components.
- Geophony: primarily reflects wind, as Sonoma County recordings show increased wind levels in the late afternoon (Quinn et al., 2022), and is considered non-ecologically meaningful ambient sound.
- Interference: reflects broad-frequency, rapid spikes in acoustic activity related to physical interferences with the ARU (e.g., branches hitting the ARU) or internal electronic malfunctions and is influenced by internal ARU self-noise. Interference results in higher acoustic activity that may be associated with events such as gusty winds, and we interpreted this as non-meaningful noise in recordings regardless of the cause of the interference event.

3.3.2.3. Acoustically-derived bird species richness

We used acoustically-derived bird species richness to measure biodiversity, which we then related to acoustic indices and soundscape components. The S2L project developed a separate CNN-based classification approach to identify 54 of the most common vocalizing bird species across Sonoma County (Clark et al., 2023). We derived bird species classification from three CNNs at 2-s intervals. For each bird species, we used the highest accuracy CNN to calculate the

presence and absence of the species in all recordings and applied a lower cutoff of n = 3 positive classifications per site to indicate a presence for a given site, resulting in 1,185 sites with bird observations. We then summed all species detections at each site to arrive at site-level species richness for 24 h of recording and a "morning-only" subset targeting high bird activity periods (4 a.m.–12 p.m.). For both summaries, all 54 bird species had presences across sites, with the minimum richness of zero species for both datasets and the highest richness of 51 species (μ = 27.1 ± 7.3) for 24-h and 48 species (μ = 21.7 ± 6.7) for morning-only data, respectively.

3.3.3. Statistical analyses

3.3.3.1. Generalized additive modeling

We used generalized additive models (GAMs) to relate soundscape components and acoustic indices with one another and with bird species richness. GAMs are an even more flexible modeling option than generalized linear models, which provide benefits over ordinary linear models, such as allowing for nonnormal error distributions and non-linear model structures (Wood, 2017).

All GAMs were fit using the R (R Core Team, 2022) package "mgcv" (Wood, 2017) using backward variable selection from full models. We generated partial dependence plots (PDPs) in R using the draw function from the "gratia" package (Simpson, 2022). We interpreted PDP slopes as the magnitude of influence a covariate had on the response, where higher absolute slope values were more influential for those areas in the covariate's domain. Comparatively, a covariate with a near-zero slope had minimal effect on the response when included. For our final GAMs, we reported error distribution, link function, adj-R^2, and deviance, where applicable. We report deviance because it can better approximate non-normal error distributions for GAMs;

however, deviance and adj-R^2 are equivalent for GAMs with a Gaussian error distribution. See Appendix 3A for the full GAM fitting procedure.

3.3.3.2. Acoustic index sound composition

We investigated how soundscape components vary with each acoustic index using GAMs. All models were nested versions of the full model:

$$E(Acoustic\ Index_i) = ARU_i + log(Minutes\ recorded_i) + f_1(Anthropophony_i) + f_2(Biophony_i) + f_3(Geophony_i) + f_4(Quiet_i) + f_5(Interference_i), \quad \text{(Eq. 1)}$$

where for each site observation i, f_j represents smooth functions for soundscape components, *ARU* has two levels (AM and LG, not individual device), and $Minutes\ recorded_i$ is the logarithm of the number of 1-min recordings. We scaled soundscape component covariates using a min-max normalization. We used PDPs and slope summaries from soundscape component models to interpret significant model covariates for each acoustic index. Because GAMs had splines fit with up to $k = 5$ basis functions, a partial effect could have positive and negative effects throughout the covariate domain. Positive effects implied a covariate resulted in increases in the response, while negative effects implied an inverse relationship.

3.3.3.3. Case study: accounting for Interference

GAM analyses indicate that the presence of ABGQI CNN-classified interference heavily influenced the ACI; thus, we performed an experiment to account for this effect. We removed any 1-min recording containing an interference sample and recalculated the average site soundscape components and ACI values. Removing recordings with interference reduced the dataset to 300,572 recordings (41.4% of data) and reduced sites to 1,194, though the average minutes per site were reduced from 608 ± 317 min to 251 ± 186 min, resulting in the rate of

soundscape components changing (Table 3-3). We followed the same procedure to model ACI with GAMs as found in the "Generalized additive modeling" section. The resulting final model reflects the relationship between soundscape components and ACI without the effect of interference.

Table 3-3. The change in the rate of soundscape components after removing 1-min recordings with Interference (mean and standard deviation).

Soundscape component	All recordings	No Interference recordings
Anthropophony	7.9% ± 11.5%	8.1% ± 12.8%
Biophony	29.4% ± 19.0%	30.7% ± 20.7%
Geophony	7.6% ± 10.2%	5.4% ± 9.8%
Quiet	25.9% ± 21.8%	30.2% ± 22.9%
Interference	15.9% ± 18.8%	---

3.3.3.4. Relationship between acoustic features and bird species diversity

We explored how soundscape components were associated with site-level bird species richness. First, we constructed a GAM to see how variance in Biophony could be explained by bird species richness (Eq. 2):

$$E(Biophony_i) = ARU_i + log(Minutes\ recorded_i) + f_1(Richness_i) \quad \text{(Eq. 2)}$$

where $Richness_i$ is each site's acoustically-derived bird species richness. We then designed three models to compare the ability of Biophony and acoustic indices to predict bird species richness and reflect their joint ability to function as biodiversity indicators. First, Biophony was used alone to model bird richness (Eq. 3):

$$E(Richness_i) = ARU_i + log(Minutes\ recorded_i) + f_1(Biophony_i) \quad \text{(Eq. 3)}$$

Recent research suggests multiple acoustic indices can better represent diversity measures like avian vocal diversity (Allen-ankins et al., 2023; Buxton et al., 2018a) and be used to aid

soundscape labeling of birds, insects, and geophony (Scarpelli et al., 2021) when used together. Therefore, we developed a GAM to relate all 15 acoustic indices to bird species richness to investigate (1) which acoustic indices best capture variation in richness and (2) the directional effect of significant acoustic indices (Eq. 4):

$$\begin{aligned} E(Richness_i) = {} & ARU_i + log(Minutes\,recorded_i) + f_1(ACI_i) + f_2(ADI_i) + f_3(AEI_i) \\ & + f_4(BI_i) + f_5(H_i) + f_6(H_{s_i}) + f_7(H_{t_i}) + f_8(M_i) + f_9(NDSI_i) \\ & + f_{10}(NDSI - \alpha_i) + f_{11}(NDSI - \beta_i) + f_{12}(R_i) + f_{13}(Rugosity_i) \\ & + f_{14}(SFM_i) + f_{15}(ZCR_i) \end{aligned} \quad \text{(Eq. 4)}$$

Our third GAM combined Biophony and all acoustic indices to assess the ability of all acoustic indicators to model bird species richness (Eq. 5):

$$\begin{aligned} E(Richness_i) = {} & ARU_i + log(Minutes\,recorded_i) + f_1(ACI_i) + f_2(ADI_i) + f_3(AEI_i) \\ & + f_4(BI_i) + f_5(H_i) + f_6(H_{s_i}) + f_7(H_{t_i}) + f_8(M_i) + f_9(NDSI_i) \\ & + f_{10}(NDSI - \alpha_i) + f_{11}(NDSI - \beta_i) + f_{12}(R_i) + f_{13}(Rugosity_i) \\ & + f_{14}(SFM_i) + f_{15}(ZCR_i) + f_{16}(Biophony_i) \end{aligned} \quad \text{(Eq. 5)}$$

We compared all model performances and variable selection for the morning-only and 24-h data to investigate whether targeting periods when bird species are more active results in a stronger relationship between acoustic features and bird species richness. We followed the approach outlined in the "Generalized additive modeling" section for all bird species richness models.

3.4. RESULTS

3.4.1. Effects of soundscape components on acoustic indices

We modeled acoustic indices using soundscape components (Eq. 1), with model mean and standard deviation deviance explained of 66.8% ± 13.9% (Table 3-4). Because of the relatively large sample size (n = 1,195 sites), covariate significance levels were frequently highly significant ($p < 0.05$) when soundscape components were retained in the final GAMs (Appendix

3A). All final GAMs contain biophony, interference, and the ARU fixed effect. However, no model contains the number of recordings. The fitted model for BI is the only model not to include anthropophony, quiet is not in the final model for ACI, and geophony is not included in NDSI-α, NDSI, SFM, and RN (all abbreviations defined in Table 3-2).

Figure 3-2 shows the influence on index values from the partial dependence on each sound component individually. Interference influences ACI most intensely based on all soundscape components' high positive slope values. Furthermore, all sounds positively affect the value of ACI, and quiet is not included in the final model. Comparatively, RN is weakly influenced by interference, with values clustered around a slope of zero, while anthropophony and biophony have more significant, generally positive effects on RN.

Table 3-4. Final GAM results and model structure, sorted by high to low deviance. Acoustic index abbreviations can be found in Table 3-2.

Acoustic Index	Error Distribution	Link function	Adj-R^2	Deviance
ZCR	Beta	Logit	0.78	80.4%
ACI	Gamma	Log	0.29	79.5%
M	Gaussian	Identity	0.79	79.3%
H	Beta	Logit	0.71	73.6%
H_s	Beta	Logit	0.71	71.7%
NDSI-α	Beta	Logit	0.65	71.4%
AEI	Beta	Logit	0.67	70.9%
H_t	Beta	Logit	0.66	68.9%
Rugosity	Beta	Logit	0.67	68.8%
NDSI-β	Beta	Logit	0.66	68.5%
NDSI	Beta	Logit	0.67	67.8%
ADI	Beta	Logit	0.64	66.6%
SFM	Gamma	Log	0.52	56.0%
BI	Gamma	Log	0.48	53.9%
RN	Gaussian	Identity	0.24	24.4%
Mean	--	--	**0.61**	**66.8%**

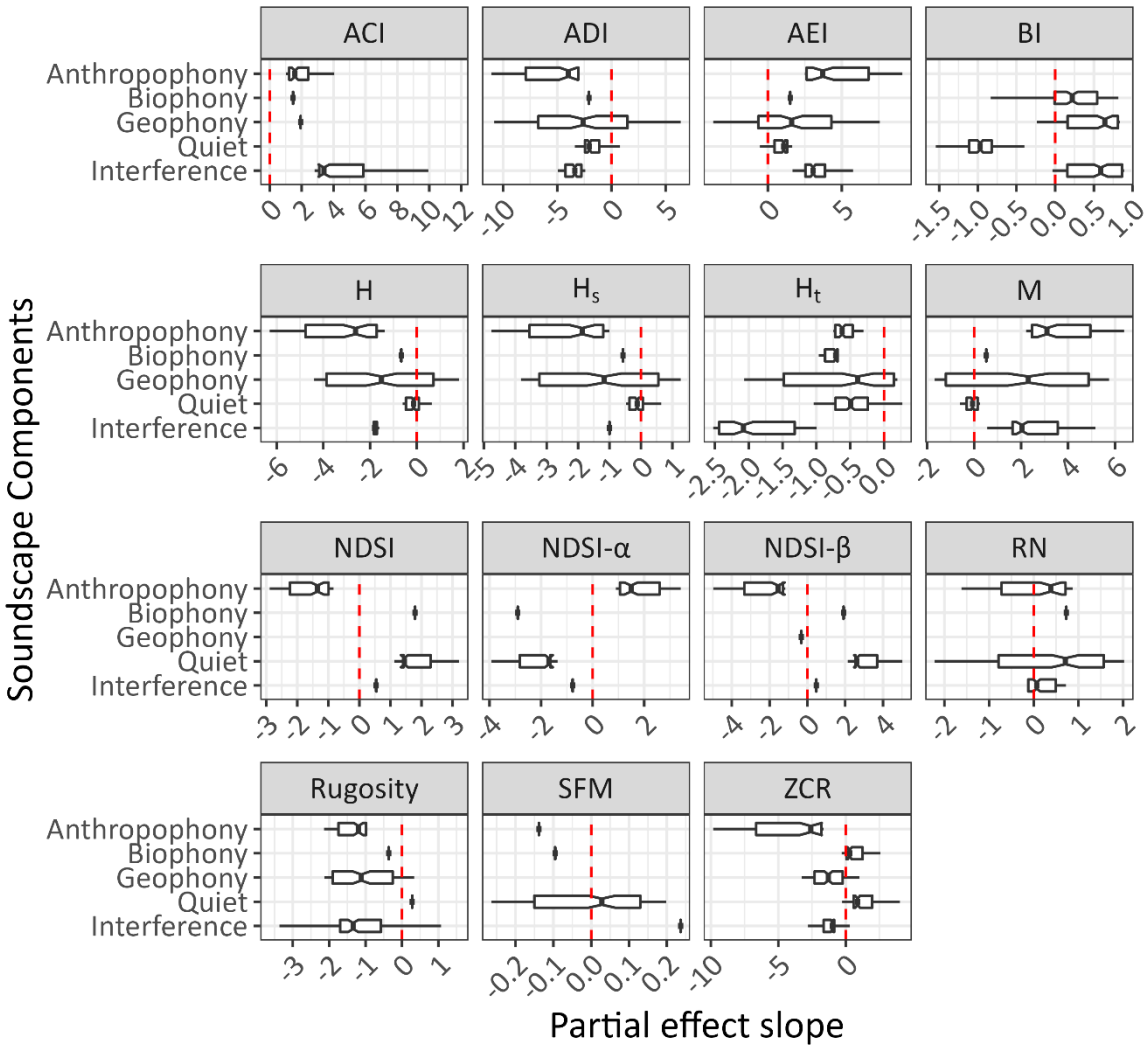


Figure 3-2. The effect of each sound component on index values in the GAM for each acoustic index, evaluated using the partial dependence of each sound component's inner 99% range of values while holding all other components to their mean value. PDPs (Appendix 3A) show the index's sensitivity to the sound component. Because these values are obtained for each component while holding the others constant, the total variance in index values for the sound component is highlighted. Higher variance in an index's partial effect slope reflects higher sensitivity to the sound component. Values of x = 0 denote a zero slope in the PDPs and do not imply non-significance.

3.4.2. Case study: accounting for ambient sounds

After removing 1-min recordings with interference, we found biophony changed from being one of the least influential covariates modeling ACI to being the most influential soundscape

component (Figure 3-3). Further, ACI became more robust to anthropophony, as indicated by lower slope values (Figure 3-4). Geophony's effect is less linear and slightly more influential, although higher slope values occur primarily in geophony's sparse higher values. Geophony's effect on ACI may have increased because it is tightly linked to non-internal ARU interference events (e.g., wind co-occurring with interference), and when these events were removed, other geophony types (e.g., rain and running water) could explain more variation in ACI. After removing interference, GAM deviance decreased slightly from 79.5% to 74.4%. Quiet is included in the final GAM, although it was non-significant ($p = 0.12$).

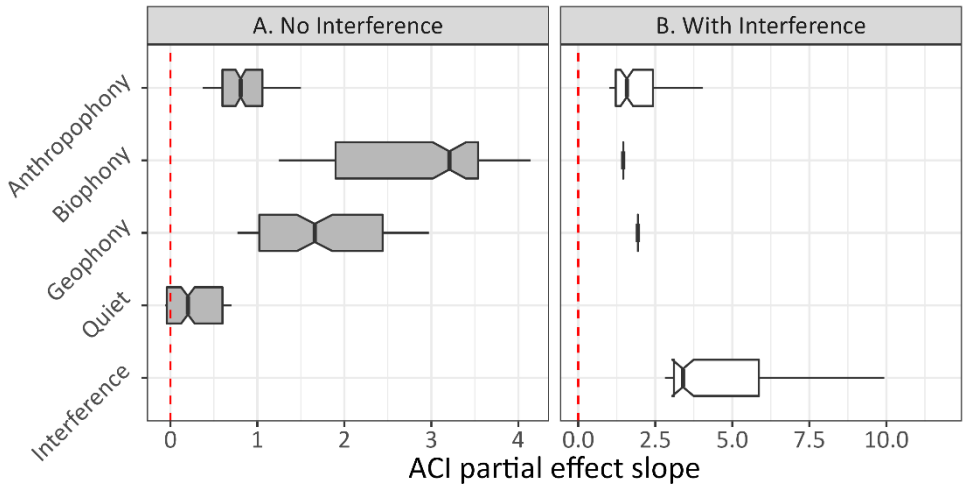


Figure 3-3. ACI GAM PDP slope summaries modeled recordings (A) without and (B) with Interference, respectively.

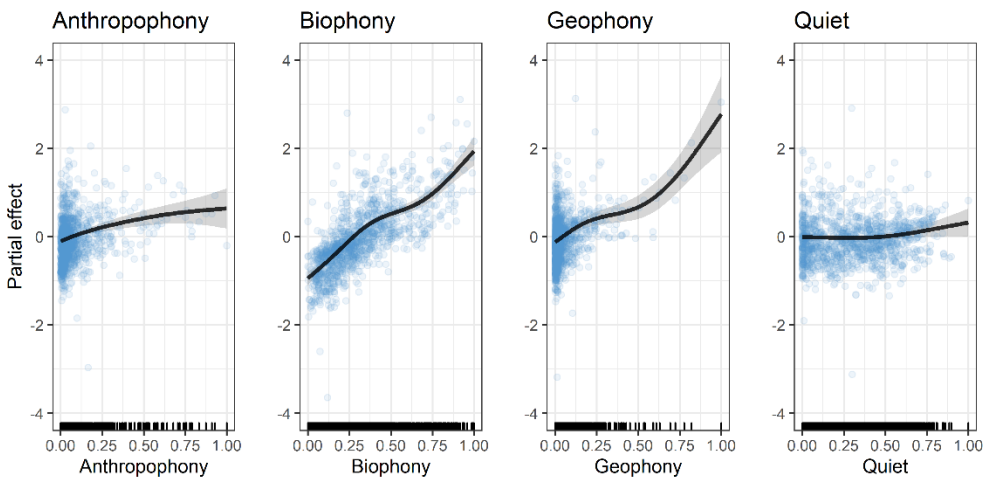


Figure 3-4. Cubic spline (black line) of PDPs of ACI modeled without Interference. The y-axis shows the mean model ACI values (centered on zero), and the x-axis shows covariate values with ticks reflecting the density of observations. Changes across a covariate's domain reflect its influence on ACI value. Shaded regions reflect 95% credible intervals of the spline, and blue points indicate partial residuals. Note that Geophony values > 0.5 are sparse yet heavily affect the behavior and precision of the spline.

3.4.3. Relationships of acoustic indices and Biophony with bird species richness

The GAM modeling biophony as a function of bird species richness (Eq. 2) had 40.3% deviance explained (i.e., a measure of model fit), 12.9% RMSE, and included both ARU and the log number of recordings for the 24-h data. The morning-only model had higher performance (deviance = 45.4%) but higher error (RMSE = 15.3%). Both datasets show a strong positive relationship between bird species richness and biophony across all species richness values. Model fits capture the mean trend in bird species richness related to biophony but underestimate the variance at intermediate species richness values (e.g., 15–35 species).

We found the highest performing GAM modeling bird species richness (according to the greatest explained deviance) was the combined acoustic indices and biophony GAM using the morning-only dataset (Table 3-5). This model resulted in the highest deviance (65.8%; RMSE =

3.9 species; Figure 3-5), 2% higher than the 24-h model. In general, the morning-only GAMs, regardless of covariates, performed better than the 24-h GAMs. Most notably, there was a 6.8% better performance for the morning-only biophony GAM (Eq. (3)) than the 24-h GAM.

The highest performing, morning-only GAM includes the log number of recordings along with biophony, ADI, AEI, BI, H, M, NDSI, NDSI-α, NDSI-β, RN, and ZCR (PDP in Appendix 3A). The number of recordings has a positive effect on bird species richness. Biophony, NDSI-β, and AEI have the largest effects of the included covariates, while ADI, BI, and RN have the smallest effects on bird species richness. AEI, BI, NDSI-β, and ZCR have concave-down shaped PDPs, but BI decreases across its entire domain. Biophony also has a concave-down PDP with the strongest positive effect from 0%–25%. NDSI-α has a concave up trend, while RN has a concave down to concave up trend across its domain, approximating a flattened sinusoid. H and M have negative linear effects, while ADI and RN have roughly linear, positive effects. Notably, NDSI has an unreliable PDP interpretation due to high concurvity and covariance structure with NDSI-α and NDSI-β. The 24-h model contained all covariates in the morningonly model in addition to ACI (Eq. 5).

Table 3-5. Performance of bird species richness models with covariates of Biophony (Eq. 3), acoustic indices (Eq. 4), and combined Biophony and acoustic indices (Eq. 5), respectively. Normalized RMSE (NRMSE) was calculated using RMSE divided by species richness range (48 for morning-only and 51 for 24-hr datasets).

	24-hr data		Morning-only data	
Model Covariates	**Deviance**	**RMSE (NRMSE)**	**Deviance**	**RMSE (NRMSE)**
Biophony (Eq. 3)	43.7%	5.4 species (10.5%)	50.5%	4.6 species (9.6%)
Acoustic Indices (Eq. 4)	50.9%	5.0 species (9.9%)	52.6%	4.5 species (9.4%)
Biophony + Acoustic Indices (Eq. 5)	63.8%	4.3 species (8.5%)	65.8%	3.9 species (8.1%)

The highest performing, morning-only GAM includes the log number of recordings along with Biophony, ADI, AEI, BI, H, M, NDSI, NDSI-α, NDSI-β, RN, and ZCR (PDP in Appendix

3A). The number of recordings has a positive effect on bird species richness. Biophony, NDSI-β, and AEI have the largest effects of the included covariates, while ADI, BI, and RN have the smallest effects on bird species richness. AEI, BI, NDSI-β, and ZCR have concave-down shaped PDPs, but BI decreases across its entire domain. Biophony also has a concave-down PDP with the strongest positive effect from 0-25%. NDSI-α has a concave up trend, while RN has a concave down to concave up trend across its domain, approximating a flattened sinusoid. H and M have negative linear effects, while ADI and RN have roughly linear, positive effects. Notably, NDSI has an unreliable PDP interpretation due to high concurvity and covariance structure with NDSI-α and NDSI-β. The 24-hr model contained all covariates in the morning-only model in addition to ACI (Eq. 5).

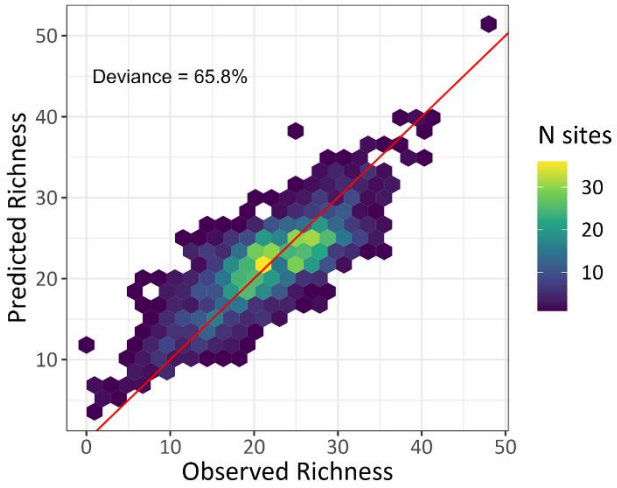


Figure 3-5. Predicted and observed bird species richness from the combined morning-only acoustic index and Biophony GAM with the log number of recordings as a fixed effect. The density of sites is shown for observed and predicted values, and the red line indicates the 1:1 fit line.

3.5. DISCUSSION

3.5.1. Contribution of soundscapes components to acoustic indices

All GAMs from Eq. 1 contained ARU effects suggesting systematic differences between ARU models. We believe this is most likely due to the hardware specification differences in the LG ARUS compared to the AM ARUs (see methods) where AM ARUs were more sensitive to distant sounds (Campos-Cerqueira and Aide, 2016) and highlights the need to include ARU type when considering sonic conditions (Haupert et al., 2022). Additionally, no acoustic index final models included the number of recordings. In our analyses, the amount of data collected did not strongly modulate the relationships between soundscape components and acoustic indices. However, Bradfer-Lawrence et al. (2019) recommend continuously sampling 120 h of data per site, or 26 weeks with the 1-min-in-10 as used here, to achieve stability in acoustic indicator variance, while our mean recording duration per site equated to 99 h of deployed recording time or 9.9 cumulative hours of recordings. We did not record for a sufficient duration to observe decreased variance at 120 h of recording, as in Bradfer-Lawrence et al. (2019), which may lead to this disparity in findings.

We found that biophony positively influenced NDSI-β, NDSI, AEI, BI, RN, ZCR, and ACI. This suggests these indices are related to vocalizing animals as these sounds were encompassed in the biophony CNN training and therefore possess ecologicallymeaningful information about animal activity and diversity. Similarly, indices with greater influence from anthropophony and quiet could provide information related to anthropogenic impact on ecosystems and areas with quiet landscapes (Pavan, 2017), which in some cases could indicate ecosystem decline if there were previously high levels of biophony (Quinn et al., 2022). Indices strongly affected by geophony and interference (ACI, AEI, M, Rugosity, H, Ht, and ADI) should be used and

interpreted judiciously because their outcomes may be more of a reflection of ambient noise than the ecology of the study domain.

Our findings corroborate some studies on soundscape components and acoustic indices yet contradict others. For example, Bradfer-Lawrence et al. (2019) found that high amounts of anthropophony and geophony led to high ADI values, while we found the inverse for anthropophony and a more complex but generally inverse relationship with geophony in our data, which aligns more with findings from Fairbrass et al. (2017). Patterns for H were consistent with prior work, namely, that anthropophony, geophony, and interference influence H to a larger degree than biophony; however, this contrasts with Ross et al. (2021), where H was relatively insensitive to confounding soundscape influences. Sueur et al. (2008) recommend a low-pass frequency filter of geophony in our dataset. Other work has demonstrated ACI to be robust against rain events and NDSI as highly sensitive to rainfall (Sánchez-Giraldo et al., 2020). Here, we contribute to these findings and demonstrate ACI's significant relationship to wind-related geophony, consistent with Depraetere et al. (2012) and NDSI's lack of a significant relationship with wind events. Overall, many of our findings relating indices to observed richness agree with findings in other studies (e.g., ACI, BI, H, and M in BradferLawrence et al., 2019). However, we provide evidence that the anthropogenic (NDSI-α) and biotic (NDSI-β) components of NDSI have strong positive relationships with anthropophony and biophony, respectively, and may be more informative as separate indices than NDSI, which integrates both components. This finding corroborates prior results that demonstrate NDSI and the biotic component of NDSI are insensitive to confounding sonic conditions (Ross et al., 2021).

We also found that removing unwanted noise from interference events resulted in a more reliable application of ACI, as recommended in prior work (Fairbrass et al., 2017). This result supports the hypothesis that acoustic indices such as ACI may be generalizable across spatial and temporal domains and therefore measure universally meaningful aspects of the soundscape, so long as extraneous biasing sounds are accounted for beforehand. This is particularly important for the AM ARUs which recorded significantly more interference events than LG ARUs and could result in erroneous interpretation of interference as biophony. Even though biophony accounted for more variance in ACI than other soundscape components when removing interference, the level of data loss here may not be an ideal solution to improve index reliability for other recording archives. To avoid this data loss, the application of ACI would be most effective in sonically consistent environments using ARUs that are not prone to rapid, broad-frequency interference events. Recording schedules could also be designed to account for data loss to achieve desired total samples after the removal of interference. Previous studies have recommended accounting for low-amplitude sound pulses when applying ACI (Farina et al., 2016), though these methods do not appear to scale to our broad-frequency interference events. We believe source separation techniques will provide better options for events such as interference while minimizing data loss (Lin and Tsao, 2020) or future statistical analysis of the interaction between interference and other soundscape components.

Indices strongly affected by multiple soundscape components are more difficult to interpret in complex landscapes that contain multiple sonic conditions across numerous land-use and land-cover types like Sonoma County, while indices affected to a lesser extent by soundscape components are more straightforward to interpret (i.e., NDSI, R, and BI). Indicators strongly

influenced by biophony and robust to the other soundscape components may be applied more broadly (NDSI and NDSI-β) than indices where biophony does not have as much influence (H, Hs, M). Depending on the acoustic index, anthropophony may be a meaningful category (e.g., NDSI-α) or a non-meaningful source of noise (e.g., ACI). We summarized the trends of acoustic indices with soundscape components (Table 3-6), and we recommend that future ecoacoustic work apply appropriate indices given a study's sonic characteristics and desired acoustic target (e.g., sound pollution or biotic events).

Table 3-6. When applying acoustic indices, consider the potential effects of soundscape components. We supply the dominant interpreted directional effects of the soundscape component on the index based on Fig. 3-2 and PDPs in Appendix 3A.

Soundscape component	Indicators positively related to the component	Indicators negatively related to the component	Indicators not sensitive to the component
Anthropophony	ACI, AEI, M, NDSI-α	ADI, H, H_s, H_t, NDSI, NDSI-β, RN, Rugosity, SFM, ZCR	BI
Biophony	ACI, AEI, BI, M*, NDSI†, NDSI-β†, RN, ZCR	ADI, H*, H_s*, H_t, NDSI-α†, Rugosity, SFM	
Geophony	ACI, AEI, BI, M	ADI, H, H_s, H_t, NDSI-β, Rugosity, ZCR	NDSI, NDSI-α, RN, SFM
Interference	ACI, AEI, BI, M, NDSI, NDSI-β, RN, SFM	ADI, H, H_s, H_t, NDSI-α, Rugosity, ZCR	

† Biophony is the most influential soundscape component, excluding Quiet
* Biophony is the least influential soundscape component, excluding Quiet

3.5.2. Predicting bird diversity with derived acoustic features

We sought to relate acoustic indices directly to species richness and followed recommendations to use indices in combination instead of independently (e.g., Buxton et al., 2018a; Bradfer-Lawrence et al., 2019). Our GAM approach with only acoustic indices (Eq. 4) resulted in comparable accuracy (adj-R^2 = 0.53; deviance/goodness of fit = 52.5%; n = 1,185) to another multi-index study that implemented Random Forest models to predict bird species vocalizations (R^2 = 0.40–0.51; Buxton et al., 2018a). Even though these findings were

comparable, the ability of acoustic indices to represent bird species richness may be limited in acoustically-complex and heterogeneous environments (Buxton et al., 2018a) based on the effects of non-ecologically meaningful sounds on acoustic index patterns (Table 3-6). However, bird species richness models are not the only utility for multi-index models. Ensemble acoustic index modeling is promising when distinguishing degraded habitats from healthy habitats in marine settings (Williams et al., 2022) and terrestrial habitats (BradferLawrence et al., 2020) and monitoring vertebrate groups (Allen-ankins et al., 2023).

To extend the ability for acoustic indices to reflect biodiversity, we leveraged the automated detection of soundscape components to provide an empirical approach to predicting bird species diversity. The combined acoustic indices and biophony GAM (Eq. 5) slightly overpredicted at low species richness values and underpredicted at higher values (Figure 3-5). This pattern, particularly at higher richness values, may reflect the saturation and the inability of recordings to capture increases in species richness (Burivalova et al., 2019) or the model's tendency to favor mean behavior over extreme values. However, the former explanation relates to sonic condition interactions, which result in the "masking" of unique signals. Notably, unlike anthropophony and geophony sources, animals are known to adjust their vocalization frequency and amplitude to increase propagation and success of signal reception (Pijanowski et al., 2011). Another consideration is that our detection data only included 54 species (a max richness of 48 in the morning data and 54 in the 24-h data), and a model with higher species richness values could resolve this potential saturation issue.

In the morning-only biophony and acoustic indices GAM, positive trends in biophony and NDSI-β support their utility as biodiversity indicators. Furthermore, reducing acoustic covariates

from 16 to 11 suggests high redundancy in acoustic indices' ability to explain variation in bird species richness. The negative trend in M aligns with other work showing that higher signal vocalizing activity leads to lower M values (Bradfer-Lawrence et al., 2019). However, this change from having a weak influence from biophony (Eq. 1) to a strong influence as a predictor of bird species richness demonstrates the need for caution when interpreting index effects.

In the 24-h dataset, acoustic indices were better predictors of bird species richness (Eq. 4) than biophony alone (Eq. 3). However, the performance for the morning-only biophony model (Eq. 3) had comparable performance to the acoustic index, 24-h GAM (Eq. 4). Our modeling is consistent with other work relating individual indices to bird richness and biophony (e.g., ADI: Machado et al., 2017; Ht, NDSI, ADI, AEI, M, ACI; Ross et al., 2021). Additionally, the ability for indices to better represent richness in a modeling framework over correlative analyses is consistent (Appendix 3A; Mammides et al., 2017). Overall, though, correlative relationships among indices and bird richness were weak here ($|\rho| \leq 0.35$), and biophony had the strongest relationship ($\rho = 0.56$ for the morning only), reinforcing our emphasis on the utility of biophony as a more robust predictor of biodiversity metrics compared to acoustic indices. The ability of biophony to explain similar levels of species richness compared to 15 acoustic indices supports the utility of biophony as a viable ecoacoustic metric on par with combined acoustic indices and allows for targeted morning-only data collection. If biophony is unavailable, acoustic indices may provide higher performance and representation of bird species richness when using a 24-h sampling approach. Adding biophony to acoustic indices in an ensemble model (Eq. 5) increased the performance beyond comparable acoustic index models of bird diversity (Eq. 4; Buxton et al., 2018a).

3.5.3. Extension of Biophony in ecoacoustics

Our analyses corroborate a finding from Fairbrass et al. (2017) and Ross et al. (2021) that acoustic indices applied in complex acoustic environments can reflect biotic activity, yet other sound sources significantly affect their replicability and make interpretation non-trivial. Namely, our analyses support NDSI-β as both relatively robust in varying sonic conditions (Ross et al., 2021) and the most representative index of bird species richness when biophony is unavailable. It is thus essential to understand assumptions and underlying effects of non-biotic sounds before interpreting index values (Gasc et al., 2015).

These analyses build on research that establishes that the cumulative amount of biophony, here quantified by our CNN, is a robust ecoacoustic indicator of biodiversity (Mullet et al., 2016; Fairbrass et al., 2019; Grinfeder et al., 2022). Compared to more traditional acoustic indices, biophony has the potential to be more directly related to animal diversity as it captures biotic events as opposed to an abstracted summary of acoustic energy as acoustic indices do, which at present can result in non-trivial relationships with underlying acoustic sources of sound. Furthermore, we believe CNN-derived metrics like biophony are potentially more robust to non-calibrated ARUs and methodological inconsistencies that influence acoustic index variation, but this issue requires formal study.

Our approach to quantifying biophony requires generalizability tests with other datasets and locations. Ideally, a future study would fine-tune the ABGQI CNN to reflect sounds in a different study domain. Even though this approach would require significant work to generate reference sound data, biophony as a general soundscape class is less intensive to generate from CNNs than those that detect species presences, which involve expert knowledge to identify

reference species vocalizations (Clark et al., 2023). Non-experts can help in the effort to generate biophony reference data, as it does not involve species identification (Snyder et al., 2022). If not used as an ecoacoustic metric, biophony is valuable as a tool to aid researchers in filtering the everexpanding size of acoustic datasets into relevant subsets for more efficient, targeted analyses (Pijanowski and Brown, 2022). In future efforts, periods with biophony may be broken into finer taxonomic granularity to understand specific family or species dynamics (Hao et al., 2022), expanding biophony's value in conservation efforts (Dumyahn and Pijanowski, 2011).

With the rapidly advancing field of deep learning, CNN classification and source separation techniques (Lin and Tsao, 2020) may soon be more user-friendly to the point of competing with acoustic index calculations. Even if adapting our CNN or the development of a new CNN is outside a project's scope, acoustic indices remain an approachable and relatively low-cost option for generating acoustic activity summaries, ensuring correct assumptions and a general understanding of how non-biotic noise influences values. Non-index quantification of soundscape dynamics using metrics like biophony may be more generalizable to measure changes in biodiversity and soundscapes incurring disturbance (e.g., increased anthropophony) and for targeted conservation and study (Grinfeder et al., 2022).

3.6. CONCLUSION

Our overarching goal was to understand how sonic conditions represented using soundscape components affect acoustic indices and improve interpretation. We applied a CNN classifier to detect soundscape components automatically, used statistical models to investigate their relationship with 15 common acoustic indices, and provided recommendations to contextualize the effects of soundscape components when applying these 15 acoustic indices. We found

combining biophony and acoustic indices particularly informative for predicting bird species richness. We also validated how acoustic indices more reliably reflect biophony when non-biotic ambient noises are quantified and excluded from models. We aimed to provide a more flexible method to measure species richness acoustic activity than species-level identification in novel acoustic datasets. Our work supports applying more automated methods, such as CNN soundscape component detection, to acoustically assess and monitor biodiversity by establishing the combination of biophony and acoustic indices as useful ecoacoustic monitoring tools for bird species richness.

3.7. APPENDIX 3A

3.7.1. Acoustic Index code-related methods

Table 3A-1. Acoustic indices and descriptors we selected for soundscape analysis (n = 15) with the R package implementation and relevant parameters used for calculations.

Index name	Abbrev.	R implementation	Parameters
Acoustic Complexity Index	ACI	seewave::ACI [1]	flim = 1-10
Acoustic Diversity Index	ADI (=H')	soundecology::acoustic_diversity [2]	max_freq = 10,000; db_threshold = -50; freq_step = 1000; shannon = TRUE
Acoustic Evenness Index	AEI	soundecology::acoustic_evenness [2]	max_freq = 10000; db_threshold = -50; freq_step = 1000
Bioacoustic Index	BI	soundecology::bioacoustic_index [2]	Defaults
Acoustic Entropy	H	seewave::H [1]	wl = 512
Temporal Entropy	H_t	seewave::th [1]	Defaults
Spectral Entropy	H_s (also H_f)	seewave::sh [1]	Defaults
Median of amplitude envelope	M	seewave::M [1]	Defaults
Normalized Difference Soundscape Index	NDSI	soundecology::ndsi [2]	fft_w = 1024; anthro_min = 1000; anthro_max = 2000; bio_min = 2000; bio_max = 10,000
Normalized Difference Soundscape Index-anthropophony	NDSI-α	soundecology::ndsi [2]	fft_w = 1024; anthro_min = 1000; anthro_max = 2000
Normalized Difference Soundscape Index - biophony	NDSI-β	soundecology::ndsi [2]	fft_w = 1024; bio_min = 2000; bio_max = 10,000
Roughness†	RN	seewave::roughness [1]	Defaults
Rugosity†	RUGO	seewave::rugo [1]	Defaults
Spectral flatness†	SFM	seewave::sfm [1]	Defaults

Zero-Crossing Rate[†]	ZCR	seewave::zcr [1]	Defaults

[†] Acoustic descriptors. Here, we refer to all indices and descriptors as acoustic indices.

[1] seewave: Sueur J, Aubin T, Simonis C (2008). seewave: a free modular tool for sound analysis and synthesis. Bioacoustics, 18: 213-226. R package version 2.2.0.

[2] soundecology: Villanueva-Rivera LJ, Pijanowski BC (2018). _soundecology: Soundscape Ecology_. R package version 1.3.3.

3.7.2. Generalized Additive Model (GAM) full methods

All GAMs were fit using the following procedure. This information supplements the overview in section 3.2.3.1 of the main manuscript.

1. If soundscape components were covariates, we normalized them from 0 to 1 using min-max normalization. This allowed for comparison between soundscape components within GAMs.
2. We fit GAMs using the applicable full model (e.g., Eq. 1), implementing GAM fitting and supporting functionality from the R package *mgcv* (Chambers & Hastie, 2017; Wood, 2017). GAMs were fit using an appropriate error structure (e.g., Gaussian for indices with all real numbers, beta for indices constrained from 0 to 1, or gamma for positive, real numbers) with corresponding link functions (Table 3-4). Model parameters were smoothed using penalized thin-plate regression splines with maximum-likelihood estimation. We penalized splines by specifying a maximum number of basis functions (k = 5) based on the assumption that soundscape components are limited to lower-order relationships with index values and more complex polynomials are not ecologically practical.

3. We ensured we did not violate model assumptions using *mgcv's gam.check* for model diagnostics. If gross outliers affected interpretation, we investigated the cause of the outlier, removed the outlier, and updated the model.
4. Once appropriate error structures were determined, we iteratively reduced model complexity if applicable, eliminating the covariate with the highest non-significant p-value. We did not reduce the model complexity if all covariates were significant ($p < 0.05$) or according to the likelihood ratio testing (LRT). We tested the updated, nested models using the LRT of nested models, *lrtest*, from the R package *lmtest* (Zeileis & Hothorn, 2002). We implemented a significance level of 0.05 and therefore rejected the null hypothesis and selected the more complex, full model if the LRT p-value ≤ 0.05 while we accepted the null hypothesis in favor of the nested model if the LRT p-value > 0.05.
5. We plotted GAM partial dependence plots (PDPs) using *gratia's draw* function. PDPs show how the mean acoustic index value (y-axis) changes across each covariate's domain (x-axis). We interpreted partial dependence slopes as the magnitude of impact a covariate had on the response where ranges further from $x = 0$ were more influential for those areas in the covariate's domain. However, a covariate with near-zero slopes across its domain tended to be non-significant, and when included, those regions had minimal effect on the response.
6. Finally, to not overstate the effects of covariates and to account for concurvity effects we used the function *concurvity* from *mgcv* and investigated pair-wise, worst-case scenario concurvity values > 0.8. In these cases, we compared the fit model PDP trends to the raw

data relationship of the high concurvity covariate with the response. If the two trends were notably divergent (i.e., the fit PDP significantly differed from the raw trend) we did not include PDP effects in our discussions. See Supplementary section 3.6.5 for high concurvity model covariates.

7. To further visualize covariate influence, we calculated the 95% confidence interval (CI) first derivative of the estimated smooths for covariates in the final model using the R package *gratia's* function *derivatives* (Simpson, 2022) which produces a median slope estimate at 200 equal-interval values in the covariate domain. Because smooth functions were limited to lower-order polynomials, tail behavior was sometimes impacted by sparse high and low covariate values. Considering this, we reported partial dependence slopes for the middle 99% of covariate values (i.e., dropping the most extreme lower and upper slope values for visualization).
8. For final GAMs, we reported error distribution, link function, adj-R^2, and deviance, where applicable. We report deviance because it better approximates non-normal error distributions for GAMs; however, deviance and adj-R^2 are equivalent for GAMs with a Gaussian error distribution.

3.7.3. Generalized additive model summaries in R

Final GAM summaries are located on the project's GitHub (github.com/CQuinn8/Ecoacoustic_indicators/tree/main/models).

3.7.4. Generalized additive model partial dependence plots

The full model (Eq.1) was used to begin fitting all GAMs and the procedure in 3.6.2 to select the final model. Below are the final model PDPs for each acoustic index model displayed alphabetically.

$$E(Acoustic\ Index_i) = ARU + log\ (Minutes\ recorded) + f_1(Anthropophony_i) + f_2(Biophony_i) + f_3(Geophony_i) + f_4(Quiet_i) + f_5(Interference_i), \quad \text{(Eq. 1)}$$

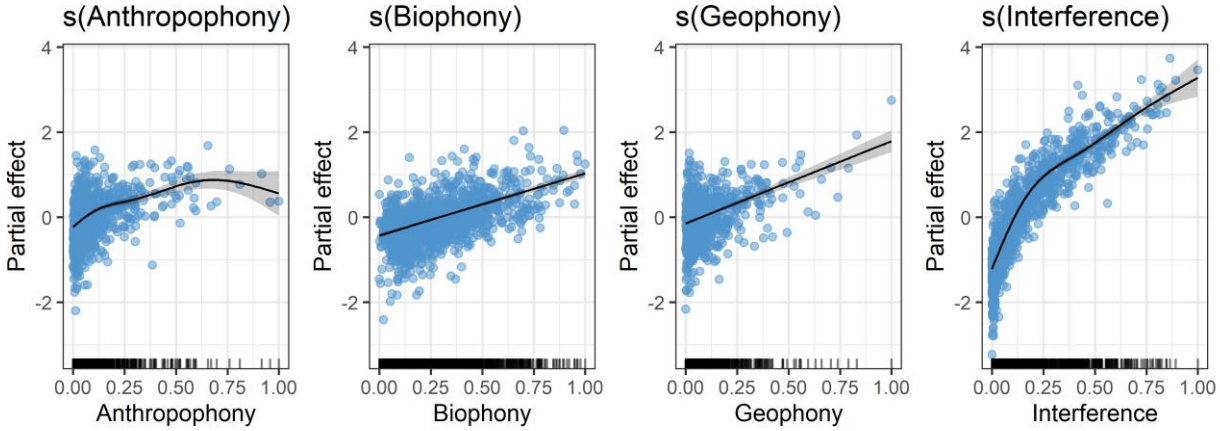


Figure 3A-1. ACI soundscape component PDP

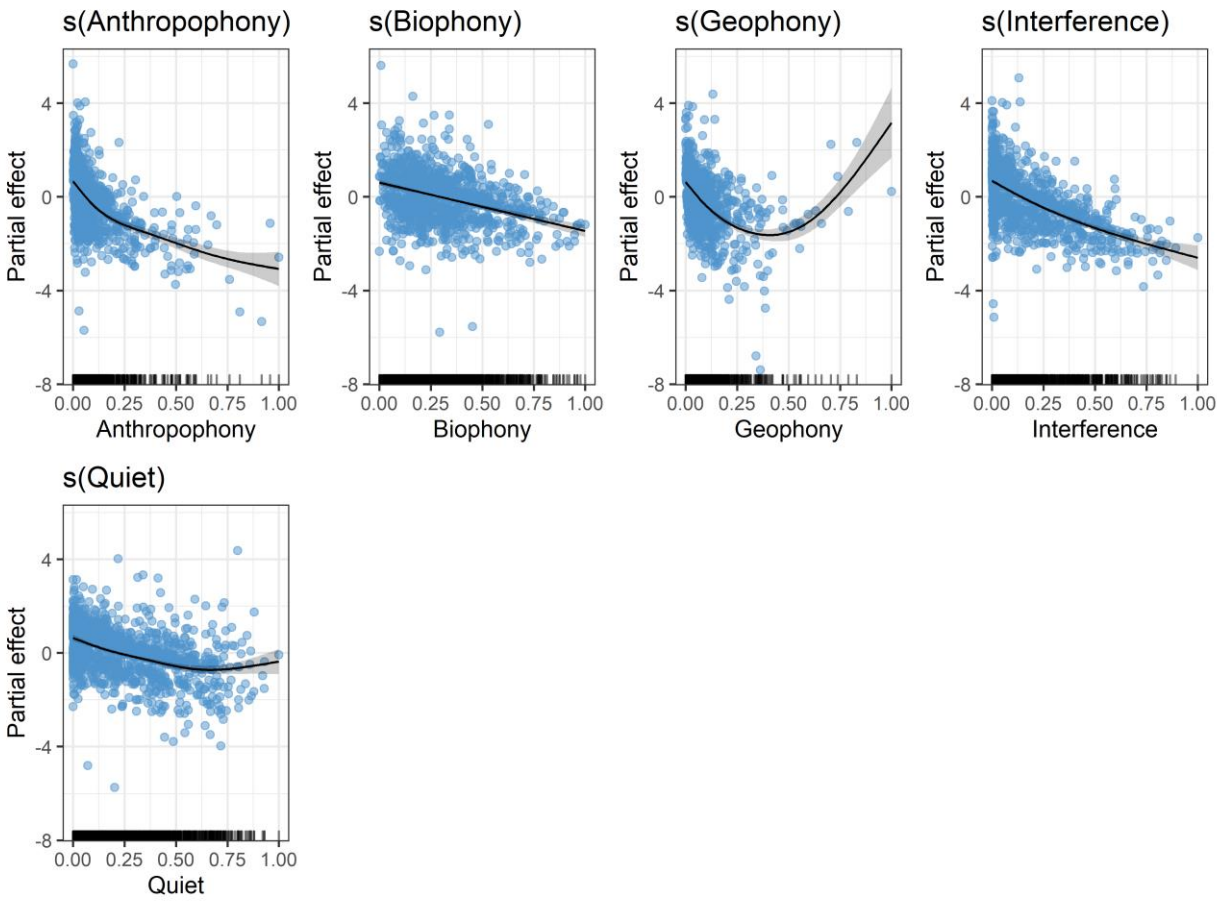


Figure 3A-2. ADI soundscape component PDP

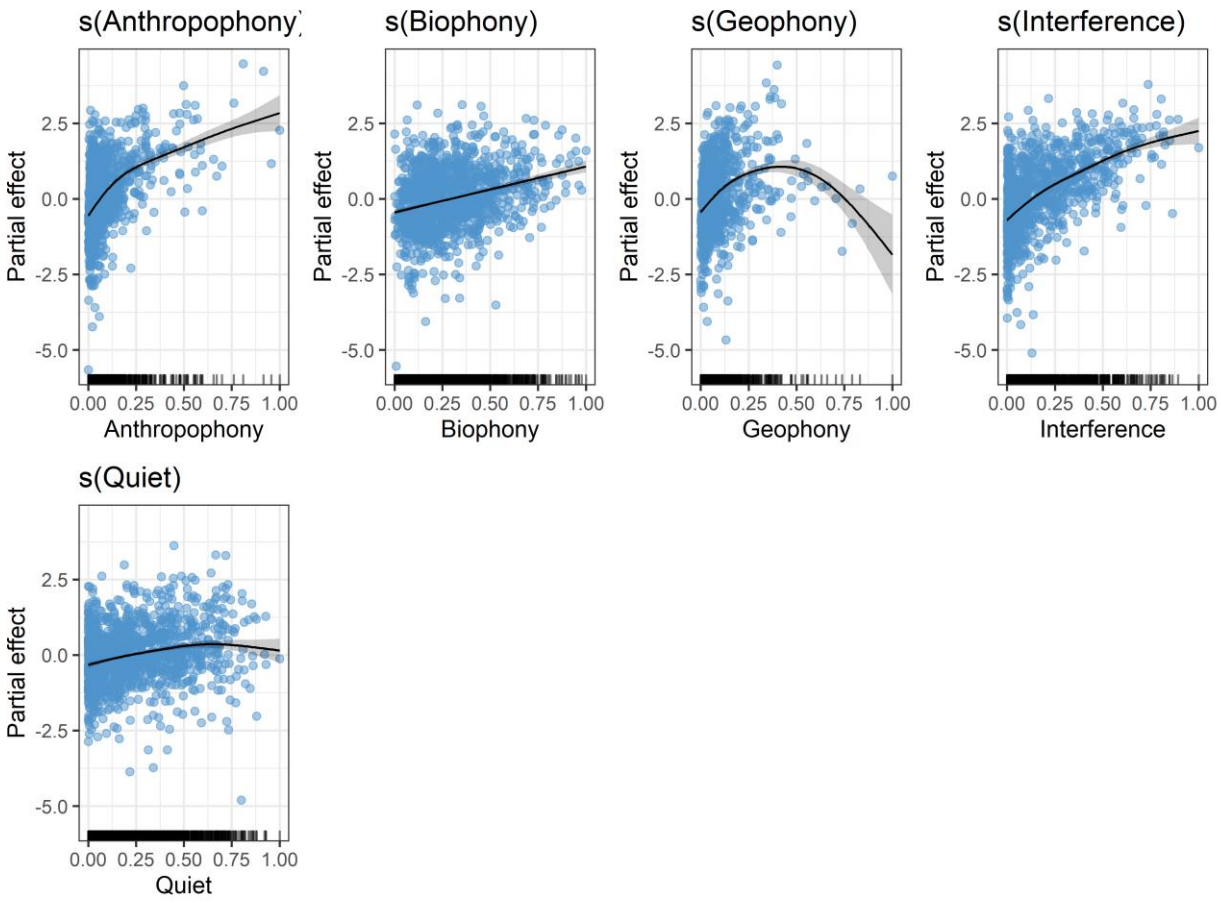


Figure 3A-3. AEI soundscape component PDP

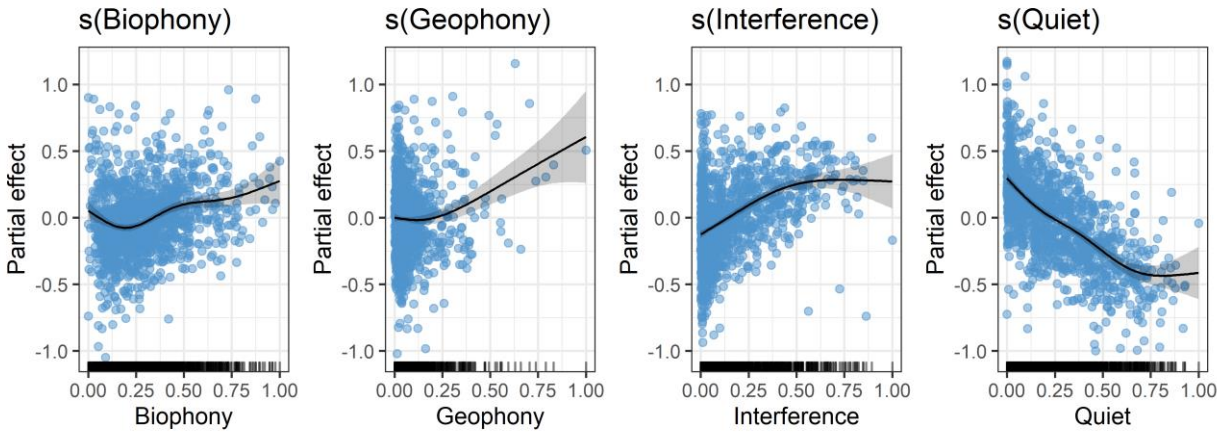


Figure 3A-4. BI soundscape component PDP

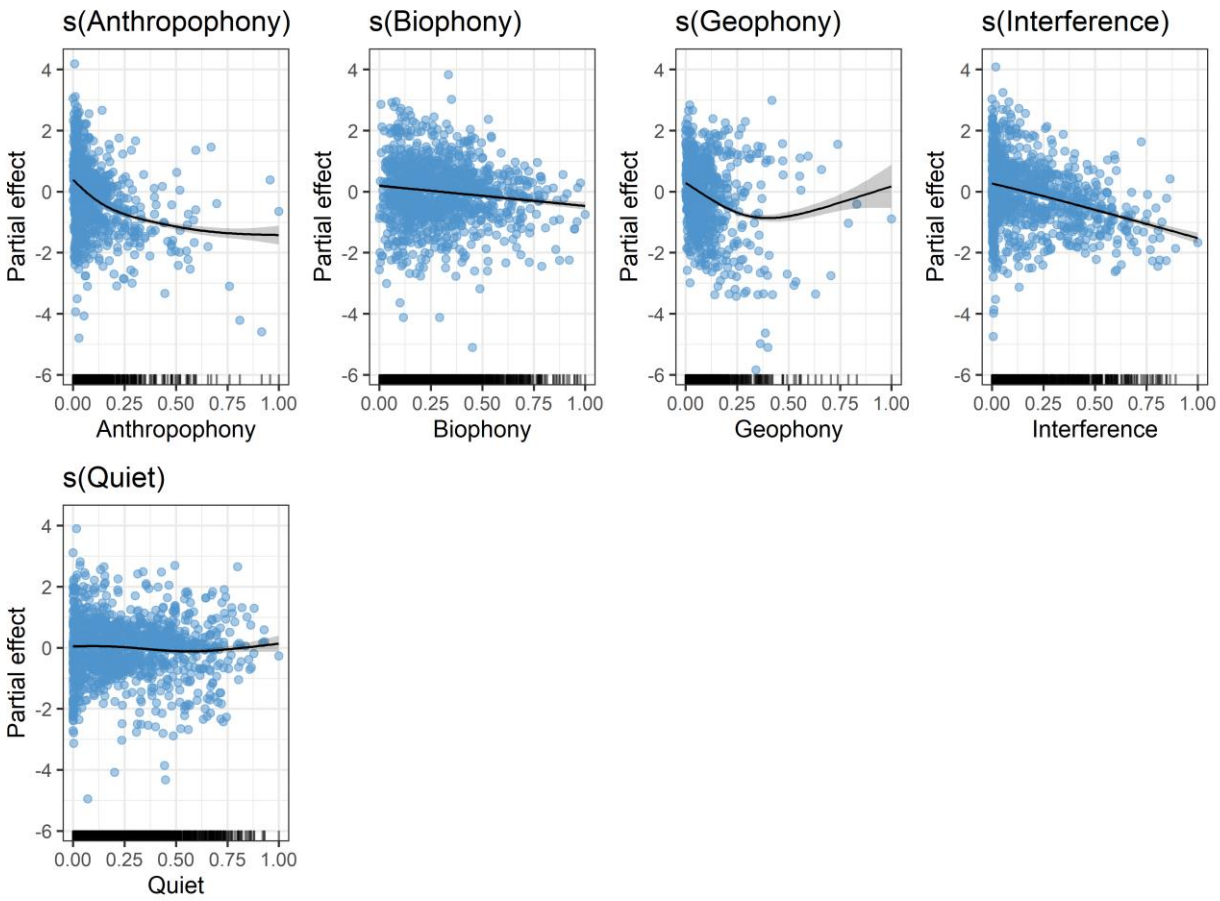


Figure 3A-5. H soundscape component PDP

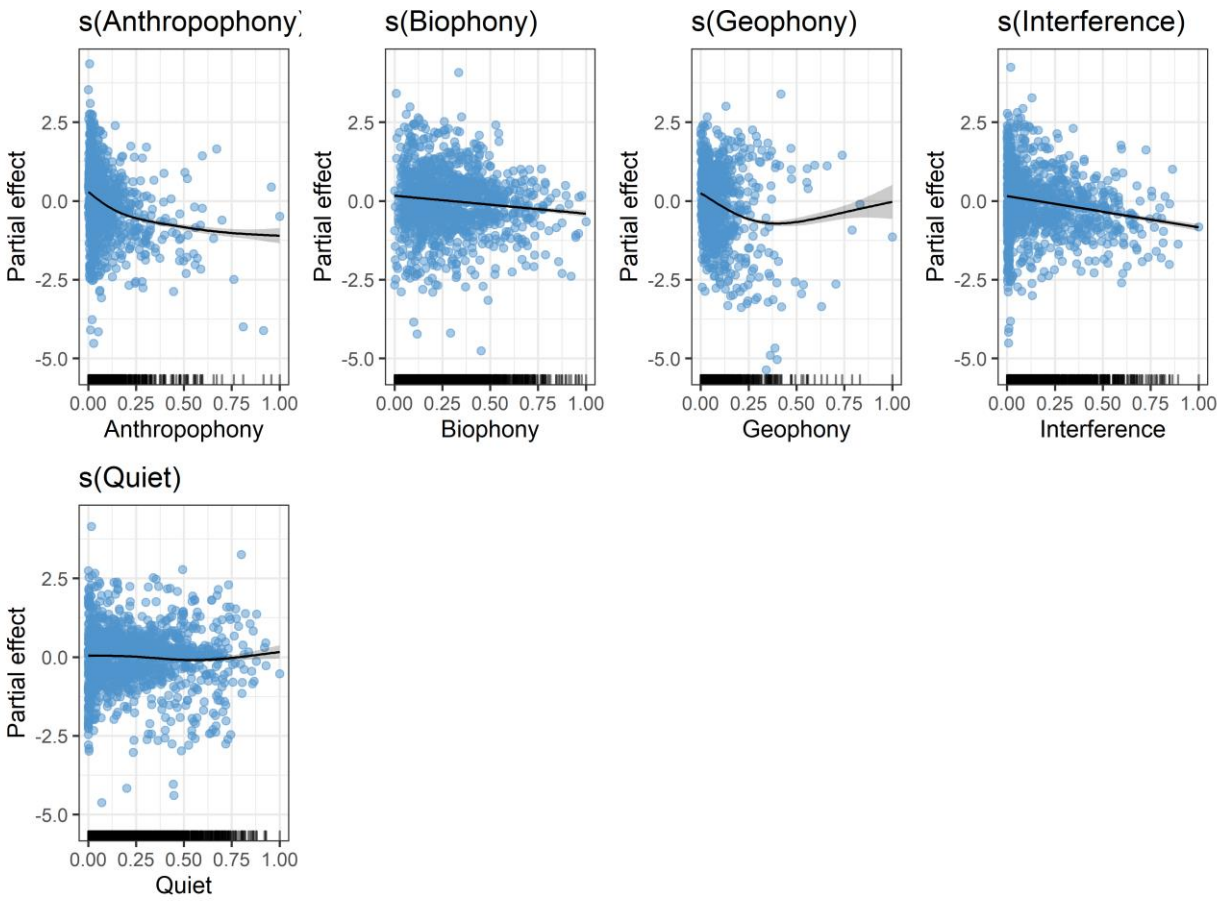


Figure 3A-6. H_s soundscape component PDP

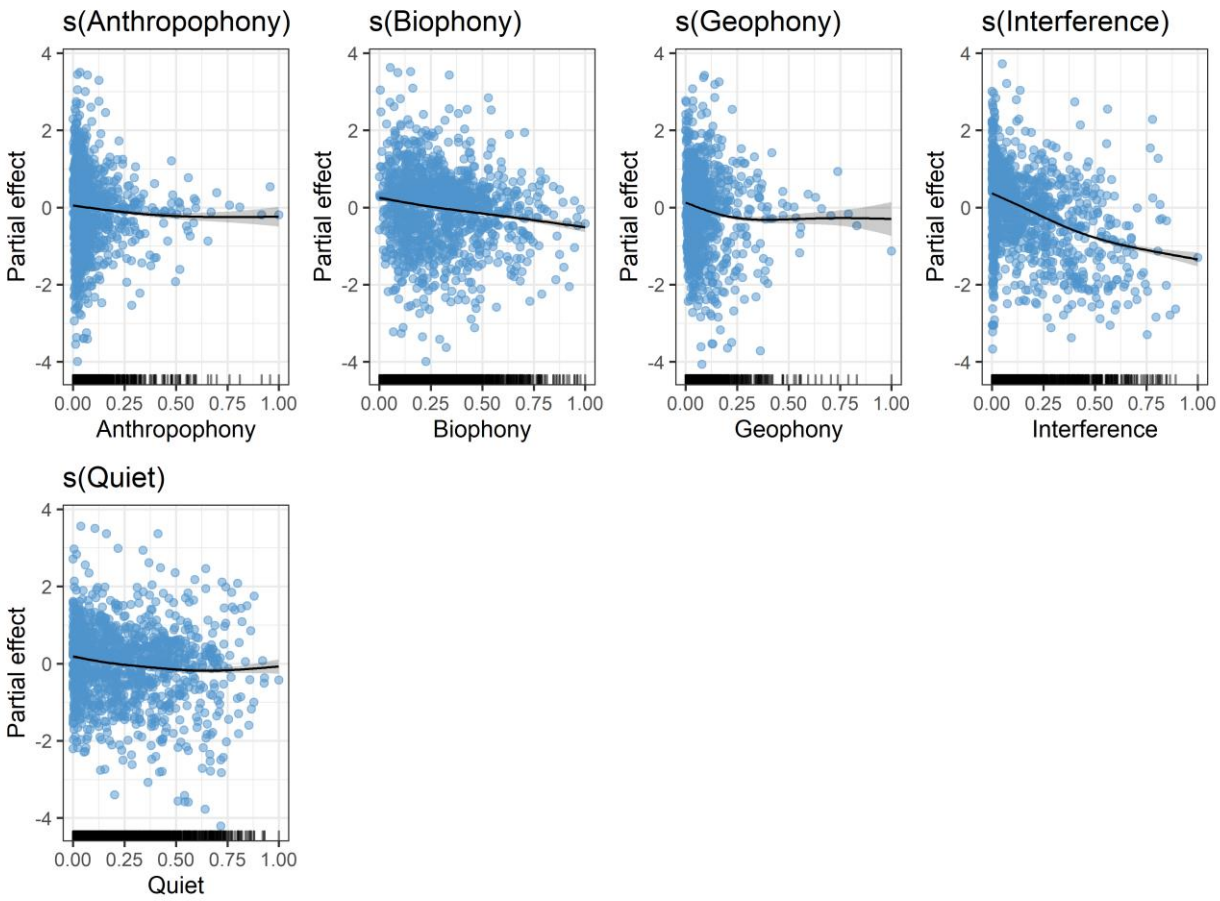


Figure 3A-7. H_t soundscape component PDP

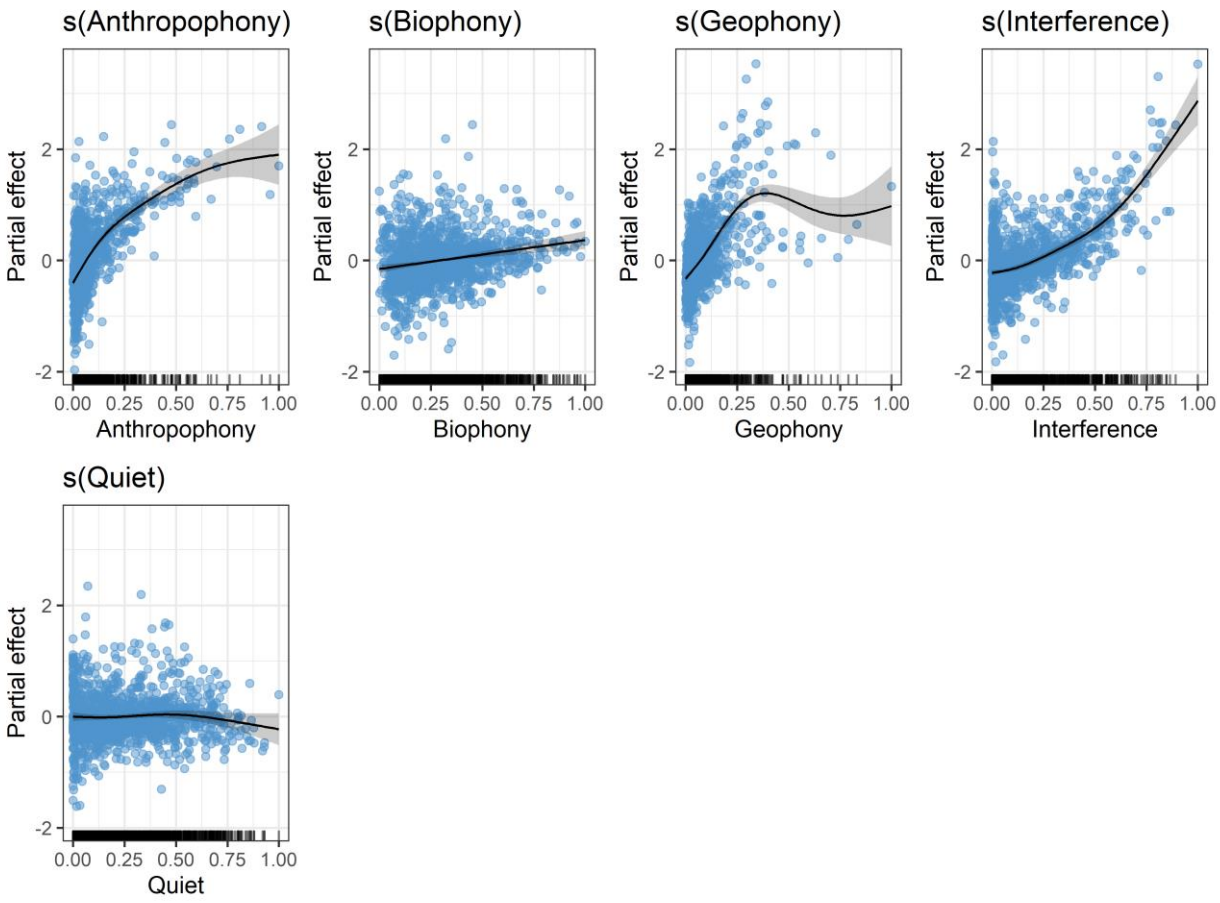


Figure 3A-8. M soundscape component PDP

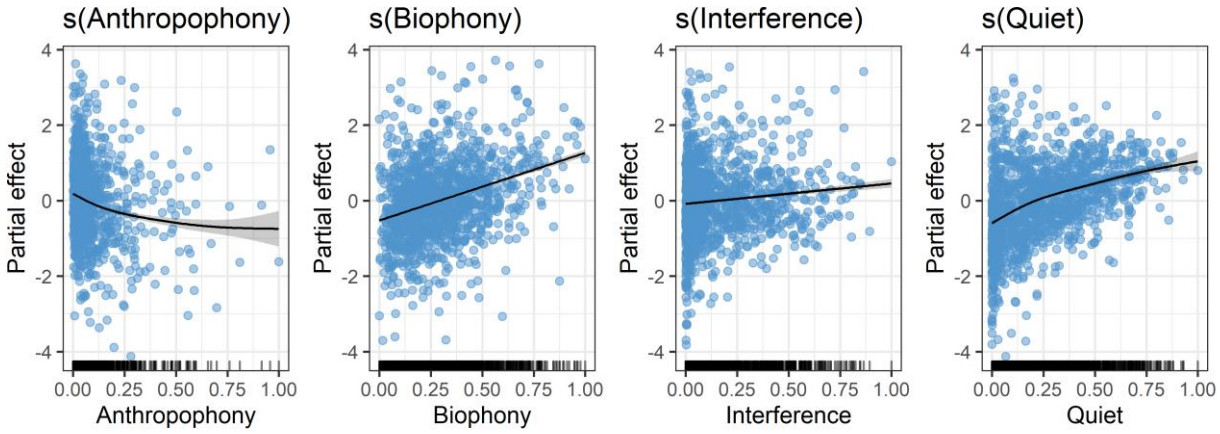


Figure 3A-9. NDSI soundscape component PDP

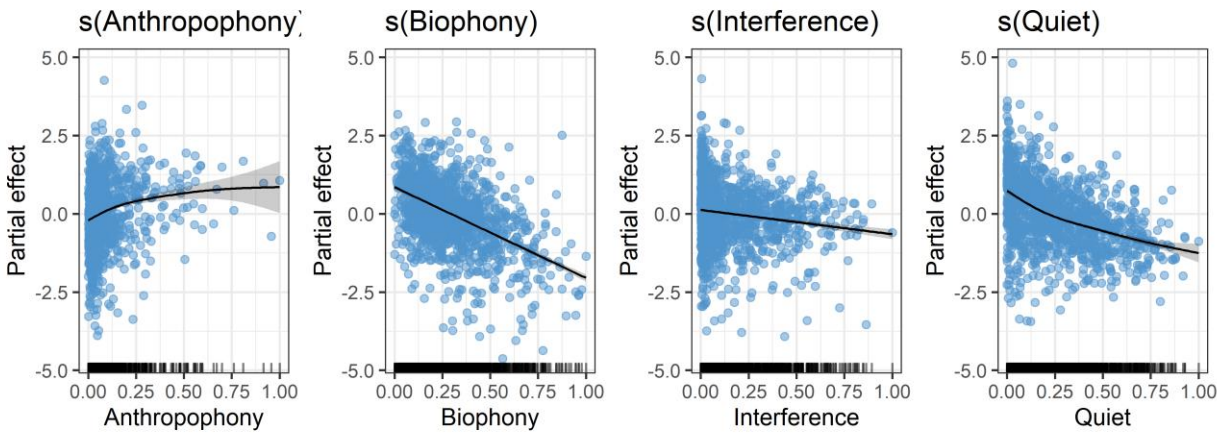


Figure 3A-10. NDSI-α soundscape component PDP

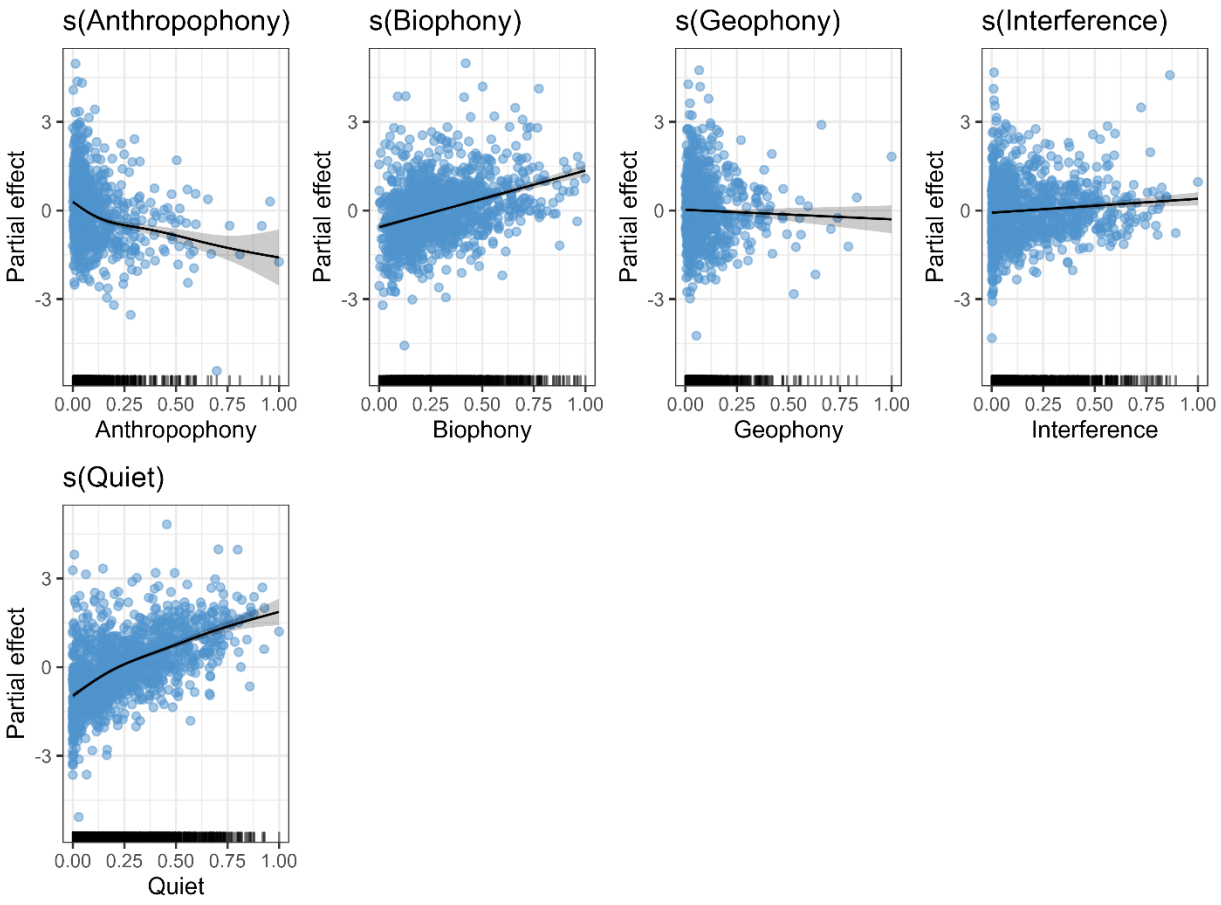


Figure 11. NDSI-β soundscape component PDP

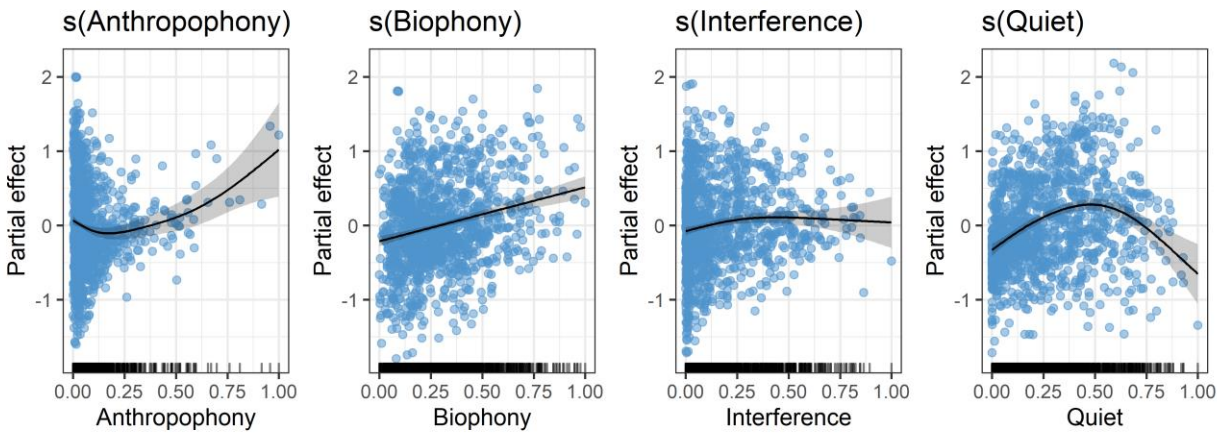


Figure 12. RN soundscape component PDP

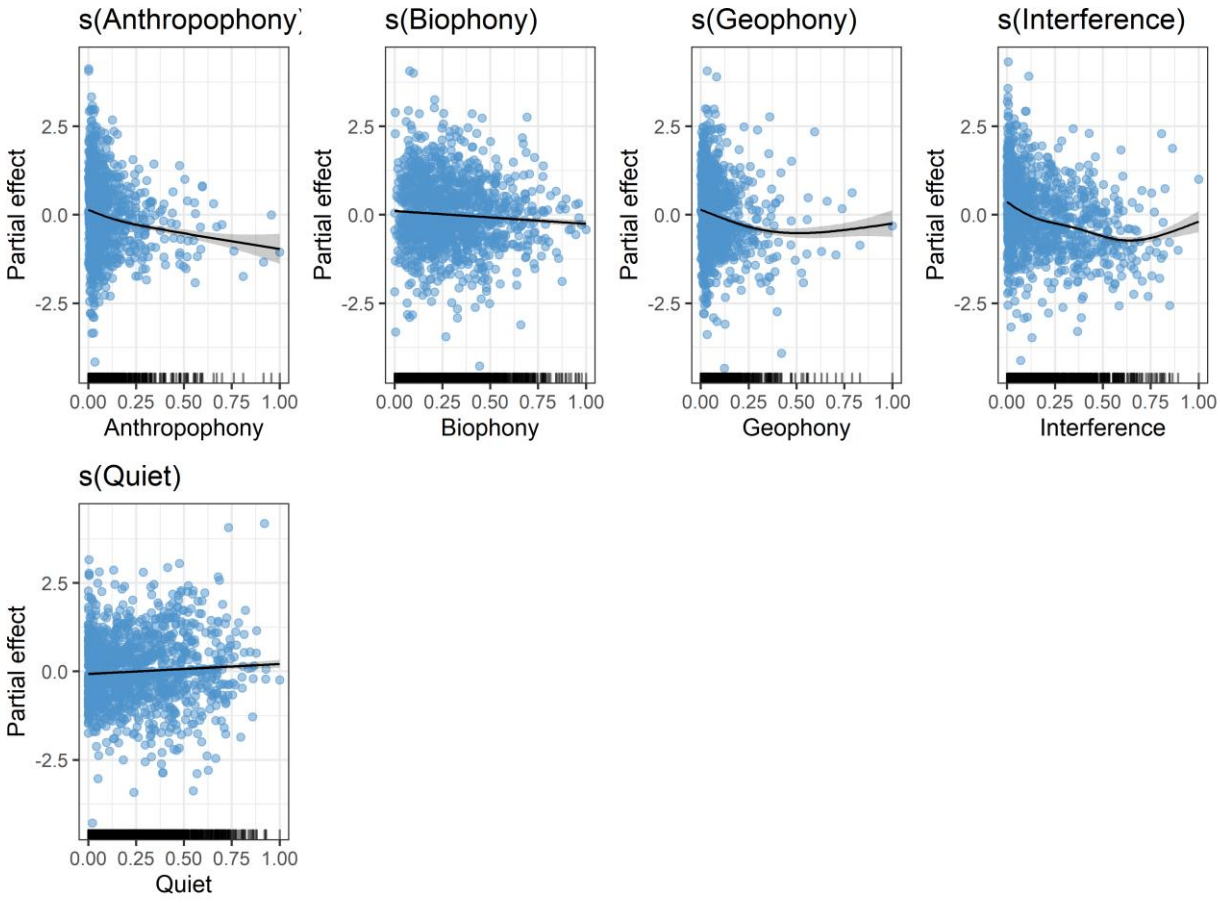


Figure 13. RUGO soundscape component PDP

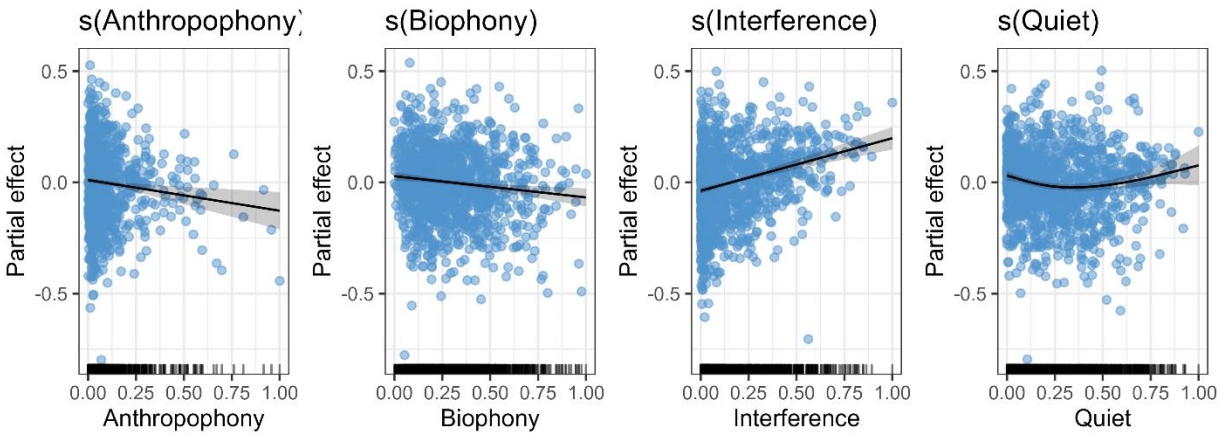


Figure 14. SFM soundscape component PDP

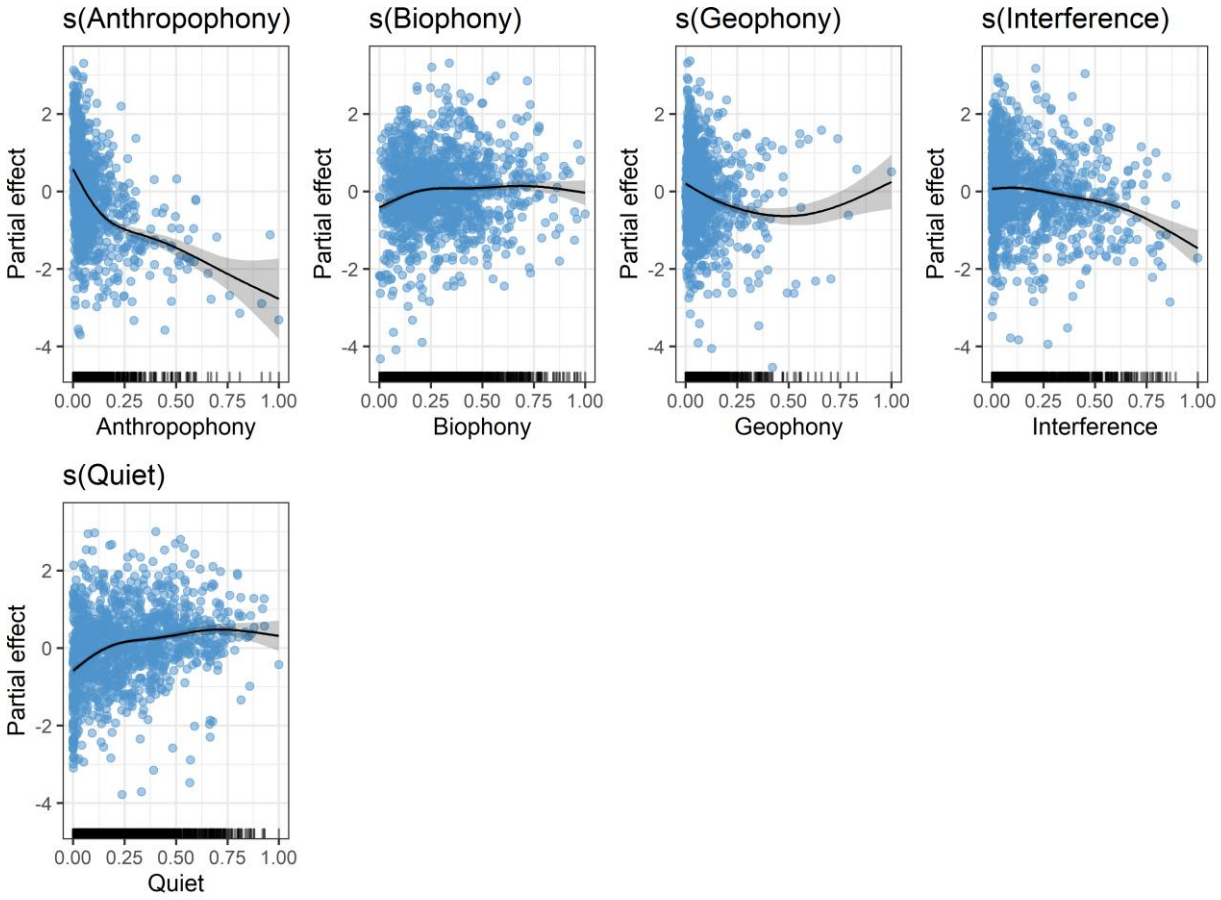


Figure 15. ZCR soundscape component PDP

3.7.5. Generalized additive model covariates with concurvity issues

We used mgcv::concurvity's worst case scenario (Wood, 2017) to assess issues with concurvity in final GAMs. Tables below highlight covariate pairs with worst case scenario concurvity values > 0.8. Covariates with structure covariation issues that impact interpretability are denoted with †.

Eq.1: No covariates with worst case scenario values > 0.80.

Eq.2(24-hr and morning-only): No covariates with worst case scenario values > 0.80.

Eq.3 (24-hr and morning-only): No covariates with worst case scenario values > 0.80.

Table 3A-2. Eq.4 (24hr): ADI and H may have oversimplified tail behavior but across each domain do not demonstrate concerning structural pattern issues. NDSI is the only covariate with concurvity and structural covariance issues.

	ADI	AEI	H	NDSI†	NDSI-α	NDSI-β
ADI	1.00	0.95	0.86	---	---	---
AEI		1.00	0.83	---	---	---
H			1.00	---	---	---
NDSI†				1.00	0.94	0.98
NDSI-α					1.00	0.91
NDSI-β						1.00

Table 3A-3. Eq.4 (morning only): H_s has issues in sparse, lower tail region potentially but overall no concerns with structure. NDSI is the only covariate with concurvity and structural covariance issues.

	ADI	AEI	H	H_s	NDSI†	NDSI-α	NDSI-β	ZCR
ADI	1.00	0.95	---	0.83	---	---	---	---
AEI		1.00	0.83	---	---	---	---	---
H			1.00	0.98	---	---	---	0.81
H_s				1.00	---	---	---	0.81
NDSI†					1.00	0.95	0.99	---
NDSI-α						1.00	0.93	---

NDSI-β					1.00	---
ZCR						1.00

Table 3A-4. Eq.5 (24 hr): ADI, NDSI and, NDSI-α hav concurvity and structural covariance issues.

	ADI†	AEI	H	NDSI†	NDSI-α†	NDSI-β
ADI†	1.00	0.95	0.86	---	---	---
AEI		1.00	0.83	---	---	---
H			1.00	---	---	---
NDSI†				1.00	0.94	0.98
NDSI-α†					1.00	0.91
NDSI-β						1.00

Table 3A-5. Eq.5 (morning only): H has issues in sparse, lower tail region potentially but overall no concerns with structure. NDSI and NDSI-α (at high domain values) are the only covariates with concurvity and structural covariance issues.

	ADI	AEI	H	NDSI†	NDSI-α	NDSI-β	ZCR
ADI	1.00	0.95	---	---	---	---	---
AEI		1.00	0.83	---	---	---	---
H			1.00	---	---	---	0.81
NDSI†				1.00	0.95	0.99	---
NDSI-α					1.00	0.93	---
NDSI-β						1.00	---
ZCR							1.00

3.7.6. Eq.5 morning-only generalized additive model partial dependence plots

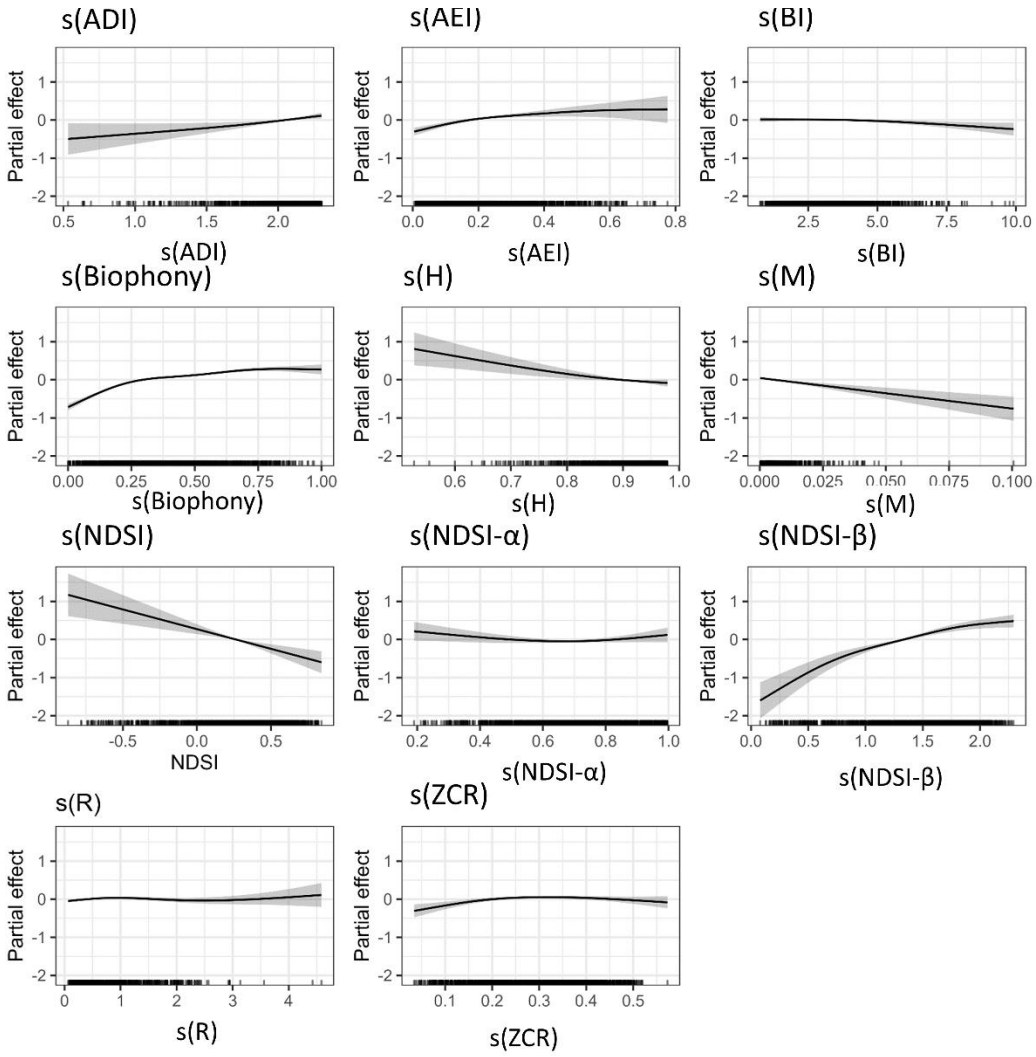


Figure 3A-16. PDPs with shaded 95% CIs for Biophony and acoustic indices in the morning-only bird species richness GAM. The final GAM included the logarithmic fixed effect number of recordings. The y-axis shows the mean model (black line) centered on zero, and the x-axis shows covariate values.

3.8. REFERENCES

Allen-ankins, S., Mcknight, D. T., Nordberg, E. J., Hoefer, S., Roe, P., Watson, D. M., et al. (2023). Effectiveness of acoustic indices as indicators of vertebrate biodiversity. *Ecological Indicators*, 147(August 2022), 109937. doi.org/10.1016/j.ecolind.2023.109937

Atauri Mezquida, D., & Llorente Martínez, J. (2009). Short communication. Platform for bee-hives monitoring based on sound analysis. A perpetual warehouse for swarm's daily activity. *Spanish Journal of Agricultural Research*, *7*(4), 824. doi.org/10.5424/sjar/2009074-1109

Boelman, N. T., Asner, G. P., Hart, P. J., & Martin, R. E. (2007). Multi-trophic invasion resistance in Hawaii: Bioacoustics, field surveys, and airborne remote sensing. *Ecological Applications*, *17*(8), 2137–2144. doi.org/10.1890/07-0004.1

Bormpoudakis, D., Sueur, J., & Pantis, J. D. (2013). Spatial heterogeneity of ambient sound at the habitat type level: Ecological implications and applications. *Landscape Ecology*, *28*(3), 495–506. doi.org/10.1007/s10980-013-9849-1

Bradfer-Lawrence, T., Bunnefeld, N., Gardner, N., Willis, S. G., & Dent, D. H. (2020). Rapid assessment of avian species richness and abundance using acoustic indices. *Ecological Indicators*, *115*(April), 106400. doi.org/10.1016/j.ecolind.2020.106400

Bradfer-Lawrence, T., Gardner, N., Bunnefeld, L., Bunnefeld, N., Willis, S. G., & Dent, D. H. (2019). Guidelines for the use of acoustic indices in environmental research. *Methods in Ecology and Evolution*, *00*, 1–12. doi.org/10.1111/2041-210X.13254

Burivalova, Z., Game, E. T., & Butler, R. A. (2019). The sound of a tropical forest. *Science*, *363*(6422), 28–29. doi.org/10.1126/science.aav1902

Buxton, R. T., Agnihotri, S., Robin, V. V., Goel, A., & Balakrishnan, R. (2018a). Acoustic indices as rapid indicators of avian diversity in different land-use types in an Indian biodiversity hotspot. *Journal of Ecoacoustics*, *2*, 1–17. doi.org/10.22261/jea.gwpzvd

Buxton, R. T., McKenna, M. F., Clapp, M., Meyer, E., Stabenau, E., Angeloni, L. M., et al (2018b). Efficacy of extracting indices from large-scale acoustic recordings to monitor biodiversity. *Conservation Biology*, *32*(5), 1174–1184. doi.org/10.1111/cobi.13119

Clark, M. L., Salas, L., Baligar, S., Quinn, C. A., Snyder, R. L., Leland, D., et al. (2023). The effect of soundscape composition on bird vocalization classification in a citizen science biodiversity monitoring project. Ecol. Inf. 75, 102065. doi:10.1016/j.ecoinf.2023. 102065

Depraetere, M., Pavoine, S., Jiguet, F., Gasc, A., Duvail, S., & Sueur, J. (2012). Monitoring animal diversity using acoustic indices: Implementation in a temperate woodland. *Ecological Indicators*, *13*(1), 46–54. doi.org/10.1016/j.ecolind.2011.05.006

Dröge, S., Martin, D. A., Andriafanomezantsoa, R., Burivalova, Z., Fulgence, T. R., Osen, K., Rakotomalala, E., et al. (2021). Listening to a changing landscape: Acoustic indices reflect bird species richness and plot-scale vegetation structure across different land-use types in north-eastern Madagascar. *Ecological Indicators*, *120*. doi.org/10.1016/j.ecolind.2020.106929

Duarte, M. H. L., Sousa-Lima, R. S. S., Young, R. J., Vasconcelos, M. F., Bittencourt, E.,

Scarpelli, M. D. A., et al. (2021). Changes on soundscapes reveal impacts of wildfires in the fauna of a Brazilian savanna. *Science of the Total Environment*, *769*, 144988. doi.org/10.1016/j.scitotenv.2021.144988

Dumyahn, S. L., & Pijanowski, B. C. (2011). Soundscape conservation. *Landscape Ecology*, *26*(9), 1327–1344. doi.org/10.1007/s10980-011-9635-x

Eldridge, A., Guyot, P., Moscoso, P., Johnston, A., Eyre-Walker, Y., & Peck, M. (2018). Sounding out ecoacoustic metrics: Avian species richness is predicted by acoustic indices in temperate but not tropical habitats. *Ecological Indicators*, *95*(September), 939–952. doi.org/10.1016/j.ecolind.2018.06.012

Fairbrass, A. J., Firman, M., Williams, C., Brostow, G. J., Titheridge, H., & Jones, K. E. (2019). CityNet—Deep learning tools for urban ecoacoustic assessment. *Methods in Ecology and Evolution*, *10*(2), 186–197. doi.org/10.1111/2041-210X.13114

Fairbrass, A. J., Rennett, P., Williams, C., Titheridge, H., & Jones, K. E. (2017). Biases of acoustic indices measuring biodiversity in urban areas. *Ecological Indicators*, *83*(August), 169–177. doi.org/10.1016/j.ecolind.2017.07.064

Farina, A., Pieretti, N., Salutari, P., Tognari, E., & Lombardi, A. (2016). The Application of the Acoustic Complexity Indices (ACI) to Ecoacoustic Event Detection and Identification (EEDI) Modeling. *Biosemiotics*, 9(2), 227–246. doi.org/10.1007/s12304-016-9266-3

Gasc, A., Pavoine, S., Lellouch, L., Grandcolas, P., & Sueur, J. (2015). Acoustic indices for biodiversity assessments: Analyses of bias based on simulated bird assemblages and

recommendations for field surveys. *Biological Conservation*, *191*, 306–312. doi.org/10.1016/j.biocon.2015.06.018

Gibb, R., Browning, E., Glover-Kapfer, P., & Jones, K. E. (2018). Emerging opportunities and challenges for passive acoustics in ecological assessment and monitoring. *Methods in Ecology and Evolution*, *2019*(September 2018), 169–185. doi.org/10.1111/2041-210X.13101

Grinfeder, E., Haupert, S., Ducrettet, M., Barlet, J., Reynet, M. P., Sèbe, F., & Sueur, J. (2022). Soundscape dynamics of a cold protected forest: dominance of aircraft noise. *Landscape Ecology*, *37*(2), 567–582. doi.org/10.1007/s10980-021-01360-1

Hao, Z., Zhan, H., Zhang, C., Pei, N., Sun, B., & He, J. (2022). Assessing the effect of human activities on biophony in urban forests using an automated acoustic scene classification model. *Ecological Indicators*, *144*(September), 109437. doi.org/10.1016/j.ecolind.2022.109437

Hill, A. P., Prince, P., Piña Covarrubias, E., Doncaster, C. P., Snaddon, J. L., & Rogers, A. (2018). AudioMoth: Evaluation of a smart open acoustic device for monitoring biodiversity and the environment. *Methods in Ecology and Evolution*, *9*(5), 1199–1211. doi.org/10.1111/2041-210X.12955

Kasten, E. P., Gage, S. H., Fox, J., & Joo, W. (2012). The remote environmental assessment laboratory's acoustic library: An archive for studying soundscape ecology. *Ecological Informatics*, *12*, 50–67. doi.org/10.1016/j.ecoinf.2012.08.001

Krause, B. L. (2002). The Loss of Natural Soundscapes. *Earth Island Journal*, *17*, 27–29.

Krause, B. L. (1993). The niche hypothesis: a virtual symphony of animal sounds, the origins of musical expression and the health of habitats. *The Soundscape Newsletter*, *6*, 6–10.

Lin, T. H., & Tsao, Y. (2020). Source separation in ecoacoustics: a roadmap towards versatile soundscape information retrieval. *Remote Sensing in Ecology and Conservation*, *6*(3), 236–247. doi.org/10.1002/rse2.141

MacArthur, R. H., & MacArthur, J. W. (1961). On Bird Species Diversity. *Ecology*, *42*(3), 594–598.

Magurran, A. E., Baillie, S. R., Buckland, S. T., Dick, J. M. P., Elston, D. A., Scott, E. M., et al. (2010). Long-term datasets in biodiversity research and monitoring: Assessing change in ecological communities through time. *Trends in Ecology and Evolution*, *25*(10), 574–582. doi.org/10.1016/j.tree.2010.06.016

Mitrović, D., Zeppelzauer, M., & Breiteneder, C. (2010). Features for Content-Based Audio Retrieval. *Advances in Computers*, *78*(10), 71–150. doi.org/10.1016/S0065-2458(10)78003-7

Moreno-Gómez, F. N., Bartheld, J., Silva-Escobar, A. A., Briones, R., Márquez, R., & Penna, M. (2019). Evaluating acoustic indices in the Valdivian rainforest, a biodiversity hotspot in South America. *Ecological Indicators*, *103*(March), 1–8. doi.org/10.1016/j.ecolind.2019.03.024

Mullet, T. C., Gage, S. H., Morton, J. M., & Huettmann, F. (2016). Temporal and spatial

variation of a winter soundscape in south-central Alaska. *Landscape Ecology*, *31*(5), 1117–1137. doi.org/10.1007/s10980-015-0323-0

Pavan, G. (2017). Fundamentals of Soundscape Conservation. *Ecoacoustics: The Ecological Role of Sounds*, (November), 235–258. doi.org/10.1002/9781119230724.ch14

Pieretti, N., Farina, A., & Morri, D. (2011). A new methodology to infer the singing activity of an avian community: The Acoustic Complexity Index (ACI). *Ecological Indicators*, *11*(3), 868–873. doi.org/10.1016/j.ecolind.2010.11.005

Pijanowski, B. C., Farina, A., Gage, S. H., Dumyahn, S. L., & Krause, B. L. (2011). What is soundscape ecology? An introduction and overview of an emerging new science. *Landscape Ecology*, *26*(9), 1213–1232. doi.org/10.1007/s10980-011-9600-8

Pijanowski, B. C., & Brown, C. J. (2022). Grand Challenges in Acoustic Remote Sensing: Discoveries to Support a Better Understanding of Our Changing Planet. *Frontiers in Remote Sensing*, 2(January), 1–9. doi.org/10.3389/frsen.2021.824848

Quinn, C. A., Burns, P., Gill, G., Baligar, S., Snyder, R. L., Salas, L., et al. (2022). Soundscape classification with convolutional neural networks reveals temporal and geographic patterns in ecoacoustic data. *Ecological Indicators*, *138*(March), 108831. doi.org/10.1016/j.ecolind.2022.108831

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: https://www.Rproject.org/.

Rappaport, D. I., Swain, A., Fagan, W. F., Dubayah, R., & Morton, D. C. (2022). Animal

soundscapes reveal key markers of Amazon forest degradation from fire and logging. *PNAS*, *119*(18), 1–11. doi.org/10.1073/pnas.2102878119

Ross, S. R. P. J., Friedman, N. R., Yoshimura, M., Yoshida, T., Donohue, I., & Economo, E. P. (2021). Utility of acoustic indices for ecological monitoring in complex sonic environments. *Ecological Indicators*, *121*(October 2020), 107114. doi.org/10.1016/j.ecolind.2020.107114

Rychtáriková, M., & Vermeir, G. (2013). Soundscape categorization on the basis of objective acoustical parameters. *Applied Acoustics*, *74*(2), 240–247. doi.org/10.1016/j.apacoust.2011.01.004

Sánchez-Giraldo, C., Bedoya, C. L., Morán-Vásquez, R. A., Isaza, C. V., & Daza, J. M. (2020). Ecoacoustics in the rain: understanding acoustic indices under the most common geophonic source in tropical rainforests. *Remote Sensing in Ecology and Conservation*, *6*(3), 248–261. doi.org/10.1002/rse2.162

Scarpelli, M. D. A., Liquet, B., Tucker, D., Fuller, S., & Roe, P. (2021). Multi-Index Ecoacoustics Analysis for Terrestrial Soundscapes: A New Semi-Automated Approach Using Time-Series Motif Discovery and Random Forest Classification. *Frontiers in Ecology and Evolution*, 9(December), 1–14. doi.org/10.3389/fevo.2021.738537

Simpson, G. L. (2022). _gratia: Graceful ggplot-Based Graphics and Other Functions for GAMs Fitted using mgcv_. R package version 0.7.3.

Snyder, R., Clark, M., Salas, L., Schackwitz, W., Leland, D., Stephens, T., et al. (2022). The Soundscapes to Landscapes Project: Development of a Bioacoustics-Based Monitoring

Workflow with Multiple Citizen Scientist Contributions. *Citizen Science: Theory and Practice*, *7*(1), 24. doi.org/10.5334/cstp.391

Sueur, J., & Farina, A. (2015). Ecoacoustics: the Ecological Investigation and Interpretation of Environmental Sound. *Biosemiotics*, *8*(3), 493–502. doi.org/10.1007/s12304-015-9248-x

Sueur, J., Farina, A., Gasc, A., Pieretti, N., & Pavoine, S. (2014). Acoustic indices for biodiversity assessment and landscape investigation. *Acta Acustica United with Acustica*, *100*(4), 772–781. doi.org/10.3813/AAA.918757

Sueur, J., Pavoine, S., Hamerlynck, O., & Duvail, S. (2008). Rapid acoustic survey for biodiversity appraisal. *PLoS ONE*, *3*(12), 1–10. doi.org/10.1371/journal.pone.0004065

Towsey, M., Wimmer, J., Williamson, I., & Roe, P. (2014). The use of acoustic indices to determine avian species richness in audio-recordings of the environment. *Ecological Informatics*, 21, 110–119. doi.org/10.1016/j.ecoinf.2013.11.007

Villanueva-Rivera, L. J., Pijanowski, B. C., Doucette, J., & Pekin, B. (2011). A primer of acoustic analysis for landscape ecologists. *Landscape Ecology*, *26*(9), 1233–1246. doi.org/10.1007/s10980-011-9636-9

Williams, B., Lamont, T. A. C., Chapuis, L., Harding, H. R., May, E. B., Prasetya, M. E., … Simpson, S. D. (2022). Enhancing automated analysis of marine soundscapes using ecoacoustic indices and machine learning. *Ecological Indicators*, 140(January), 108986. doi.org/10.1016/j.ecolind.2022.108986

Wood, C. M., Popescu, V. D., Klinck, H., Keane, J. J., Gutiérrez, R. J., Sawyer, S. C., & Peery,

M. Z. (2019). Detecting small changes in populations at landscape scales: a bioacoustic site-occupancy framework. *Ecological Indicators*, *98*(July 2018), 492–507. doi.org/10.1016/j.ecolind.2018.11.018

Wood, S.N. (2017) Generalized Additive Models: An Introduction with R (2nd edition). Chapman and Hall/CRC.

Zeileis, A. & Hothorn, T. (2002). Diagnostic Checking in Regression Relationships. R News 2(3), 7-10. CRAN.R-project.org/doc/Rnews/

CHAPTER 4: SOUNDSCAPE MAPPING: UNDERSTANDING GEOSPATIAL PATTERNS IN SOUNDSCAPES IN THE CONTEXT OF WILDFIRE DISTURBANCE

4.1. ABSTRACT

1. Context: Soundscape ecology has provided innovative solutions for quantifying biodiversity across various landscapes. Acoustic indices can transform vast quantities of acoustics data into interpretable numeric summaries, while soundscape components can capture specific sound signals like Anthropophony (anthropogenic sound) and Biophony (animal sound).
2. Objectives: We used seven ecoacoustic metrics and acoustically-derived bird species richness across a heterogeneous landscape to investigate potential systematic spatiotemporal patterns and spatial relationships with remote sensing and environmental characteristics. We provide a case study into the ability of predicted Biophony to capture changes in biotic acoustic activity pre- and post-wildfire.
3. Methods: We assessed ecoacoustic metric relationships with environmental characteristics using a Random Forest approach that provided predictor importance and partial dependencies. The Random Forest models generated prediction surfaces across Sonoma County, CA, USA, for four target periods. Annual Biophony surfaces were also generated across the recording dataset archive (2017-2021) for three wildfires from the past five years.
4. Results: Predictive models captured landscape and diel differences in ecoacoustic patterns with varying performance (avg. $R^2 = 0.36 \pm 0.11$) depending on metric and period of day. Climate and forest structure metrics were the most important groups for

predicting seven of the eight ecoacoustic responses. We demonstrated spaceborne Lidar could contextualize forest structure with soundscape dynamics. We highlighted landscapes where acoustic indices and soundscape components may be less reliable in confidently indicating anticipated patterns. For example, the normalized difference soundscape index did not capture expected higher levels of Biophony in riparian corridors, but highlighted low to high development gradients. In contrast, low values of the acoustic entropy index corresponded with urban areas and denuded hilltops, but high values did not systematically relate to Biophony or bird species richness. We also provided evidence that Biophony may capture the reorganization of acoustic communities following wildfires – showing an upward trend 1-2 years post-wildfire, particularly in more severely burned areas.

5. Implications: These results support using ecoacoustic metrics for acoustic-diversity distribution mapping to identify spatiotemporal patterns, but multiple metrics are needed to understand the influences of ecosystem characteristics on soundscapes in heterogeneous landscapes. We highlight the importance of climate and forest structure characteristics when modeling ecoacoustic metrics and also how these characteristics can help understand Biophonic patterns following wildfire disturbance.

4.2. INTRODUCTION

Within the past decade, ecoacoustics has emerged as a practical approach to monitoring ecosystems and their biodiversity (Sueur et al., 2008; Sueur and Farina, 2015; Villanueva-Rivera et al., 2011). Ecoacoustics takes advantage of summarizing differences in the passive acoustic environment, differences due to vocalizing animals, such as birds, insects, and frogs, as well as

other natural and anthropogenic sounds, collectively referred to as a soundscape. Soundscapes and their inherent acoustic activity serve as proxies for community-level biodiversity by recording unique activity signatures in space and time (Sueur et al., 2014). For example, research advances confirm that acoustics are at least as effective as traditional point counts for detecting bird species (Celis-Murillo et al., 2012; Furnas and Callas, 2015; Towsey et al., 2014) at a maximum effective radius of approximately 50-100 m (Darras et al., 2018). However, site-based estimates of biodiversity derived from acoustic sensors correspond to a relatively small footprint, in some cases as small as a 40 m radius (Hao et al., 2021), making it challenging to create continuous and extensive maps of acoustically-informed biodiversity (Buxton et al., 2018; Lomolino et al., 2015). Such spatial context is vital in understanding patterns in biodiversity (Holgate et al., 2021). Remote sensing and landscape characteristics can provide the link to generalize discrete acoustic observations over larger, regional-scale geographic regions and understand how soundscapes vary across heterogeneous landscapes. Spatiotemporal soundscape maps can provide metrics for animal monitoring within a landscape ecology framework (Farina et al., 2011; Turner, 1989), be used to monitor anthropogenic noises in the context of urban planning (Barber et al., 2011), or even identify stressed plant conditions (Khait et al., 2019).

The spatial context of recording location is frequently incorporated in experimental design to characterize human impact (Fairbrass et al., 2017) or variation in habitat and vegetation characteristics (Boelman et al., 2007; Tucker et al., 2014). Work has strengthened the connections between environmental characteristics and acoustic diversity, represented by acoustic indices (Machado et al., 2017; Müller et al., 2022; Rodriguez et al., 2014) or derived acoustic features (Sethi et al., 2020). Early ecoacoustic work demonstrated increased forest

structure complexity (e.g., canopy cover and tree height) related positively to bird abundance and the bioacoustic index at discrete recording locations (Boelman et al., 2007). Elsewhere, the acoustic diversity index (ADI) was directly compared to and positively linked with vertical forest structure, but no study-specific relationship was made between ADI and its ability to indicate biodiversity (Pekin et al., 2012). Sethi et al. (2020) used acoustic features from deep learning models to visualize different habitat types represented by global location as a proxy for habitat type, aboveground biomass, seasonality, and diurnal patterns. Do Nascimento et al. (2020) provide a synopsis of acoustic indices used to investigate relationships with habitat and vegetation structure. However, continuous mapping of these acoustic diversity relationships and contextualization with environmental characteristics across heterogenous landscapes remains limited. Still, continuous acoustic mapping has the potential to aid in measuring acoustic-informed biodiversity at larger spatial scales than discrete site information.

Some of the earliest efforts to map soundscapes focused on urban noise pollution assessments that laid out an understanding of urban soundscape dynamics related to geospatial characteristics (de Kluijver and Stoter, 2003). In ecoacoustics, work began by integrating spatial data into the ecoacoustic framework by interpolating sounds with landscape characteristics (Matsinos et al., 2008), soundscape patterns along urban-rural gradients in the context of land use and road proximity (Joo et al., 2011), deploying recorders in distinct habitat types (Bormpoudakis et al., 2013), or relating local Landsat optical remote sensing data of land cover types to recording locations (Krause et al., 2011). Soon after these studies, a large-spatial scale effort generated a continuous acoustic diversity map using sound pressure levels across the conterminous USA (Mennitt et al., 2014). However, this mapping study was limited to

soundscape data from US National Parks and urban cores, generalizing across regions with sparse observations and interpreting patterns at the country scale. Ecoacoustic work still requires understanding the complex interactions of soundscapes and diverse landscape characteristics (Barbaro et al., 2022) to generalize correctly.

At a local scale, Mullet et al. (2016) modeled Biophony, Anthropophony, and Geophony, finding that Biophony was strongly associated with river and wetland features where recordings were primarily composed of bird vocalizations. Anthropophony was associated with roads, infrastructure, and oil/gas compressors and was present in regions where development was nonexistent due to airplane and snowmobile activity (Mullet et al., 2016). Small-scale studies have also generated predicted or interpolated layers of ecoacoustic metrics with landscape characteristics (Hao et al., 2021; Matsinos and Tsaligopoulos, 2018; Turner et al., 2018), remote sensing vegetation metrics (Benocci et al., 2022; Krause et al., 2011), and to identify biodiversity hotspots and hot moments (Holgate et al., 2021).

Not only are links between habitat and vegetation characteristics being made, but work linking disturbance or human presence is increasing in prevalence. A study found that acoustic indices aid in delineating levels of wildness defined as areas without human impacts and areas of increased ecological naturalness (Carruthers-Jones et al., 2019). Work has also highlighted the effects of logging and wildfire degradation on soundscapes that did not correlate with aboveground biomass recovery (Rappaport et al., 2020). Only recently, though, have studies begun to explore how soundscapes respond to specific disturbances like wildfires. Soundscapes have been shown to capture post-fire reorganization in soniferous communities in a space-for-time experiment in the tropics (Rappaport et al., 2020), biotic recovery about one-year post-fire

in Brazilian savannas (Duarte et al., 2021), and phenological patterns of the bioacoustic index post-fire in Arizona, USA's Chiricahua Mountains (Gasc et al., 2018). Meanwhile, in the Western US, wildfires have emerged as a significant threat to wildlife and human spaces as human development moves into the wildland-urban interface (WUI) and fires grow in size and intensity associated with drought conditions (Keeley and Syphard, 2021). Studying the response of holistic monitoring, like ecoacoustic metrics, can be a potentially valuable tool in assessing wildfire impacts. However, validation in landscapes with varying fire histories is still required to understand the complex nature of post-fire community behavior.

Ecoacoustic studies tend to focus on limited spatial scales due to the discrete nature of recording and the effort required to relate acoustic activity to meaningful biodiversity signals. Generating spatially explicit maps of biodiversity and anthropogenic activity using reliable ecoacoustic metrics can be valuable for identifying sources of noise pollution, tracking environmental disturbance, for urban planning and design to target areas that may be more sensitive to changes in noise, for human and animal wellbeing, and in quantifying cultural value. Predictive models using landscape characteristics can provide insight into what landscape characteristics are strongly associated with levels of biodiversity, highlight areas where human impact overlaps regions of rich biodiversity, and identify naturally quiet landscapes. These products become more valuable as the human population expands and landscapes change due to natural disturbances like wildfire.

Here we use a suite of climatic and remote sensing landscape characteristics to predict the distribution of ecoacoustic metrics and bird species richness across Sonoma County, California, USA, to understand (1) how these landscape characteristics influence spatiotemporal patterns of

soundscapes, (2) compare how ecoacoustic metrics vary across the county, and (3) provide insights into the sensitivity of predicted soundscapes to capture or reflect changes in Biophony due to wildfire.

4.3. METHODS

4.3.1. Study site

This study used data from the Soundscapes to Landscapes project (Snyder et al., 2022; soundscapes2landscapes.org, hereafter S2L), which focused on collecting acoustic data across the diverse habitats and landscapes of Sonoma County, California, USA (Figure 4-1; 4,152 km^2). Sonoma County is a Mediterranean coastal environment with a complex, developed, and natural landscape. The developed spaces include dense urban mesh to development along the wildland-urban interface (WUI), including residential areas, vineyards, and agriculture. Natural landscapes range from coastal shores to low mountainous terrain (up to approx. 1,369 m) with complex mixed oak-hardwood forests, riparian corridors, and Redwood conifer forests. Among these landscapes are a range of protected spaces and a persistent fire history with numerous fires primarily in the eastern areas of the County's oak-hardwood forests into the WUI that occur in late summer and fall in recent history.

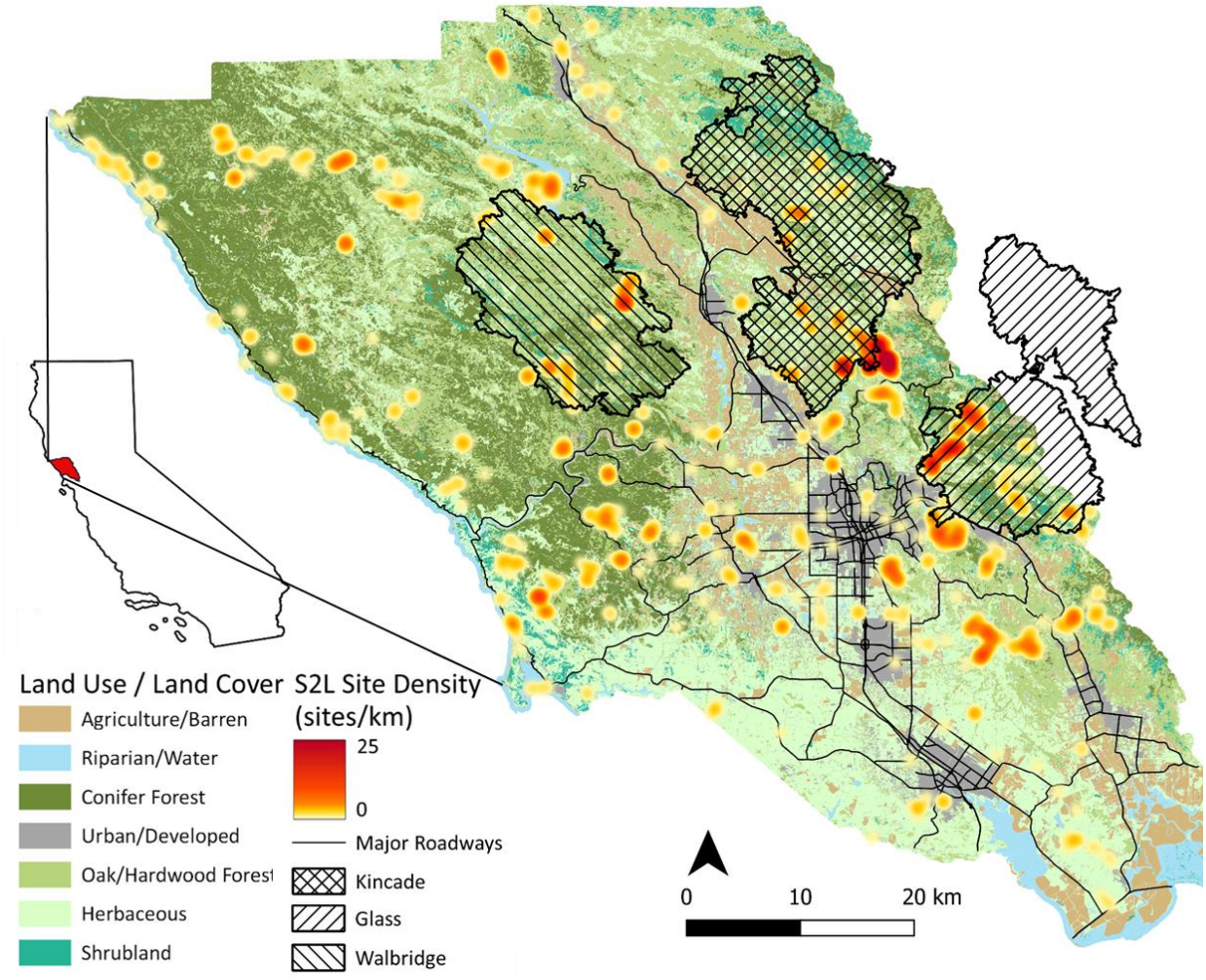


Figure 4-1. Sonoma County S2L site density for the 2017-2021 seasons. Three wildfire boundaries used in the wildfire case study are shown.

The S2L acoustic data campaigns targeted the bird-breeding season from late March to June through years 2017-2021 (Figure 4-1). At each site, autonomous recording units (ARUs) recorded 1 min every 10 min. There were two different ARU models: AudioMoths (AM) programmed with an upper-frequency range of 24 kHz (sampling rate = 48,000 Hz, digitization depth = 16-bit, signal-to-noise ratio = 44.2 dB; Hill et al., 2018; openacousticdevices.info/audio) and LG phone recorders with an upper-frequency range of 22.05 kHz (sampling rate = 44,100 Hz, digitization depth = 16-bit). ARUs were affixed at approximately 1.5 m height to the nearest

woody vegetation or post to the target site coordinates. We used 1,170 possible unique sites and 710,178 min of acoustic data for analyses, but due to some sites not containing recordings across the full 24-hr schedule, the number of sites varied across target periods: dawn = 1,166 sites, midday = 1,167 sites, dusk = 1,170 sites, and night = 1,149 sites (periods explained in section 2.2).

4.3.2. Ecoacoustic metrics

We used eight previously derived ecoacoustic metrics to represent acoustic activity diversity (Quinn et al., 2022; Quinn et al., in review) and acoustically derived bird species richness (Clark et al., 2023). The ecoacoustic metrics include the soundscape components Anthropophony, Biophony, Geophony, and Quiet (ABGQ) and the acoustic indices include acoustic complexity index (ACI) (Pieretti et al., 2011), acoustic entropy (H) (Sueur et al., 2008), and normalized difference soundscape index (NDSI) (Kasten et al., 2012). Soundscape component presence or absence was derived from a convolution neural network (CNN) at a 2-sec window (Quinn et al., 2022). Acoustic indices were calculated for each 1-min recording. ACI and H were computed in the R package *seewave* using a frequency limit of 1 to 10 kHz (flim) for ACI and window length (wl) of 512 for H. NDSI was computed using the R package *soundecology* with a fast fourier transform window of 1024 (fft_w), anthro_min of 1 kHz, anthro_max of 2 kHz, bio_min of 2 kHz, and bio_max of 10 kHz.

The acoustically derived bird species richness was derived from a mixture of three CNN experts (Clark et al., 2023). The top-performing CNN was used to classify each 2-sec audio clip for 54 possible bird species. We then used these presence and absence classifications at a site level to estimate the number of bird species at each site. If a species had more than three

presences at a site it was considered present. This was based on a sensitivity analysis using one to five presences per site to indicate an overall settling on n = 3 to account for the infrequent false positives generated by the mixture of CNNs that would result in a false presence for n = 1 or n = 2.

All ecoacoustic metrics were averaged to target periods delineated based on sun position for each site, providing a more stable estimate of soundscapes than non-aggregated features (Bradfer-Lawrence et al., 2020). These periods included: dawn chorus - one hour before and three hours after the astronomical end of the night; midday – two hours before and two hours after solar maximum; dusk chorus – three hours before and one hour after the onset of astronomical night; and night – two hours before and two hours after sun nadir, the darkest moment at night. Relative sun positions were calculated for every site and recording day using the R package *suncalc* (Thieurmel and Elmarhraoui, 2022) with the America/Pacific timezone and site longitude/latitudes in the WGS84 coordinate system.

4.3.3. Spatial predictors

We gathered a set of 123 potential spatial predictors primarily from remote sensing sources or using thematic mapping products derived initially using remote sensing data (Table 4-1). Original spatial resolution varied from 1 m to 463.83 m, and all variables were resampled using bilinear interpolation (*resample* from *terra*; Hijmans, 2022) to our base modeling resolution of 90 m. We refer to this 90 m grid as our reference grid used to standardize the spatial variables' cell alignment, extent, projection, and resolution. We used the 90 m scale to represent an assumed effective radius of the ARUs recorded (Darras et al., 2018; Shaw et al., 2022).

Select variables were also of interest to summarize at multiple scales. We used focal windows to calculate the mean and standard deviation values at medium and coarse scales using radii of 225 m and 495m, respectively. The coarser scale (990 m) was evenly divisible by 30 centered on a single pixel and represents an approximate maximum effective range of ARU recording (Krause et al., 2011), while the intermediate scale (450 m) was included to bridge the minimum and maximum ARU ranges. These multi-scale statistical summaries were done for select anthropogenic and forest structure measurements to represent the effects of distance on noise attenuation due to human density and physical barriers, respectively, and to capture larger-scale spatial patterns. We grouped variables into broad groups that describe their characteristics: anthropogenic, climate, geomorphology, phenology, and structure. These groups are used for feature selection and variable importance interpretation below.

4.3.3.1. Anthropogenic predictors

Anthropogenic variables included building density, distance to buildings, distance to streets, nighttime lights, and the proportion of area covered by agricultural land use. We calculated the density of building footprints using the *Buildings* shapefile from the County of Sonoma GIS (hub.arcgis.com/datasets/sonomacounty::buildings/about). We used the 90 m reference grid to sum the number of building footprints within each raster cell with the *count point in polygon* function in *QGIS* (v.3.20.3) and assessed varying density by then using a circular focal window and the 90 m building density raster for the medium and coarse resolutions using *focal and focalWeight* from the R package *terra* (Hijmans, 2022). Distance to buildings was calculated in *QGIS* with the *raster proximity* tool using the same *Buildings* shapefile from the Sonoma County GIS. Distance to roads is a source of Anthropophony from traffic (Mullet et al., 2016; Quinn et

al., 2022; Ware et al., 2015); we calculated distance to roads using the *Streets* shapefile from the Sonoma County GIS (accessed January-2023; Merrick & Company Project Number 50013734) in *QGIS* after filtering out non-developed roads (CLASS: Unimproved, Fire Road, Private (Surface Type = Dirt), Path Multi-Use, Path Pedestrian). Nighttime light data were included as a source of human impact using the Suomi National Polar-orbiting Partnership Visible Infrared Imaging Radiometer Suite (VIIRS) Day-Night Band sensor's Stray Light Corrected monthly composite (Mills et al., 2013). We computed mean annual nighttime radiance for the 2017-2021 field seasons using the original 463.83 m spatial resolution on Google Earth Engine (Gorelick et al., 2017) exported at 90 m. To calculate the percent agricultural cover, we used the Sonoma County Fine-scale Vegetation and Habitat Map (sonomavegmap.org) filtered to agricultural land cover types in *QGIS*. Because of the fine spatial resolution of the LULC map (1 m), we simplified polygons with a 5 m tolerance during simplification for more efficient computation. We derived the percent of the shapefile area intersecting each 90 m reference grid cell. We calculated the medium and coarse metrics for percent agriculture as agricultural machinery in less structurally complex environments may result in longer distance attenuation of point sources of Anthropophony. Ten anthropogenic variables were created.

4.3.3.2. Climate predictors

Climate variables were derived from the California Basin Characterization Model (BCM) monthly climate data from 2017-2021 at a spatial resolution of 270 m (Flint et al., 2013). We aggregated data into corresponding water year quarters: October – December (Q1), January – March (Q2), April – June (Q3), and July – September (Q4), resulting in four sets of six predictors for a total of 24 predictors per water year summarizing the included S2L campaign.

Each quarter's climate predictors included precipitation (PPT), average minimum temperature (TMN), average maximum temperature (TMX), potential evapotranspiration (PET), actual evapotranspiration (AET), and climatic water deficit (CWD).

4.3.3.3. Geomorphology predictors

Topographic characteristics are commonly associated with species distribution models (Hao et al., 2019), so we included a set of six geomorphological variables derived from a 1 m airborne lidar digital elevation model (DEM) resampled to 30 m. The 30 m DEM was used to generate slope and topographic position index, both continuous and classified, using *Topography Tools* for *ArcGIS 10.3*. In Sonoma, distance to the coast is correlated with the spatial pattern of habitat types and was calculated using the *Euclidian Distance* tool in *ArcGIS Pro*. Distance to streams was derived with *Euclidean Distance* from the Sonoma County lidar-derived stream centerlines, filtered to 1st-order streams (geomorphological data: sonomavegmap.org).

4.3.3.4. Phenology predictors

We included phenology to represent dynamics in vegetation productivity across each year (Radeloff et al., 2019). Phenometrics were derived using Framework for Operational Radiometric Correction for Environmental monitoring (FORCE), which aids in processing medium-resolution multispectral satellite imagery (Frantz, 2019). We used FORCE to perform atmospheric correction of available Landsat 7 and 8 and Sentinel 2A and B imagery from January 1, 2016 to February 1, 2022. We then used FORCE Time Series Analysis (TSA) to stack Normalized Difference Vegetation Index (NDVI) data and perform Radial Basis Function interpolation to standardize the data to an 8-day time sequence. These time-series NDVI data were then used to calculate seven annual phenology variables spanning a water year (October

before the calendar year to September of the calendar year). Annual metrics were gridded at 30 m and included dynamic habitat indices (DHI) cumulative, minimum, and variance (Hobi et al., 2017) and four NDVI metrics (5th, 50th, and 95th percentiles, and seasonal difference between mean of June and mean of November data; Burns et al., 2020).

4.3.3.5. Forest structure predictors

We incorporated forest structure variables based on data from the Global Ecosystem Dynamics Investigation (GEDI) Lidar sensor aboard the International Space Station (ISS) (Dubayah et al., 2020). GEDI is a sampling sensor where measurements are collected in discrete, ~ 25 m diameter "shots" along the ISS path resulting in non-continuous data with gaps between samples. Because our modeling approach utilizes continuous data to generate mapping products, we developed continuous interpolations of forest structure. Furthermore, because GEDI was launched in December 2018 and creating interpolated products required repeat observations, we created a time-series relationship using select spectral bands from the Landsat Collection 2 surface reflectance imagery and continuous change detection and classification (CCDC; Zhu and Woodcock, 2014). This method related quality GEDI shots determined using quality and degrade filters from the GEDI shots between April 17 2019 and September 30 2021 and days of year 105 to 243 to focus on leaf-on conditions. These quality GEDI shots were collocated with Sonoma airborne lidar data from 2012 to improve geolocation accuracy and uploaded to Google Earth Engine (GEE) for time-series analysis. We paired GEDI shots with Landsat surface reflectance time-series modeled harmonic coefficients from the CCDC, from which we used the first order harmonic coefficient to fit a Random Forest classifier (RF) to predict structure metrics (Breiman, 2001). GEDI metrics were generated at 30 m resolution every June 15th during the S2L recording

campaign from 2017 to 2021. Further information on forest structure methods can be found in Appendix 4A. We then generated the medium and coarse mean and standard deviation values to represent horizontal structural heterogeneity and bilinearly resampled all structure layers to 90 m.

We derived the multi-scale variables to serve as proxies for physical limitations on sound propagation. This was based on the assumption that an area with a denser surrounding canopy will have a reduced effective recording radius compared to an open-canopy structural setting like in grasslands, meadows, and some conifer forests (Do Nascimento et al., 2020; Pekin et al., 2012), although evidence counter to this also exists (Darras et al., 2016). This resulted in 15 unique annual GEDI metrics (relative height at 5^{th}, 25^{th}, 50^{th}, 75^{th}, 95^{th}, 98^{th}, and 100^{th} percentiles, percent tree cover, plant area index (PAI), foliage height diversity (FHD) of PAI at 1 m height, plant area volume density (PAVD) from 0 to 5 m and 5 to 10 m, and the vertical distribution ratio (VDR) of the bottom, middle, and top thirds of the vertical profile) with 60 additional mean and standard deviation layers totaling 75 structure metrics per year.

Table 4-1. Predictor variables for geospatial modeling with predictor category, units, description and data source.

Category	Predictor	Units	Description
Anthropogenic	Building density† [1]	Count	Number of building footprints
	Distance to buildings [1]	m	Distance to nearest building footprint
	Distance to streets [2]	m	Distance to developed streets
	Percent ag. LULC† [3]	%	The area of agricultural LULC within each 90 m cell
	VIIRS Night lights [4]	$nW/cm^2/sr$	Summarizes mean annual upward radiance at night
Climate	AET [5]	mm	Quarterly water year average actual evapotranspiration (AET)
	CWD [5]	mm	Quarterly water year climatic water deficit (CWD); PET - AET
	PET [5]	mm	Quarterly water year potential evapotranspiration (PET)

	PPT [5]	mm	Quarterly water year precipitation (PPT)
	TMN [5]	°C	Quarterly water year minimum temperature (TMN)
	TMX [5]	°C	Quarterly water year maximum temperature (TMX)
Geomorphology	Distance to coast [3]	m	Distance to the Sonoma County coastline (west boundary of county)
	Distance to streams [3]	m	Distance to the centerline of 1st order streams
	Elevation [3]	m	Digital elevation model, meters above sea level
	Slope [3]	Degrees	Rate of change of elevation
	TPI [3]	Continuous	Unclassified topographic positon index (TPI) value
	TPI classified [3]	Categorical	TPI categorized into flat, ridge, upper hillslope, middle hillslope, lower hillslope, and valley (6 classes)
Phenology	DHI cumulative [6]	--	Dynamic Habitat Index (DHI); The area under the annual phenological curve
	DHI minimum [6]	--	The minimum value of the annual phenological curve
	DHI variance [6]	--	The coefficient of variation of the annual phenological curve
	NDVI 5th percentile [6]	--	The 5th percentile value of Normalized Difference Vegetation Index (NDVI) using FORCE annual phenological time-series
	NDVI 50th percentile [6]	--	The 50th percentile value of NDVI using FORCE annual phenological time-series
	NDVI 95th percentile [6]	--	The 95th percentile value of NDVI using FORCE annual phenological time-series
	NDVI season difference [6]	--	The difference in mean June and mean November NDVI values using FORCE annual phenological time-series
Structure	RH 5%† [7]	m	Relative height (RH) at the 5th percentile of the returned energy profile (REP)
	RH 25%† [7]	m	RH at the 25th percentile of the REP
	RH 50%† [7]	m	RH at the 50th percentile of the REP
	RH 75%† [7]	m	RH at the 75th percentile of the REP
	RH 95%† [7]	m	RH at the 95th percentile of the REP
	RH 98%† [7]	m	RH at the 98th percentile of the REP
	RH 100%† [7]	m	RH at the 100th percentile of the REP
	Cover† [7]	%	Proportion of unit with canopy present when considered vertically

PAI† [7]	m^2/ m^2	Plant area index (PAI)
FHD† [7]	--	Foliage height diversity (FHD)
PAVD 0-5† [7]	m^2/ m^3	Areal density of vegetation from 0 to 5 m
PAVD 5-10† [7]	m^2/ m^3	Areal density of vegetation from 5 to 10 m
VDR bottom† [7]	--	Vertical distribution ratio (VDR) of the bottom third of the REP
VDR middle† [7]	--	VDR of the middle third of the REP
VDR top† [7]	--	VDR of the top third of the REP

† Predictor spatial statistics mean and standard deviation for medium (radius = 225 m) and coarse (radius = 495 m) scales calculated. This resulted in four additional variants of each predictor.
[1] Buildings: hub.arcgis.com/datasets/sonomacounty::buildings/about
[2] Streets: Merrick & Company Project Number 50013734
[3] LULC, Coast, Streams, DEM: sonomavegmap.org
[4] VIIRS: (Mills et al., 2013)
[5] BCM: (Flint et al., 2013)
[6] Phenology: (Frantz, 2019; Radeloff et al., 2019)
[7] GEDI: (Dubayah et al., 2020)

4.3.3.6. Site extraction of predictor data

We extracted co-occurring spatial predictor data using the coordinate pair of each site and the underlying 90 m cell values. Annual predictors (structure, climate, phenology, and nighttime light) were temporally matched and extracted for the year data collection at a site occurred, while all other predictors (anthropogenic excluding nighttime lights and geomorphology) were extracted without temporally matching.

4.3.4. Spatial modeling

We implemented RF regression models, a non-parametric approach that can predict the non-normally distributed ecoacoustic metrics while identifying important predictors and accounting for interactions. RF has also been successfully applied for species distribution modeling (Lorena et al., 2011) and geospatial sound level modeling (Mennitt et al., 2014). RF models allowed us to (1) understand prediction performance for each ecoacoustic metric, (2) investigate what remote sensing predictors were most important in explaining ecoacoustic metrics, (3) interpret the

effects of top-performing predictors, and (4) spatially model ecoacoustic metrics across periods of the day. RF models were trained, and products were generated (e.g., variable importance, variable effects, and prediction surfaces) on Northern Arizona University's high-performance computing cluster, Monsoon, while all other analyses were completed locally using the R environment (R Core Team, 2022).

4.3.5. Variable selection

When using RF models, variable importance interpretation is sensitive to highly correlated predictors in the dataset (e.g., Gregorutti et al., 2017). A range of methods exist for RF variable selection, but agreement on efficacy varies depending on the analysis (Speiser et al., 2019), and computational efficiency should be considered when not testing these methods explicitly. Because of this, we implemented variance inflation factoring (VIF) with study-specific adjustments, a commonly used feature selection method (Zuur et al., 2010). Predictor selection consisted of three steps for each response to reduce the potential set of 123 predictors:

1. We believe the non-informed selection of predictors by VIF that eliminates highly correlated predictors without expert opinion introduces potential difficulty in interpretation. Due to this, but still considering the effort to model eight responses with 123 predictors, we used a semi-informed method where we first built univariate RF regressors using *caret's ranger* RF implementation for each response and predictor combination. The R package *caret* enables machine learning pipelines for data splitting, feature selection, model tuning, and model evaluation (Kuhn, 2022). *Caret* allows for 238 unique model implementations, and *ranger* RF is a fast implementation of random forests that is computationally efficient at handling high-dimension data (Wright and Ziegler, 2017),

allowing for tuning of 9,840 univariate RF models in this step. Site average response values were used to keep by-response predictor groups consistent, making variable importance comparable across period-of-day models. We calculated mean test R^2 after k = 10 cross-validation folds to estimate model fit and tuned the minimum nodes in each RF fold while the predictor was centered and scaled.

2. The univariate RF R^2 values were then used to rank predictors within their respective groupings (anthropogenic, climate, geomorphology, phenology, and structure) to reduce within-group redundancy. We used *spatialRF's auto_vif* function (Benito, 2021), which ranks predictors to prioritize which predictors should be preferentially retained. We viewed this as advantageous because standard VIF procedures do not integrate the predictor's relationship to the response, allowing for a semi-informed selection of predictors. We used a cutoff threshold of 3, recommended when the ecological signal may be weak or highly variable (Zuur et al., 2010).
3. Finally, we used the within-group thinned predictors and again used the univariate RF R^2 to rank the remaining predictors with a more conservative threshold cutoff of 10 in the *auto_vif* function. This step removed any highly-correlated predictors affecting variable importance interpretations and was the final predictor set for each response.

4.3.6. Cross-validation (CV)

We implemented a geographically informed cross-fold validation (referred to here as geo-CV) method to evaluate final model performance and assess model stability. Geo-CV adds to the strength of k-fold CV, during which data is separated into k groups where one group is withheld to evaluate model performance as the test set, and the remaining k-1 groups are used for model

training. However, when modeling involves spatial data, geographic context is vital to incorporate when assessing model performance due to issues with spatial autocorrelation (Roberts et al., 2017). We found that more rigid approaches like spatial blocking, where a fixed resolution grid is overlaid across the study domain, resulted in highly uneven fold membership and did not reflect the original S2L sampling regime. This weakness led us to develop a custom clustering approach using *ArcGIS's find point clusters* function with a distance of 500 m from site to site to determine a cluster. We chose 500 m matching our coarse scale focal window for predictors (495 m), a distance that most acoustic attenuation would be minimal, and recordings would have a lower likelihood of overlapping (Haupert et al., 2022). The clusters were then used in a grouped CV approach using *caret's groupKFold* function (Kuhn, 2022) and k = 10 to generate the geo-CVs. We then repeated the grouped k-fold splitting ten times using randomly generated starting seed values to randomize site assignment to geo-CV folds. The random seed and k = 10 approach resulted in 100 unique folds to train the RF models. Even with the more flexible clustering approach, the folds were not even because clusters varied in the number of constituent sites, so the training set size ranged from 932 (79.7%) to 1,129 (96.5%) sites among folds.

4.3.7. Modeling

We used RF regression models implemented with *caret's ranger* RF (Kuhn, 2022; Wright and Ziegler, 2017) to predict ecoacoustic metrics using VIF-selected predictor sets. *Caret's* training protocol allows for utilizing a tuning grid to optimize the RF hyperparameters *mtry* and *min.node.size*. The number of trees was fixed at 500, predictors were centered and scaled, and hyperparameter selection was evaluated internally using a repeated-CV approach of 5 folds with

3 repeats. We assessed variable importance with the permutation option. The repeated geo-CV setup for each period and response resulted in 3,200 final models, a product of 4 periods, 8 responses, 10 geo-CV folds, and 10 repeats. Each unique period and response combination had 100 final RF models used for evaluation. Period-of-day models were used based on the assumption that sound-level models including the time of day as a predictor performed worse than the individual time-of-day models (Mennitt and Fristrup, 2016). This structure reduced the effect of temporal autocorrelation if the time of day was treated as a predictor (Roberts et al., 2017). However, we did not explicitly test temporal assumptions here.

4.3.8. Model evaluation

We used R^2 and root mean square error (RMSE) values of the withheld test sets in each geo-CV fold (n = 100) to assess model performance.

Variable importance was calculated using permutation importance for each RF model to assess the contribution of each predictor. When a predictor is permuted, the deterioration in model performance (i.e., increase in error) is measured. The larger the decrease in performance, the more important the predictor is in the model. We summarized each raw predictor permutation value across folds using mean and standard deviation to summarize variable importance. The average global permutation values were then ranked to assess the overall importance of each response's predictor set. To determine importance, we (1) identified the top five predictors that were consistently the most important and (2) summarized predictor permutation values by group (anthropogenic, climate, geomorphology, phenology, and structure) to understand the relative importance of each predictor group to ecoacoustic metrics.

To interpret the effects of variable importance on responses, we generated partial dependency plots (PDPs) for each response's top five predictors. Visual aids such as PDPs highlight relationships between predictors and responses that are otherwise difficult to interpret due to RF models' non-parametric, complex nature (Maxwell et al., 2018). We used the *partial* function from the R package *pdp* (Greenwell, 2017) by modeling the predicted partial response for each predictor from 100 equally spaced observations along the predictor's training domain.

Prediction surfaces allow for the extension of discrete, site-based observations to a spatial surface. We used the 90 m, 2021 spatial predictors to predict the spatial patterns of period and response using *raster's predict* function (Wright and Ziegler, 2017). Because of our repeated geo-CV method, we created an average behavior surface using the median value and assessed by-pixel variation using the interquartile range (IQR) of the 100 final models for each period and response combination. These prediction maps allowed us to spatially and temporally compare the patterns across responses and times of the day.

4.3.9. Soundscape recovery following wildfires

We assessed how Biophony changed following wildfires using predicted surfaces. To do this, we averaged period-of-day surfaces for 2017 through 2021 and targeted three wildfires that did not experience extensive reburn during the recording period or recently. We targeted the Kincade Fire, the Glass Fire, and the Walbridge Fire (Figure 4-1). The Kincade Fire began in October 2019, burning 77,758 acres in the eastern part of Sonoma in mixed hardwood, conifer, shrubland, and grass terrain that overlapped in the south with the 2017 Tubbs fire, the north with the 2017 Pocket fire, and the Geysers Fire in the north in 2004. The Glass fire began in September 2020, burning 67,484 acres (26,876 acres in Sonoma) in complex landscapes dominated by hardwoods

and conifer forests with smaller shrub and grasslands, agriculture, and urban land cover areas overlapping in the south with the 2017 Nuns fire. The Walbridge Fire burned 55,209 acres within the LNU Lightning Complex in August 2020 in a complex rugged hill system in the Austin Creek State Recreation Area in Sonoma and had a historically longer reburn interval, with the most recent historic fire occurring between 1944. The Walbridge fire land cover is similar to the Glass Fire, with the southern portion of the fire dominated by conifer forest and the northern portion of the fire mixing with hardwood forest.

We used Monitoring Trends in Burn Severity (MTBS) classified difference normalized burn ratio (dNBR) fire severity maps created using bi-temporal Landsat data targeting differences in vegetation before and after wildfire (Eidenshink et al., 2007) to characterize the burn severity. We rescaled the classified severity maps from 30 m to 90 m using a majority vote to identify burn severity classes considering unburned/low, low, moderate, and high severity areas. We also generated a non-burned buffer using a 1 km ring located 1 km from the fire perimeter to assess anticipated year-to-year Biophony patterns (i.e., a "pseudo-control"). We removed mentioned recent historic fires from the target fire areas and pseudo-control rings to avoid recent reburn signals (Kincade: Geysers, Pocket, and Tubbs; Glass: Tubbs and Nuns). Predicted Biophony was extracted for each fire area and non-burned buffer area and related to the underlying burn severity class to compare the average acoustic biotic behavior following wildfire. We did not assess other ecoacoustic metrics such as Anthropophony or Quiet as these may relate more to the change in forest structure and therefore increased sound attenuation following severe burns, while Biophony includes this effect as well as the potential recovery of soniferous species such as insects and birds (Duarte et al., 2021; Gasc et al., 2018).

4.4. RESULTS

4.4.1. Modeling performance

RF models predicted Anthropophony with the best performance across periods (mean and sd $R^2 = 0.55 \pm 0.13$), particularly for the dawn period ($R^2 = 0.67 \pm 0.09$; Figure 4-2). Overall, dawn had the best performance among responses ($R^2 = 0.41 \pm 0.13$).

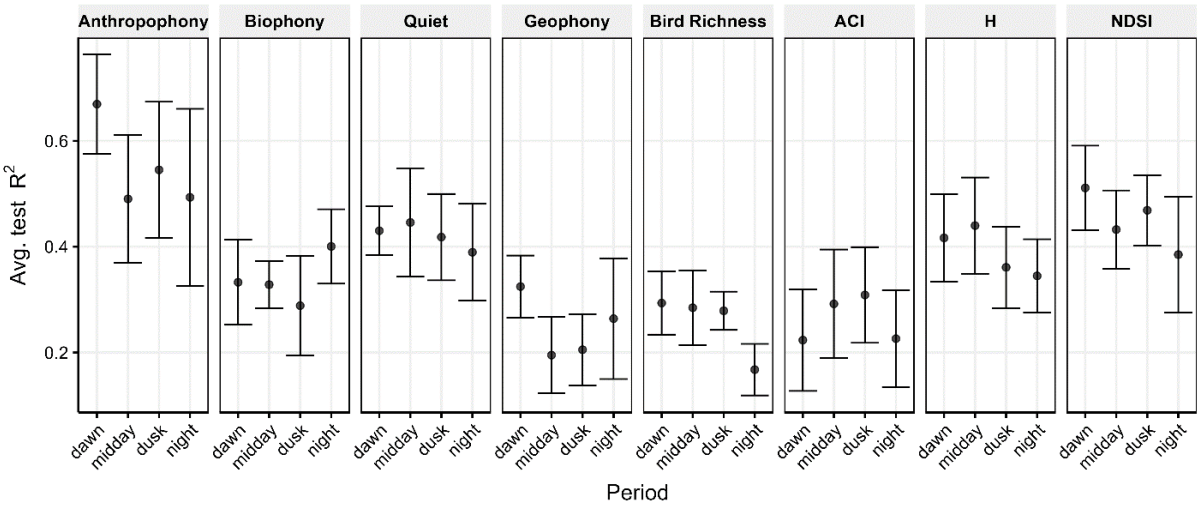


Figure 4-2. RF model performance is summarized as the test R^2 for the 100 final models for each period of the day. Center points represent the mean R^2 with one standard deviation whiskers.

4.4.2. Predictor importance

The predictor selection method reduced our predictor datasets from 123 predictors to 26 for Anthropophony, 27 for Biophony, 26 for Geophony, 15 for Quiet, 27 for bird species richness, 22 for ACI, 27 for H, and 26 for NDSI (list of predictors in Appendix 4A). Two predictors were shared among the eight responses (slope and VIIRS average nighttime radiance), while 27 predictors were unique to only one response, and 51 predictors were not selected for any model. Of the 51 non-included predictors, 49 were structural metrics, one was anthropogenic (90m

building density), and one was phenological (NDVI median). The reduction in structural metrics was due to the higher proportional representation in the original predictor set, primarily due to the multiple spatial summary variants for each unique metric leading to high levels of correlation. The only structural metrics not represented in any final predictor set (i.e., all five summary variants were dropped) were PAI and RH98; however, RH95, RH98, and RH100 are extremely highly correlated and all approximate canopy height (e.g., 90m average $\rho = 0.999$).

Overall we observed climate predictors most frequently represented in the top five most important predictors across responses with precipitation of quarter one in five of eight response top five predictors (Table 4-2). Structure was the second most represented predictor group occurring in six responses. In only two responses, anthropogenic (Anthropophony and NDSI) and phenology (ACI and Biophony) were included as top predictors least often.

Table 4-2. Top five predictor importances from average permutation importance (per period = 100 models, summary = 400 models).

Predictor rank	1	2	3	4	5
Anthropophony	Nighttime lights	PAVD 5-10m$_{990\ m\ mean}$	RH 25%$_{990m\ sd}$	Slope	PPT Q1
Biophony	RH 95%$_{450m\ mean}$	DHI$_{cumulative}$	DHI$_{variance}$	TMX Q3	TMN Q1
Geophony	PET Q4	TMX Q2	TPI	DEM	RH 25%$_{450m\ sd}$
Quiet	VDR Bottom$_{450m\ mean}$	Elevation	PPT Q1	PET Q3	PPT Q2
Bird Spp. Richness	TMX Q1	Nighttime lights	Distance from the coast	PPT Q2	PPT Q1
ACI	DHI$_{minimum}$	RH 75%$_{990m\ mean}$	Distance from the coast	PPT Q2	Elevation
H	PPT Q2	PPT Q1	RH 50%$_{990m\ mean}$	PAVD 0-5m$_{990m\ mean}$	Cover
NDSI	Elevation	PPT Q3	Nighttime lights	PPT Q1	TMX Q4

The average importance of predictor groups highlighted differences in the expected prevalence of predictors if predictor groups had an equal influence on responses (Figure 4-3). Bird species richness showed a notably higher prevalence of climate predictors in the final

models and a decrease in all other predictor groups. In general, climate (Biophony, Quiet, Geophony, bird richness, and NDSI), structure (Anthropophony and H), and geomorphology (ACI) groups had the highest prevalence of predictors in the top five model predictors for each response. Anthropogenic predictors ranked in the top five less frequently for all responses except for Anthropophony, while geomorphology and phenology predictors displayed a decrease in prevalence for three and four responses, respectively. Quiet showed the most similarity in prevalence between the final VIF predictor set and the prevalence of the top five predictors.

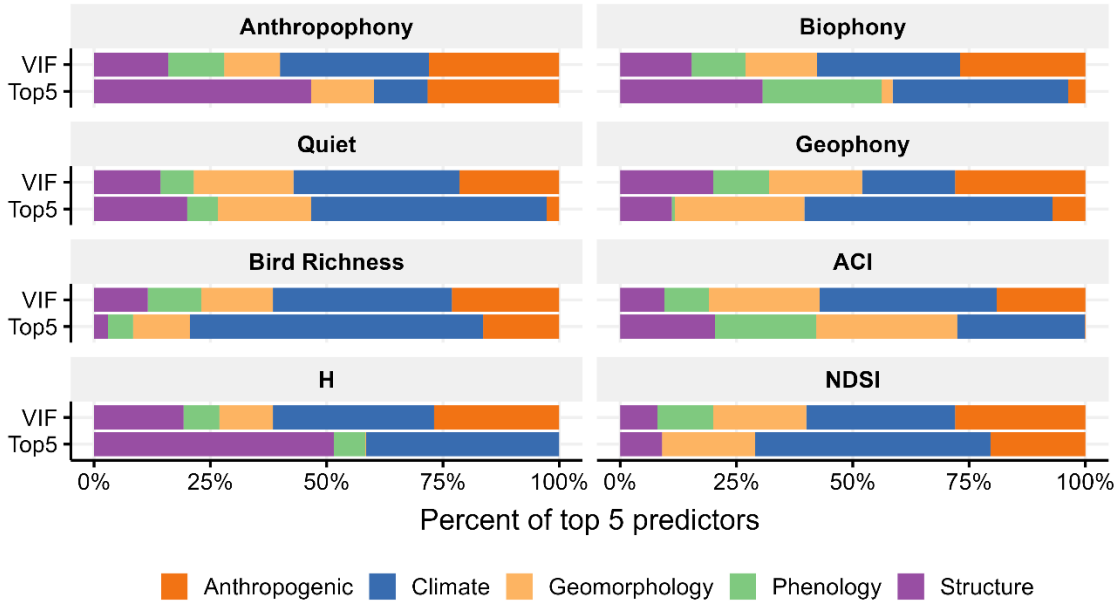


Figure 4-3. Prevalence of predictor groups in top five ranked predictors (n = 400 model folds per response) compared to anticipated predictor prevalence in VIF final predictor set if predictors were equally influential on responses.

4.4.3. Predictor Effects

We used the overall top five rankings to generate PDPs for each response period to interpret the relationship between predictors and responses (Figure 4-4). PDPs revealed that most periods

had a consistent effect but in select cases deviated, particularly for Biophony's DHI predictor effects.

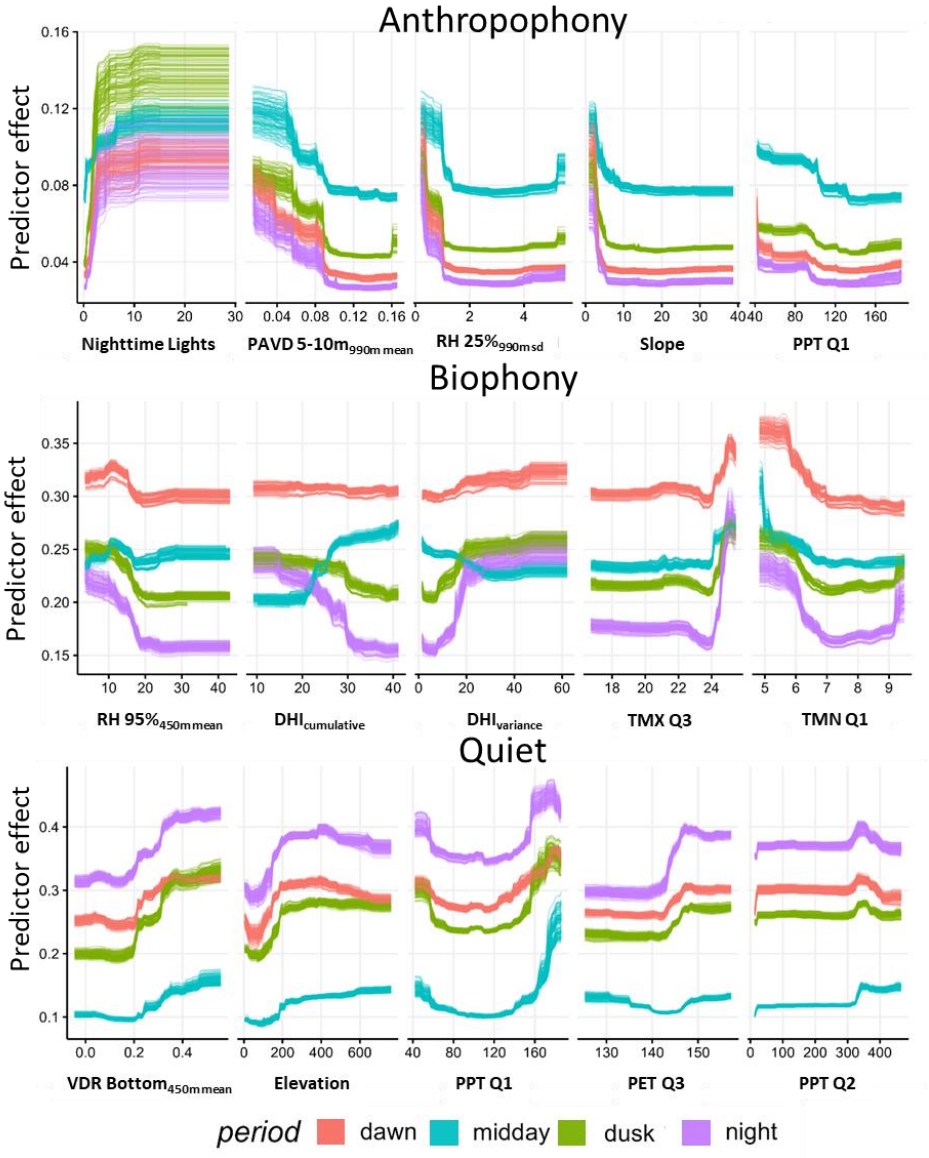


Figure 4-4. Partial dependency plots (PDP) for three responses. Predictors shown represent the top five predictors for each response, on average (i.e., across periods of day).

4.4.4. Spatiotemporal patterns

We found the RF models captured spatial variations in all ecoacoustic metrics. Generally, the anthropogenic impact was concentrated in the urban corridors that travel north to south across Sonoma County. Biotic-related acoustic activity peaked at the interface of urban-natural spaces (Figures 4-5, 4-6, and 4-7).

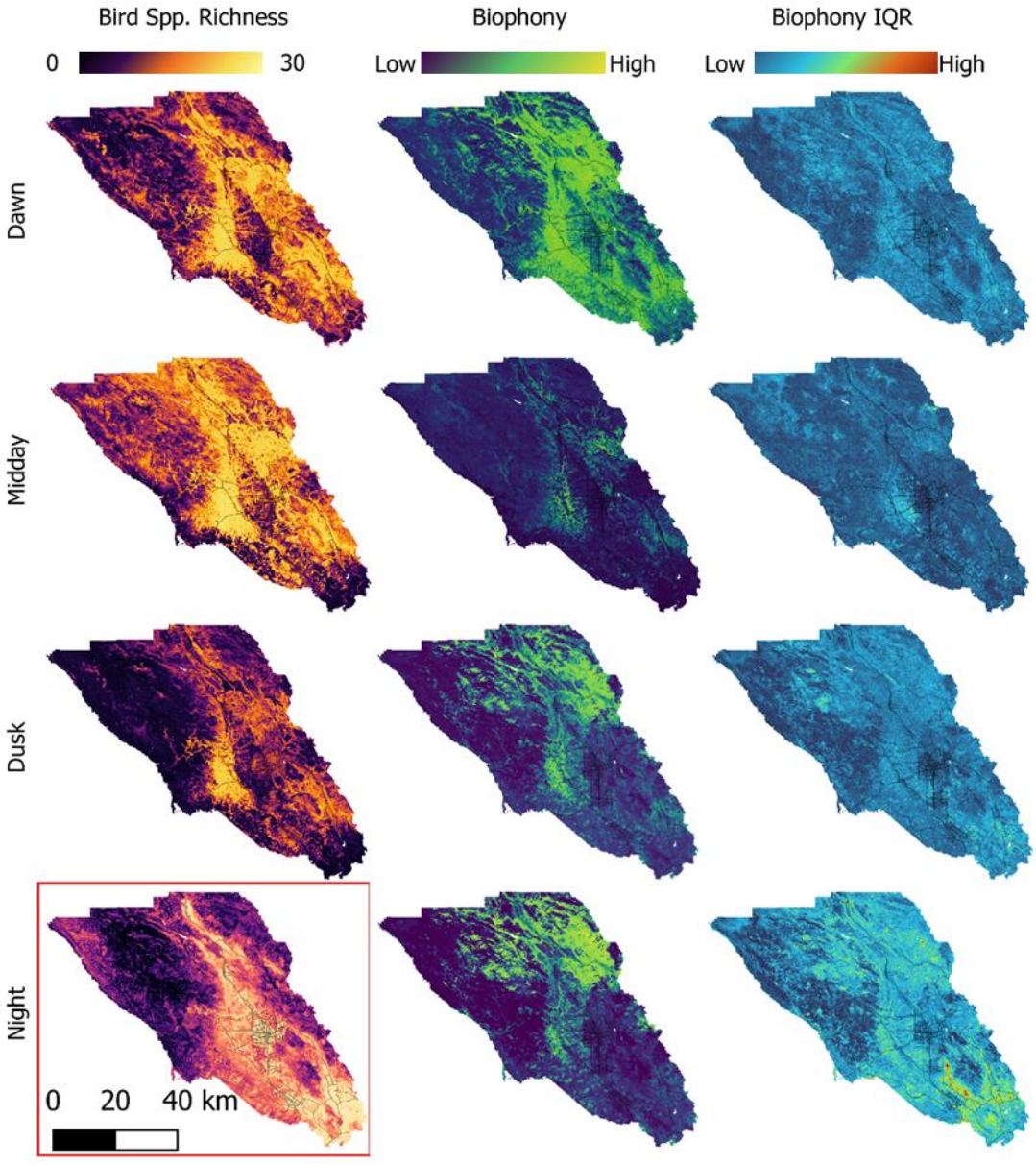


Figure 4-5. Bird species richness, Biophony, and variation in Biophony (IQR = interquartile range) patterns for periods across Sonoma County, showing highway, freeway, and primary arterial roads in light grey. The red-outlined bird species richness at night map was rescaled (low = 0, high = 10) because values were too low to use the same visualization as the other periods. Scales are normalized within each response where minimum = 0 and maximum = lowest high predicted value across periods (richness = 30, Biophony = 0.65, Biophony IQR = 0.1).

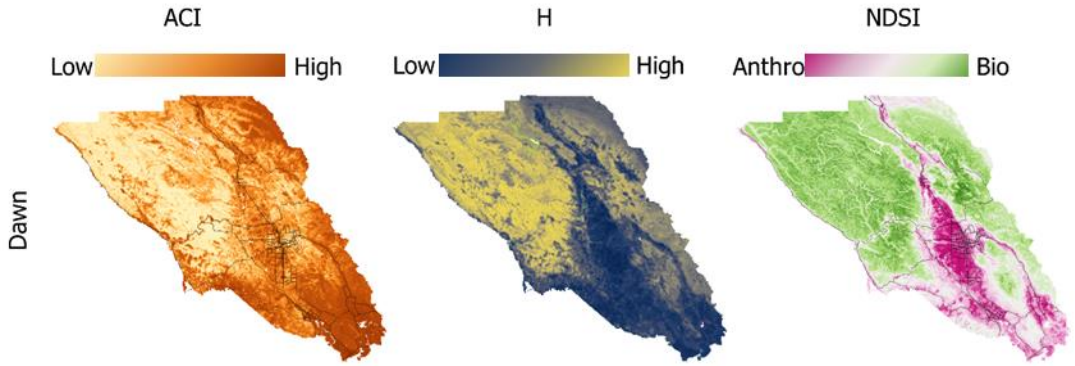


Figure 4-6. ACI, H, and NDSI patterns for dawn across Sonoma County, showing highway, freeway, and primary arterial roads in light grey. Variability across periods was minimal for acoustic indices, and midday, dusk, and night are not shown here. Scales are normalized within each response where minimum = highest minimum predicted value (ACI = 57, H = 0.77, NDSI = -0.4) and maximum = lowest high predicted value across periods (ACI = 77, H = 0.97, NDSI = 0.6).

High levels of Biophony appear along riparian zones, in oak-hardwood forests, and along transitions from urban or agriculture to woodlands (i.e., wildland-urban interfaces; Figure 4-5). Temporally, Biophony follows anticipated patterns with widespread high levels in the dawn period and concentrated higher levels at dusk and night, while levels were lowest across the county for the midday period. We included the variation in predicted Biophony to demonstrate our error associated with prediction surfaces (Figure 4-5 – Biophony IQR). Night had the highest variation, particularly in southern riparian, wetland, and agrarian-adjacent areas. In contrast, variation was lower for other periods across most of the County, with locally high variation in grassland areas.

The largest shift in countywide Quiet was from low midday levels (i.e., large amounts of sound) to high levels at night (Figure 4-7). This trend is most notable in forested land cover areas, with the Quietest regions occurring in the western, conifer-dominated forests.

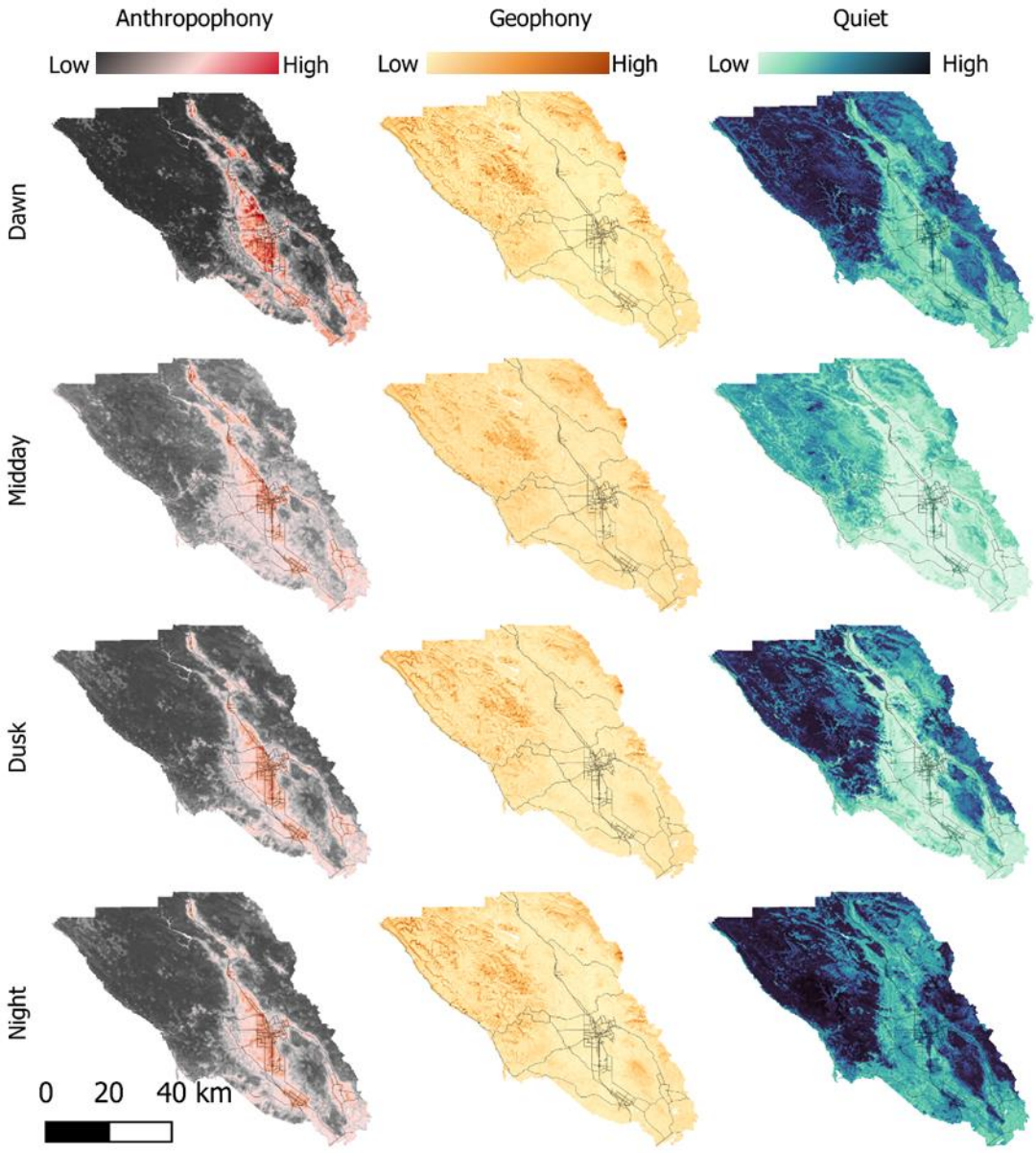


Figure 4-7. Anthropophony, Geophony, and Quiet patterns for periods across Sonoma County, showing highway, freeway, and primary arterial roads in light grey. Scales are normalized within each response where minimum = highest minimum predicted value (all = 0) and

maximum = lowest high predicted value across periods (Anthropophony = 0.4, Geophony = 0.45, Quiet = 0.8).

Spatially, we found Anthropophony follows expectations concentrated in higher industrial and commercial development areas, urban residential spaces, and along major roadways (Figure 4-7). However, these high-Anthropophony areas had temporal patterns that did not match expectations. The dawn urban cores showed high levels of Anthropophony, while we anticipated midday and dusk to be the highest. These concentrated, high Anthropophony areas were primarily agrarian and adjacent to primary roadways. Midday and dusk Anthropophony, though slightly lower in magnitude than the concentrated dawn levels, had a more considerable spatial extent into less developed urban/suburban areas and along more minor roadways, particularly during the midday period extending westward from the center of the county.

Bird species richness had the highest predicted values (richness = 33) during the midday period, while the anticipated highest period, dawn, had a slightly lower peak (richness = 30) and similar overall patterns across the county (Figure 4-5). The highest areas of richness occurred in residential neighborhoods, similar to patterns in Biophony, but also in riparian corridors and prominently along the wildlife corridors that span Pepperwood and Modini Preserves on the eastern side of the county that roughly overlap with areas burned by the Kincade fire (Figure 4-1).

NDSI levels reflect the general expected pattern that Anthropophony drives lower values towards -1 while Biophony drives values towards 1, and a balance of the soundscape sources results in values around 0 (Figure 4-6). Lower values highlight similar urban cores as described for Anthropophony's patterns; however, riparian corridors, even in low anthropogenic impact

regions, have some of the lowest values implying high Anthropophony and low Biophony. This may be due to the acoustic similarity, particularly in the S2L dataset, for Geophony in the 0-2 kHz frequency band, the frequency range which is used to represent Anthropophony as well.

ACI had patterns that diverged from other ecoacoustic metrics in that the southern portion of the county, the urban cores, some riparian corridors, and some hilltops to the northeast were highest for midday (Figure 4-6). Night had the lowest average values, but high ACI reflected select agricultural areas and hilltops to the northeast. Dawn and dusk had similar patterns and were midpoints between the extreme night and midday patterns.

H had the highest levels across the county for the nighttime and dawn, though values in urban cores were lower at dawn. Values were lowest at midday and rose from dusk to night (Figure 4-6). The lowest values coincided with urban areas, while the highest values were in the mosaicked oak-woodland regions in the east and more so in the conifer forests of the western part of the county. Local regions of lower values, comparable to urban areas, were interspersed in the conifer forests on higher relief, bare-vegetation hills, and ridgelines.

The highest levels of Biophony/bird species richness and Anthropophony correlation were in riparian corridors and urban-adjacent spaces, where anthropogenic effects propagate to higher Biophonic-rich areas. Biophony and bird species richness levels were lower in select recent wildfire regions such as the Nuns (2017) and Glass (2020) fires. Bird diversity was negatively related to H in many regions, though there were local areas of high correlation in the suburban/residential regions where high bird diversity and high values in H co-occurred.

4.4.5. Soundscape recovery following wildfires

Wildfire soundscape analyses revealed there might be detectable deviations in Biophony levels following wildfires (Figure 4-8). Because we do not explicitly model sound propagation, having a pure pre- and post-fire biotic recovery signal is complex. However, our models predicted values for wildfire-impacted spaces caused by structural changes, topography, phenology, human impact, and climate with soundscapes, not explicitly disturbance.

The S2L recording dataset covers one (Walbridge and Glass - 2020) to two (Kincade - 2019) years following wildfire disturbance here. We show dusk patterns in Figure 4-8, highlighting a general trend for all periods of Biophony following wildfire disturbance. Trends during these periods display a relative increase in Biophony, particularly for moderate and high severity areas for the Glass and Walbridge fires and all severities for the Kincade. This increase in post-fire predicted Biophony is most notable for dusk and night periods and, to a lesser degree, for dawn. The midday patterns show an inverse pattern, with lower predicted post-fire Biophony than other periods.

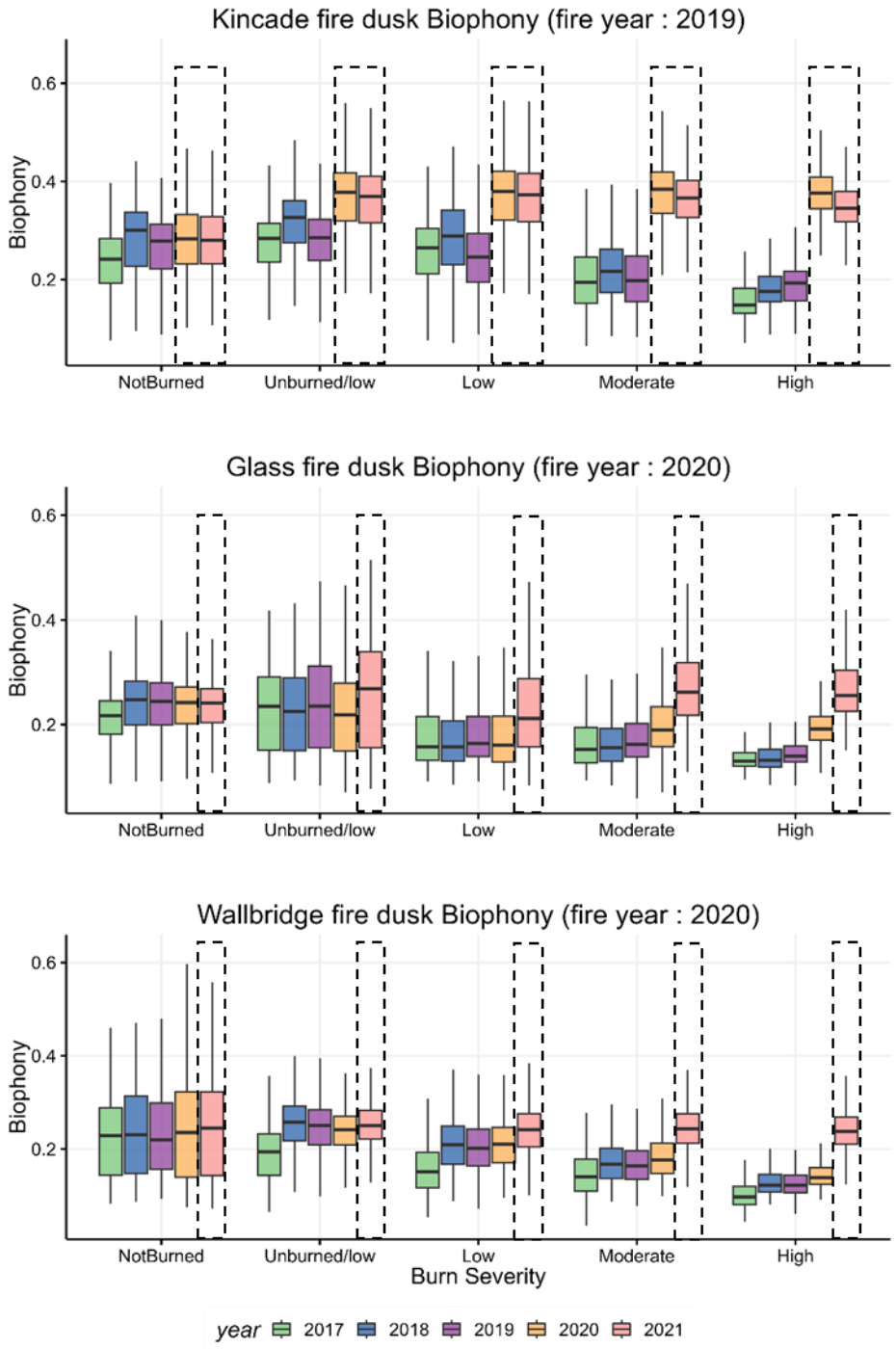


Figure 4-8. Wildfire predicted annual Biophony for the dusk period within the burned perimeter of three historic wildfires stratified by MTBS burn severity and a non-burned, pseudo-control sample 1-2 km away from the fire perimeter with recent-reburn areas removed from analysis. Post-burn years are outlined in black dashed boxes.

4.5. DISCUSSION

The need for cost-effective biomonitoring continues at increasing rates, with widespread recognition that quantifying changes in animal communities occur from changing landscapes associated with changing land use and climate (Schmitz et al., 2023). Our analyses highlight the ability of ecoacoustic metrics to capture spatiotemporal differences across habitats and periods of day related to human impact and wildfire disturbance reflected by changes in remotely sensed vegetation. We linked ecoacoustic metrics to varying degrees with environmental characteristics, demonstrating the helpfulness in determining acoustic-informed biodiversity across heterogeneous landscapes using citizen-science-driven data collection. Within these analyses, we used a geographically-informed method for error assessment (geo-CV) and solar position to identify acoustic periods of interest. We also compared the patterns of three commonly used acoustic indices with soundscape components and bird species richness to determine differences in spatiotemporal patterns and identify regions with higher model confidence. These analyses support the quantification of spatial and temporal ecoacoustics biomonitoring techniques that can facilitate the understanding of vocalizing species dynamics by highlighting core and edge communities.

4.5.1. Spatial patterns

Relationships among environmental variables and acoustic indices have been numerous over the last decade (e.g., Benocci et al., 2020; Hao et al., 2021; Holgate et al., 2021; Turner et al., 2018), with a select few studies modeling both spatial and temporal patterns (Holgate et al., 2021; Mullet et al., 2016). More recently, work has established the utility of ecoacoustic metrics in an SDM framework (Desjonquères et al., 2022), which we believe is a novel direction for

ecoacoustic work for aiding in conservation and biomonitoring. However, the success of acoustic SDMs relies on the ability to use reliable ecoacoustic metrics that are derived in a semi- or fully-automated approach (Desjonquères et al., 2022) which we supply here.

Our analyses suggest that spatial acoustic patterns are strongly driven by climate variables, which mirrors patterns previously documented in bird SDMs (Burns et al., 2020). Similarly, SDM studies incorporating forest structure metrics find them vital descriptors of bird species habitat selection (Goetz et al., 2010; Vierling et al., 2008) and national scale richness patterns (Goetz et al., 2014). However, we find here that structural metrics are the least significant predictor group for bird species richness but, when present, have a positive effect. When considering the top-five most influential predictors, increased structural complexity (e.g., Figure 4-4) corresponded with decreased Anthropophony, Biophony, and ACI and increased values of Quiet and H. The relationship between structure and ecoacoustic metrics supports prior ecoacoustic work finding structure was linked to ecoacoustic patterns (Boelman et al., 2007; Darras et al., 2016; Dröge et al., 2021) and are expanded by the inclusion of more complex structural metrics that convey vertical and horizontal heterogeneity characteristics (Darras et al., 2016). For example, Darras et al. (2016) found that simple measures of forest structure predicted sound transmission poorly among different habitat types. This finding is partly corroborated by Haupert et al. (2022), who found habitat differences accounted for a smaller proportion of sound attenuation than geometric attenuation, which accounted for up to 80% of attenuation below 10 kHz. The inclusion and importance of more complex metrics that quantify variation in structure may add nuance to these findings that, instead of actual habitat differences, indicate edges or transition zones between habitat types control soundscape patterns of core habitats.

For example, Anthropophony patterns corresponded with the presence of and transition to forested areas from other, less structurally complex LULCs represented by variation in the mid-canopy (e.g., RH $25\%_{990m\ sd}$) and were not significantly affected beyond these edges. Additionally, Quiet was strongly positively modulated by the structural complexity of the bottom of the forest canopy (Figure 4-4 – VDR $Bottom_{450m\ mean}$), which was linked to both less developed landscapes but also landscapes with lower sound transmission distance, particularly for Biophony which had a higher frequency than Anthropophony (Haupert et al., 2022). On average, canopy height was the most important predictor for Biophony (RH $95\%_{450m\ mean} > 20m$), had the highest effect on Biophony for the dawn period, and lower Biophony values corresponded with conifer forests. Conifer forests in Sonoma have lower structural complexity in the middle and upper canopy but higher complexity in the lower third canopy. This spatial pattern of low biodiversity is seen in the bird species richness data (though structural information is not as important) and follows previous bird species mapping efforts in Sonoma (Burns et al., 2020). If low structural complexity was the primary driver of patterns increasing sound transmission, we would anticipate these regions (conifers and urban settings) would have the highest levels of Biophony. Instead, we see the highest levels of acoustic-informed biodiversity along developed-natural landscape gradients.

Among other spatial relationships, we found bird species richness and Biophony were negatively associated with the distance from structures and roads, which follows previous work (Buxton et al., 2019; Mullet et al., 2016), although this relationship was less important for both ecoacoustic responses when considered in this analysis. Climate data, known to strongly affect animal vocalizing behavior (Krause and Farina, 2016) and drive weather patterns that affect

sound levels (Mennitt et al., 2014), has strengthened the interpretation of soundscape spatial patterns as an essential set of predictors. Our results also support previous sound-level mapping efforts that demonstrated the importance of including nighttime lights in models (Mennitt and Fristrup, 2016), particularly, in our case, for Anthropophony and bird species richness. The nighttime lights data highlight Sonoma's highly urbanized commercial spaces that strongly modulate Anthropophony, where a quick rise in anthropogenic noise co-occurs, while any light above residential levels decreases bird species richness. Incorporating climate and human impact into soundscape modeling is vital to understanding these critical global drivers altering species distribution (Pimm et al., 2014).

In addition to analyzing the most important predictors, the least important can help illuminate landscape context patterns. For example, agricultural proximity was not crucial for predicting Anthropophony, and building density was more important than the distance to the nearest building, while the distance to streets was even less important for predicting Anthropophony. This pattern persists for all responses regarding agricultural information when included in models (i.e., not Quiet), possibly implying that agrarian land use is either not important in modeling soundscapes or is reflected in other, non-thematic layers. This finding suggests that other studies may avoid using thematic layers, which can differ across study regions in quality or availably relative to remote sensing predictors. Conversely, many of the most important predictors in these analyses are available for large areas worldwide. While we used climate datasets specifically generated for California, USA, we also used predictors found in global datasets like the monthly climate variables from WorldClim (www.worldclim.org). VIIRS nighttime lights and high-resolution DEMs (e.g., 1-3 m) are other example data sets available for

large areas worldwide. Forest structure can be derived using our continuous GEDI modeling process, but coarser, 1 km, continuous products exist for canopy height and other structure metrics (Dubayah et al., 2021c). The availability of such datasets that reflect the most important variables related to soundscapes here could facilitate extending these analyses to other regions.

To our knowledge, our generation of error maps (Figure 4-5 – Biophony IQR) is the first effort to quantify the spatial and temporal error in soundscape mapping patterns. This information can be helpful in biomonitoring approaches as a likelihood probability can be derived whether an observed value in Biophony, or other ecoacoustic metrics, deviates from a central tendency in the data (Sethi et al., 2020). Maps can also aid in targeting areas of interest (e.g., highly variable regions adjacent to agrarian space in the southern part of Sonoma at night).

4.5.2. Temporal patterns

The temporal patterns in ecoacoustic metrics are linked to the changing behavior of animal and human communities across the diurnal cycle (Fairbrass et al., 2019). We provide predictive maps that highlight the changes in soundscape structure with time of day (hot moments) in different spaces (hotspots) (Holgate et al., 2021). For example, regardless of time of day we see that Biophony remains high in urban-adjacent and eastern mixed oak/hardwood forests while similar hotspots and hot moments of bird species richness co-occur with Biophony in the center of the county (Figure 4-5). Anthropophony and Quiet follow anticipated patterns and have opposing trends throughout the day though Quiet is affected by some high levels of Biophony at night in the northeast. Our bird species richness models also captured strong diurnal effects with low levels at night and higher levels during dawn and midday periods. Acoustic indices were not strongly linked to bird richness or Biophony for any period or broadly across the county, but

local areas have higher correlation at specific periods. For example, NDSI captures an increase in nighttime Biophony in the northeastern portion of the county. Interestingly, Geophony, ACI, and H have little variation across periods when compared to the range of values for Anthropophony, Biophony, Quiet, and bird species richness. The only exception is that ACI has observably lower levels at night but the spatial patterns make interpretation non-trivial. From our models, soundscape components and diversity metrics have a greater degree of sensitivity to time of day and across heterogeneous landscapes, while acoustic indices lack reliability in capturing daily acoustic activity cycles in these predictive models.

In addition to temporal patterns, our sampling recording scheme of 1 min every 10 min captured anticipated diel patterns in soundscapes. The effectiveness of sampling recordings counters previous work that posits continuous soundscapes should be used to capture differences between day and night (e.g., Dröge et al., 2021). Our ability to discern patterns could be due to the volume of data (1,195 sites totaling over 700,000 minutes of date) that we extracted mean period values for each location from or the use of solar positioning to determine periods that better summarize diel trends. We initially attempted to model hourly patterns and found this potentially introduced unintended smoothing effects on soundscape patterns. Light intensity at sunrise is one of many factors hypothesized to drive dawn chorus patterns, and using more informed temporal windows can aid in capturing temporal soundscape dynamics (Farina et al., 2015).

Furthermore, our data spanned over four months of the year, during which physiological effects of sunrise and sunset would shift across time of day, and biotic soundscape patterns relate strongly to such diurnal cycles (Farina et al., 2015). Though we did not model seasonal changes

in soundscapes here, modeling approaches like these could highlight seasonal shifts in Biophony based on the high importance of seasonal predictors like phenology and quarterly climate. Patterns in post-wildfire recovery may also be more nuanced when explored across seasons (Gasc et al., 2018).

4.5.3. Acoustic Indices

Among the three acoustic indices included here, NDSI has the most consistent spatial and temporal patterns with Anthropophony and Biophony, two soundscape components the index summarizes. ACI had the lowest average performance for dawn and did not perform well across other periods. ACI and H also had non-trivial spatial patterns at a broad level, poorly capturing anticipated human and biodiversity gradients with slight variation across times of day compared to other ecoacoustic metrics. NDSI values appear reliable across forested and developed areas and higher at the edges of urban areas where the highest Biophony levels occur. However, NDSI requires further validation in riparian areas where Biophony can be high, but instead, NDSI portrayed a strong anthropogenic impact and low biotic activity. Both H and NDSI captured lower bird species richness and Biophony well in urban settings, but the correlations degrade outside these higher biodiversity areas. Comparatively, ACI correlated with lower values of Biophony and especially bird species richness in the county's forested and less developed regions but poorly captured the transition from less-developed LULCs to urban areas. Given these patterns with Biophony and bird species richness, the acoustic indices have variable reliability in capturing patterns in biodiversity, and both environmental setting and time of day are important considerations when applied. ACI and H are more reliable at capturing low levels of biotic activity, especially outside of developed urban settings, and degrade in reliability in

urban settings, reinforcing prior findings that acoustic indices are poorly constrained in urban settings (Fairbrass et al., 2017). We also recommend using NDSI's individual Anthropophonic and Biophonic components, which can capture their respective activity more reliably than the composite NDSI (Quinn et al., in review). Using these narrower frequency band indices also supports recent work suggesting narrowing frequency bands to apply indices correctly (Pérez-Granados and Traba, 2021). Modeling these indices here was not done due to the computational intensity of analyses though an exhaustive investigation of the spatiotemporal reliability of acoustic indices is necessary in diverse landscapes and across seasons and times of day.

Even though spatiotemporal patterns in indices did not follow soundscape component or bird species richness patterns as closely as we anticipated, our findings regarding predictor importance agree with several studies that demonstrate forest structure is associated with acoustic indices (Do Nascimento et al., 2020; Fuller et al., 2015; Ng et al., 2018). We found structure was relatively more important in predicting all indices than their prevalence in the predictor set would suggest (Figure 4-3). For example, Do Nascimento et al. (2020) found ACI was negatively related to canopy cover, and we observed negative relationships between ACI and relative height. Structural predictors also comprised three of the top five predictors for H, all having a roughly positive relationship. However, we found structural predictors not as influential for NDSI. Our results help emphasize that structural settings greatly influence the patterns of some acoustic indices expanding prior work that has identified structural linking across other ecosystems (Boelman et al., 2007; Chen et al., 2021; Do Nascimento et al., 2020). However, further work is needed to verify the reliability of indices when used across heterogeneous landscapes, as patterns shown here are poorly linked to Biophony and bird species richness.

4.5.4. Biophony following wildfire

We found Biophony one to two years post-wildfire increased across burn severities for the Kincade Fire and increased one-year post-wildfire for moderate and high severity areas in the Glass and Walbridge fires. We did not observe differences in bird species richness across burn severities as observed in other arid California fire studies (Tingley et al., 2016), which may imply that in 1-2 years following a fire, insect, amphibian, or mammal reorganization dominates the signal. This increase in Biophony reinforces prior work that found a rise in Biophonic activity following wildfires (Duarte et al., 2021; Gasc et al., 2018), even though recovery timelines differ. Duarte et al. (2021) found soundscape differences in burned and unburned Brazilian savannas were minimal during the day one year after fire, but differences remained at night, a pattern also observed in Arizona Sky Islands post-wildfire (Gasc et al., 2018). We found the same signal where Biophony was higher 1-year post-fire for dawn, dusk, and especially at night, while midday Biophony was lower than in pre-burn and unburned areas. Additionally, bird species richness was lowest at night across Sonoma, concentrated in urban and urban adjacent spaces, and therefore the increased Biophony may be due to increased activity in insect or amphibian activity (based on the low prevalence of mammal calls in the S2L dataset; Quinn et al., 2022). Based on 2021 prediction surfaces (Figure 4-6), the only acoustic index that captured a change in acoustic activity in wildfire-impacted areas was ACI, where values across all periods were higher than surrounding regions (most notable for the Kincade fire).

4.5.5. Limitations

In our opinion, the most significant limitation of this study was the lack of finer-taxonomic Biophony. Functional groupings (e.g., insects, amphibians, birds) would have allowed for a more

informed interpretation of diel patterns and wildfire successional recovery, but would require extensive improvements in data tagging that are currently beyond the scope of this study.

To our knowledge, AudioMoth ARUs have not been thoroughly investigated regarding their effective recording radius, which is a critical component in estimating biodiversity (Shaw et al., 2022; Yip et al., 2017). Because of this and our inability to model sound transmission differences among habitats, we treated acoustic observations as a site feature and used a nominal radius (45 m) to generate prediction surfaces, though, beyond 50 m, differences in ARU models appear to be minimal in signal detection (Shaw et al., 2022). These analyses build from a larger project that utilized citizen science efforts to gather ARU-based observations. This meant that the deployment of ARUs was not done in a manner to allow for the estimation of sound attenuation, especially across the range of land cover and climate conditions during the five-year study. In our models, we attempted to account for variation in sound attenuation by including predictors such as horizontal heterogeneity of forest structure and topographic position. We also focused analyses on frequencies below 11 kHz (10 kHz for NDSI and 11 kHz for ABGQ), meaning that geometric attenuation may drive sound attenuation across habitats (up to 80% of attenuation below 10 kHz) and that habitat differences account for a smaller portion of attenuation than other factors (Haupert et al., 2022). Darras et al. (2018) found that 95% of bird detections were included within 40 m of ARU, and our modeling resolution of 90 m cells roughly approximates a high confidence radius for bird biophony detection range, though under-estimates the effects of lower frequency sound sources. This may be why we see an over-prediction in NDSI, preferentially reflecting Anthropophonies, as they primarily occur below 5 kHz and therefore propagate further than the higher-frequency Biophony sounds.

4.6. CONCLUSION

Our results provide the first contextualization of a very large acoustic dataset and spatial predictors developed to investigate regional-scale patterns of soundscape components and their response to wildfire and burn severity. Utilizing the relative prevalence of predictor groups will be helpful as an among-study tool for comparison of important environmental characteristics. Creating soundscape component maps can aid in identifying areas of attention where changes in animal communities occur at the edge of human impact or natural disturbances like wildfires. These mapping efforts highlight the ability to spatially scale-up discrete recording locations using predictive, remote sensing modeling approaches. Continuous ecoacoustic spatiotemporal mapping can be used as a non-invasive method to study biodiversity and the effects of disturbance across heterogeneous habitats as urban development expands, protected areas are created, and climate change continues.

4.7. APPENDIX 4A

4.7.1. GEDI methodology

We used three GEDI version 2 products for this analysis: GEDI L1B (Dubayah et al. 2021a), GEDI L2A (Dubayah et al. 2021b), and GEDI L2B (Dubayah et al. 2021d). We selected shots between April 17 2019 and September 30 2021 and days of year 105 to 243 to focus on leaf-on conditions. We applied quality and degrade filters to select high quality GEDI L2 shots. We used the L1B dataset in conjunction with airborne lidar (ALS) acquired in 2013 (sonomavegmap.org/) to "co-locate" GEDI waveforms with the ALS point cloud (Hancock et al. 2019). We excluded ALS tiles which overlapped areas that had burned since 2013 since the structural changes may lead to poor matching between GEDI and ALS. Even after removing areas that burned there is still a tremendous amount of ALS data, so we only used a subset of the ALS data for orbital co-location. Specifically for each orbit, we chose the 50 ALS tiles with the most quality GEDI shots. We only selected shots from orbits and applied the suggested co-location X and Y offsets in the correlation value exceeded 0.8.

We uploaded the co-located GEDI shots to Google Earth Engine (GEE) for the continuous structure mapping step. We paired high quality GEDI L2 shots with Landsat Collection 2 surface reflectance time series modeled harmonic coefficients from the continuous change detection and classification (CCDC; Zhu and Woodcock 2014) algorithm in GEE. We ran CCDC from January 1 2002 to June 30 2022 with the following Landsat Collection 2 spectral bands/indices: green, red, near-infrared (NIR), short-wave infrared (SWIR) 1, SWIR 2, enhanced vegetation index (EVI), tasseled-cap wetness (TCW), and normalized burn ratio (NBR). We only selected first order harmonic coefficients (intercept, slope, root-mean-squared-error, cosine, and sine) in order

to simplify the predictive model. We also used a set of ancillary topographic predictor variables, namely slope, tangent of slope, aspect, height above nearest drainage (HAND), compound topographic index (CTI), terrain ruggedness index (TRI), and topographic position index (TPI). Slope, tangent of slope, and aspect were computed from the national elevation dataset 3DEP ⅓ arc-second (10 m) DEM. HAND was derived from the 30 m SRTM DEM (Donchyts et al. 2016). CTI, TRI, and TPI were derived from the 90 m MERIT DEM (Amatulli et al. 2020). We selected shots on low to moderate slopes, removing any shots where the average ground slope was greater than 30 degrees.

We utilized the gee-ccdc-tools repository (Arévalo et al. 2020) to extract modeled time series harmonic coefficients and predict GEDI structure at multiple time periods. To account for GEDI's uneven shot distribution we used Uber's H3 hexagonal grid system (h3geo.org/) to randomly subsample each hexagon (resolution 6; 3.2 km edge length) containing at least 10 quality GEDI shots. We used the 25th percentile of all per grid cell shot counts to determine the subsampling value of 130 shots per grid. Grids were then randomly divided into training (70%) and testing (30%). To account for potential impacts of spatial autocorrelation on model performance we applied an inward buffer of 2,000 m to each testing hexagon grid, ensuring that shots used for model evaluation are spatially independent. We randomly limited the total number of possible shots to use for training (n = 10,000) and testing (n = 2,000).

We predicted the structure metrics listed in Table 4A-1 at June 15 for years 2017 to 2022 using the Random Forest classifier available in GEE. We used 200 trees and 21 variables per split, keeping the random seed constant for each metric.

Table 4A-1. GEDI metrics predicted with Random Forest and corresponding evaluation metrics

GEDI Metric	Units	RMSE	Median absolute error
RH 5	meters above ground	2.2	1.06
RH 25	meters above ground	4.08	2.30
RH 50	meters above ground	5.31	3.06
RH 75	meters above ground	6.31	3.70
RH 95	meters above ground	7.43	4.38
RH 100	meters above ground	7.98	4.64
VDR_B	-	0.17	0.11
VDB_M	-	0.14	0.09
VDR_T	-	0.17	0.11
cover	-	0.22	0.13
PAI	m2/m2	1.31	0.83
FHD	-	0.4	0.21
PAVD 0 to 5	m2/m3	0.09	0.05
PAVD 5 to 10	m2/m3	0.09	0.05

4.7.2. Final random forest, VIF-thinned predictors

Table 4A-2. Final predictors for random forest models after VIF feature reduction organized by predictor. Predictor is 90 m unless otherwise marked. This visualization orients attention to the prevalence of specific predictors in models, not the predictors for each response.

Predictor	Response1	Response2	Response3	Response4	Response5	Response6	Response7	Response8
DHIcumulative	Biophony	---	---	---	---	---	---	---
DHIminimum	Quiet	ACI	---	---	---	---	---	---
DHIvariance	Biophony	Geophony	H	---	---	---	---	---
NDVI05pct	Anthropophony	NDSI	bird richness	---	---	---	---	---
NDVI95pct	Anthropophony	Geophony	ACI	NDSI	bird richness	---	---	---
NDVISeasDiff	Anthropophony	Biophony	Geophony	H	NDSI	bird richness	---	---
Nighttime lights	Anthropophony	Biophony	Quiet	Geophony	ACI	H	NDSI	bird richness
AET Q1	Anthropophony	ACI	NDSI	bird richness	---	---	---	---
AET Q2	ACI	H	NDSI	bird richness	---	---	---	---
AET Q3	Biophony	---	---	---	---	---	---	---

Table 4A-2. Final predictors for random forest models after VIF feature reduction organized by predictor. Predictor is 90 m unless otherwise marked. This visualization orients attention to the prevalence of specific predictors in models, not the predictors for each response.

Predictor	Response1	Response2	Response3	Response4	Response5	Response6	Response7	Response8
AET Q4	Biophony	ACI	NDSI	---	---	---	---	---
CWD Q1	Anthropophony	H	bird richness	---	---	---	---	---
CWD Q2	Geophony	---	---	---	---	---	---	---
CWD Q3	Anthropophony	H	NDSI	---	---	---	---	---
CWD Q4	H	bird richness	---	---	---	---	---	---
PET Q1	Geophony	ACI	H	---	---	---	---	---
PET Q2	Anthropophony	Biophony	---	---	---	---	---	---
PET Q3	Biophony	Quiet	---	---	---	---	---	---
PET Q4	Geophony	---	---	---	---	---	---	---
PPT Q1	Anthropophony	Biophony	Quiet	H	NDSI	bird richness	---	---
PPT Q2	Biophony	Quiet	ACI	H	bird richness	---	---	---
PPT Q3	NDSI	---	---	---	---	---	---	---
PPT Q4	Geophony	ACI	bird richness	---	---	---	---	---
TMN Q1	Biophony	Quiet	---	---	---	---	---	---
TMN Q2	Biophony	Quiet	ACI	bird richness	---	---	---	---
TMN Q3	Anthropophony	NDSI	---	---	---	---	---	---
TMN Q4	H	bird richness	---	---	---	---	---	---
TMX Q1	Anthropophony	ACI	bird richness	---	---	---	---	---
TMX Q2	Anthropophony	Geophony	---	---	---	---	---	---
TMX Q3	Anthropophony	Biophony	---	---	---	---	---	---
TMX Q4	Quiet	Geophony	H	NDSI	---	---	---	---
building density 450m	Biophony	Geophony	ACI	bird richness	---	---	---	---
building density 990m	Anthropophony	H	NDSI	---	---	---	---	---
Cover 90m	H	---	---	---	---	---	---	---
Cover 990m sd	Biophony	bird richness	---	---	---	---	---	---
Elevation	Quiet	Geophony	ACI	NDSI	---	---	---	---
Distance to coast	Biophony	Geophony	ACI	H	NDSI	bird richness	---	---
Distance to streams	Anthropophony	Biophony	Geophony	ACI	H	NDSI	bird richness	---
Distance to buildings	Anthropophony	Biophony	Quiet	Geophony	H	NDSI	bird richness	---
Distance to streets	Anthropophony	Biophony	Quiet	Geophony	H	NDSI	bird richness	---
Fhd pai 450m sd	bird richness	---	---	---	---	---	---	---
Fhd pai 90m	ACI	---	---	---	---	---	---	---
Fhd pai 990m sd	H	---	---	---	---	---	---	---
Pavd 0-5m 450m mean	Biophony	bird richness	---	---	---	---	---	---
Pavd 0-5m 450m sd	Geophony	---	---	---	---	---	---	---
Pavd 0-5m 90m	NDSI	---	---	---	---	---	---	---
Pavd 0-5m 990m mean	Geophony	H	---	---	---	---	---	---
Pavd 0-5m 990m sd	Anthropophony	---	---	---	---	---	---	---
Pavd 5-10m 990m mean	Anthropophony	---	---	---	---	---	---	---
Pavd 5-10m 990m sd	Geophony	ACI	---	---	---	---	---	---
Pct ag 450m mean	Anthropophony	ACI	bird richness	---	---	---	---	---
Pct ag 450m sd	Biophony	Geophony	H	NDSI	---	---	---	---
Pct ag 90m	Anthropophony	Biophony	Geophony	ACI	H	NDSI	---	---

Table 4A-2. Final predictors for random forest models after VIF feature reduction organized by predictor. Predictor is 90 m unless otherwise marked. This visualization orients attention to the prevalence of specific predictors in models, not the predictors for each response.

Predictor	Response1	Response2	Response3	Response4	Response5	Response6	Response7	Response8
Pct ag 990m mean	Geophony	H	NDSI	bird richness	---	---	---	---
Pct ag 990m sd	Anthropophony	Biophony	---	---	---	---	---	---
RH100 450m sd	Anthropophony	---	---	---	---	---	---	---
RH100 90m	Geophony	---	---	---	---	---	---	---
RH25 450m sd	Geophony	---	---	---	---	---	---	---
RH25 990m mean	NDSI	---	---	---	---	---	---	---
RH25 990m sd	Anthropophony	---	---	---	---	---	---	---
RH50 450m mean	bird richness	---	---	---	---	---	---	---
RH50 990m mean	H	---	---	---	---	---	---	---
RH5 990m sd	NDSI	---	---	---	---	---	---	---
RH75 990m mean	ACI	---	---	---	---	---	---	---
RH95 450m mean	Biophony	---	---	---	---	---	---	---
Slope	Anthropophony	Biophony	Quiet	Geophony	ACI	H	NDSI	bird richness
TPI	Anthropophony	Biophony	Geophony	H	NDSI	bird richness	---	---
TPI classified	Quiet	ACI	---	---	---	---	---	---
VDR bot. 450m mean	Quiet	---	---	---	---	---	---	---
VDR mid. 450m sd	Biophony	---	---	---	---	---	---	---
VDR top 450m sd	H	---	---	---	---	---	---	---
VDR top 990m sd	Quiet	---	---	---	---	---	---	---

4.8. REFERENCES

Amatulli, G., McInerney, D., Sethi, T., Strobl, P., Domisch, S., 2020. Geomorpho90m, empirical evaluation and accuracy assessment of global high-resolution geomorphometric layers. Sci. Data, 7(1), 1-18. doi.org/10.1038/s41597-020-0479-6

Arévalo, P., Bullock, E. L., Woodcock, C. E., Olofsson, P., 2020. A suite of tools for continuous land change monitoring in google earth engine. Front. in Clim., 2, 576740. doi.org/10.3389/fclim.2020.576740

Barbaro, L., Sourdril, A., Froidevaux, J.S.P., Cauchoix, M., Calatayud, F., Deconchat, M., Gasc, A., 2022. Linking acoustic diversity to compositional and configurational heterogeneity in mosaic landscapes. Landsc. Ecol. 37, 1125–1143. doi.org/10.1007/s10980-021-01391-8

Barber, J.R., Burdett, C.L., Reed, S.E., Warner, K.A., Formichella, C., Crooks, K.R., Theobald, D.M., Fristrup, K.M., 2011. Anthropogenic noise exposure in protected natural areas: Estimating the scale of ecological consequences. Landsc. Ecol. 26, 1281–1295. doi.org/10.1007/s10980-011-9646-7

Benito, B.M., 2021. R package spatialRF: Easy Spatial Regression with Random Forest. doi.org/10.5281/zenodo.4745208

Benocci, R., Brambilla, G., Bisceglie, A., Zambon, G., 2020. Eco-acoustic indices to evaluate soundscape degradation due to human intrusion. Sustain. 12, 1–19. doi.org/10.3390/su122410455

Benocci, R., Potenza, A., Bisceglie, A., Roman, H.E., Zambon, G., 2022. Mapping of the Acoustic Environment at an Urban Park in the City Area of Milan, Italy, Using Very Low-Cost Sensors. Sensors 22, 1–23. doi.org/10.3390/s22093528

Boelman, N.T., Asner, G.P., Hart, P.J., Martin, R.E., 2007. Multi-trophic invasion resistance in Hawaii: Bioacoustics, field surveys, and airborne remote sensing. Ecol. Appl. 17, 2137–2144. doi.org/10.1890/07-0004.1

Bormpoudakis, D., Sueur, J., Pantis, J.D., 2013. Spatial heterogeneity of ambient sound at the habitat type level: Ecological implications and applications. Landsc. Ecol. 28, 495–506. doi.org/10.1007/s10980-013-9849-1

Bradfer-Lawrence, T., Bunnefeld, N., Gardner, N., Willis, S.G., Dent, D.H., 2020. Rapid assessment of avian species richness and abundance using acoustic indices. Ecol. Indic. 115, 106400. doi.org/10.1016/j.ecolind.2020.106400

Breiman, L., 2001. Random forests. Mach. Learn. 5, 5–32. doi.org/10.1007/9781441993267_5

Burns, P., Clark, M., Salas, L., Hancock, S., Leland, D., Jantz, P., Dubayah, R., Goetz, S.J., 2020. Incorporating canopy structure from simulated GEDI lidar into bird species distribution models. Environ. Res. Lett. 15. doi.org/10.1088/1748-9326/ab80ee

Buxton, R.T., McKenna, M.F., Clapp, M., Meyer, E., Stabenau, E., Angeloni, L.M., Crooks, K., Wittemyer, G., 2018. Efficacy of extracting indices from large-scale acoustic recordings to monitor biodiversity. Conserv. Biol. 32, 1174–1184. doi.org/10.1111/cobi.13119

Buxton, R.T., McKenna, M.F., Mennitt, D., Brown, E., Fristrup, K., Crooks, K.R., Angeloni, L.M., Wittemyer, G., 2019. Anthropogenic noise in US national parks – sources and spatial extent. Front. Ecol. Environ. 559–564. doi.org/10.1002/fee.2112

Carruthers-Jones, J., Eldridge, A., Guyot, P., Hassall, C., Holmes, G., 2019. The call of the wild: Investigating the potential for ecoacoustic methods in mapping wilderness areas. Sci. Total Environ. 695, 133797. doi.org/10.1016/j.scitotenv.2019.133797

Celis-Murillo, A., Deppe, J.L., Ward, M.P., 2012. Effectiveness and utility of acoustic recordings for surveying tropical birds. J. F. Ornithol. 83, 166–179. doi.org/10.1111/j.1557-9263.2012.00366.x

Chen, Y.F., Luo, Y., Mammides, C., Cao, K.F., Zhu, S., Goodale, E., 2021. The relationship between acoustic indices, elevation, and vegetation, in a forest plot network of southern China. Ecol. Indic. 129, 107942. doi.org/10.1016/j.ecolind.2021.107942

Clark, M.L., Salas, L., Baligar, S., Quinn, C.A., Snyder, R.L., Leland, D., Schackwitz, W., Goetz, S.J., Newsam, S., 2023. Ecological Informatics The effect of soundscape composition on bird vocalization classification in a citizen science biodiversity monitoring project. Ecol. Inform. 75, 102065. doi.org/10.1016/j.ecoinf.2023.102065

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 Ecol. Evol. 9, 1928–1938. doi.org/10.1111/2041-210X.13031

Darras, K., Pütz, P., Fahrurrozi, Rembold, K., Tscharntke, T., 2016. Measuring sound detection spaces for acoustic animal sampling and monitoring. Biol. Conserv. 201, 29–37. doi.org/10.1016/j.biocon.2016.06.021

de Kluijver, H., Stoter, J., 2003. Noise mapping and GIS: Optimising quality and efficiency of noise effect studies. Comput. Environ. Urban Syst. 27, 85–102. doi.org/10.1016/S0198-9715(01)00038-2

Desjonquères, C., Villén-Pérez, S., Paulo, D.M., Márquez, R., Beltrán, J.F., Llusia, D., 2022. Acoustic species distribution models (aSDMs): A framework to forecast shifts in calling

behaviour under climate change. Methods Ecol. Evol. 13, 1–14. doi.org/10.1111/2041-210X.13923

Do Nascimento, L.A., Campos-Cerqueira, M., Beard, K.H., 2020. Acoustic metrics predict habitat type and vegetation structure in the Amazon. Ecol. Indic. 117, 106679. doi.org/10.1016/j.ecolind.2020.106679

Donchyts, Gennadii, Hessel Winsemius, Jaap Schellekens, Tyler Erickson, Hongkai Gao, Hubert Savenije, and Nick van de Giesen, 2016. Global 30m Height Above the Nearest Drainage (HAND), Geophysical Research Abstracts, Vol. 18, EGU2016-17445-3, 2016, EGU General Assembly.

Dröge, S., Martin, D.A., Andriafanomezantsoa, R., Burivalova, Z., Fulgence, T.R., Osen, K., Rakotomalala, E., Schwab, D., Wurz, A., Richter, T., Kreft, H., 2021. Listening to a changing landscape: Acoustic indices reflect bird species richness and plot-scale vegetation structure across different land-use types in north-eastern Madagascar. Ecol. Indic. 120. doi.org/10.1016/j.ecolind.2020.106929

Duarte, M.H.L., Sousa-Lima, R.S.S., Young, R.J., Vasconcelos, M.F., Bittencourt, E., Scarpelli, M.D.A., Farina, A., Pieretti, N., 2021. Changes on soundscapes reveal impacts of wildfires in the fauna of a Brazilian savanna. Sci. Total Environ. 769, 144988. doi.org/10.1016/j.scitotenv.2021.144988

Dubayah, R., Blair, J.B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., Hofton, M., Hurtt, G., Kellner, J., Luthcke, S., Armston, J., Tang, H., Duncanson, L., Hancock, S., Jantz, P., Marselis, S., Patterson, P.L., Qi, W., Silva, C., 2020. The Global Ecosystem Dynamics

Investigation: High-resolution laser ranging of the Earth's forests and topography. Sci. Remote Sens. 1, 100002. doi.org/10.1016/j.srs.2020.100002

Dubayah, R., Luthcke, S., Blair, J., Hofton, M., Armston, J., Tang, H., 2021a. GEDI L1B Geolocated Waveform Data Global Footprint Level V002 [Data set]. NASA EOSDIS Land Processes DAAC. Accessed 2022-07-11 from doi.org/10.5067/GEDI/GEDI01_B.002

Dubayah, R., Hofton, M., Blair, J., Armston, J., Tang, H., Luthcke, S., 2021b. GEDI L2A Elevation and Height Metrics Data Global Footprint Level V002 [Data set]. NASA EOSDIS Land Processes DAAC. Accessed 2022-07-11 from doi.org/10.5067/GEDI/GEDI02_A.002

Dubayah, R.O., Luthcke, S.B., Sabaka, T.J., Nicholas, J.B., Preaux, S., Hofton, M.A., 2021c. GEDI L3 Gridded Land Surface Metrics, Version 1. doi.org/doi.org/10.3334/ORNLDAAC/1865

Dubayah, R., Tang, H., Armston, J., Luthcke, S., Hofton, M., Blair, J., 2021d. GEDI L2B Canopy Cover and Vertical Profile Metrics Data Global Footprint Level V002 [Data set]. NASA EOSDIS Land Processes DAAC. Accessed 2022-07-11 from doi.org/10.5067/GEDI/GEDI02_B.002

Eidenshink, J., Schwind, B., Brewer, K., Zhu, Z.-L., Quayle, B., Howard, S., 2007. A project for monitoring trends in burn severity. Fire Ecol. 3, 3–21.

Fairbrass, A.J., Firman, M., Williams, C., Brostow, G.J., Titheridge, H., Jones, K.E., 2019. CityNet—Deep learning tools for urban ecoacoustic assessment. Methods Ecol. Evol. 10, 186–197. doi.org/10.1111/2041-210X.13114

Fairbrass, A.J., Rennett, P., Williams, C., Titheridge, H., Jones, K.E., 2017. Biases of acoustic indices measuring biodiversity in urban areas. Ecol. Indic. 83, 169–177. doi.org/10.1016/j.ecolind.2017.07.064

Farina, A., Ceraulo, M., Bobryk, C., Pieretti, N., Quinci, E., Lattanzi, E., 2015. Spatial and temporal variation of bird dawn chorus and successive acoustic morning activity in a Mediterranean landscape. Bioacoustics. doi.org/10.1080/09524622.2015.1070282

Farina, A., Lattanzi, E., Malavasi, R., Pieretti, N., Piccioli, L., 2011. Avian soundscapes and cognitive landscapes: Theory, application and ecological perspectives. Landsc. Ecol. 26, 1257–1267. doi.org/10.1007/s10980-011-9617-z

Flint, L.E., Flint, A.L., Thorne, J.H., Boynton, R., 2013. Fine-scale hydrologic modeling for regional landscape applications: The California Basin Characterization Model development and performance. Ecol. Process. 2, 1–21. doi.org/10.1186/2192-1709-2-25

Frantz, D., 2019. FORCE-Landsat + Sentinel-2 analysis ready data and beyond. Remote Sens. 11. doi.org/10.3390/rs11091124

Fuller, S., Axel, A.C., Tucker, D., Gage, S.H., 2015. Connecting soundscape to landscape: Which acoustic index best describes landscape configuration? Ecol. Indic. 58, 207–215. doi.org/10.1016/j.ecolind.2015.05.057

Furnas, B.J., Callas, R.L., 2015. Using automated recorders and occupancy models to monitor common forest birds across a large geographic region. J. Wildl. Manage. 79, 325–337. doi.org/10.1002/jwmg.821

Gasc, A., Gottesman, B.L., Francomano, D., Jung, J., Durham, M., Mateljak, J., Pijanowski, B.C., 2018. Soundscapes reveal disturbance impacts: biophonic response to wildfire in the

Sonoran Desert Sky Islands. Landsc. Ecol. 33, 1399–1415. doi.org/10.1007/s10980-018-0675-3

Goetz, S.J., Steinberg, D., Betts, M.G., Holmes, R.T., Doran, P.J., Dubayah, R., Hofton, M., 2010. Lidar remote sensing variables predict habitat use by a breeding passerine bird. Ecology 91, 1173–1181.

Goetz, S.J., Sun, M., Zolkos, S., Hansen, A., Dubayah, R., 2014. The relative importance of climate and vegetation properties on patterns of North American breeding bird species richness. Environ. Res. Lett. 9. doi.org/10.1088/1748-9326/9/3/034013

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sens. Environ. 202, 18–27. doi.org/10.1016/j.rse.2017.06.031

Greenwell, B.M., 2017. pdp: An R Package for Constructing Partial Dependence Plots. R J. 9, 421–436.

Gregorutti, B., Michel, B., Saint-Pierre, P., 2017. Correlation and variable importance in random forests. Stat. Comput. 27, 659–678. doi.org/10.1007/s11222-016-9646-1

Hancock, S., Armston, J., Hofton, M., Sun, X., Tang, H., Duncanson, L. I., Kellner, R.J., Dubayah, R., 2019. The GEDI simulator: A large-footprint waveform lidar simulator for calibration and validation of spaceborne missions. Earth and Space Science, 6(2), 294-310. doi.org/10.1029/2018EA000506

Hao, T., Elith, J., Guillera-Arroita, G., Lahoz-Monfort, J.J., 2019. A review of evidence about use and performance of species distribution modelling ensembles like BIOMOD. Divers. Distrib. 25, 839–852. doi.org/10.1111/ddi.12892

Hao, Z., Wang, C., Sun, Z., van den Bosch, C.K., Zhao, D., Sun, B., Xu, X., Bian, Q., Bai, Z., Wei, K., Zhao, Y., Pei, N., 2021. Soundscape mapping for spatial-temporal estimate on bird activities in urban forests. Urban For. Urban Green. 57, 126822. doi.org/10.1016/j.ufug.2020.126822

Haupert, S., Sèbe, F., Sueur, J., 2022. Physics-based model to predict the acoustic detection distance of terrestrial autonomous recording units over the diel cycle and across seasons: Insights from an Alpine and a Neotropical forest. Methods Ecol. Evol. 2023, 614–630. doi.org/10.1111/2041-210X.14020

Hijmans, R., 2022. terra: Spatial Data Analysis.

Hill, A.P., Prince, P., Piña Covarrubias, E., Doncaster, C.P., Snaddon, J.L., Rogers, A., 2018. AudioMoth: Evaluation of a smart open acoustic device for monitoring biodiversity and the environment. Methods Ecol. Evol. 9, 1199–1211. doi.org/10.1111/2041-210X.12955

Hobi, M.L., Dubinin, M., Graham, C.H., Coops, N.C., Clayton, M.K., Pidgeon, A.M., Radeloff, V.C., 2017. A comparison of Dynamic Habitat Indices derived from different MODIS products as predictors of avian species richness. Remote Sens. Environ. 195, 142–152. doi.org/10.1016/j.rse.2017.04.018

Holgate, B., Maggini, R., Fuller, S., 2021. Mapping ecoacoustic hot spots and moments of biodiversity to inform conservation and urban planning. Ecol. Indic. 126. doi.org/10.1016/j.ecolind.2021.107627

Joo, W., Gage, S.H., Kasten, E.P., 2011. Analysis and interpretation of variability in soundscapes along an urban-rural gradient. Landsc. Urban Plan. 103, 259–276. doi.org/10.1016/j.landurbplan.2011.08.001

Kasten, E.P., Gage, S.H., Fox, J., Joo, W., 2012. The remote environmental assessment laboratory's acoustic library: An archive for studying soundscape ecology. Ecol. Inform. 12, 50–67. doi.org/10.1016/j.ecoinf.2012.08.001

Keeley, J.E., Syphard, A.D., 2021. Large California wildfires: 2020 fires in historical context. Fire Ecol. 17. doi.org/10.1186/s42408-021-00110-7

Khait, I., Lewin-Epstein, O., Sharon, R., Saban, K., Perelman, R., Boonman, A., Yovel, Y., Hadany, L., 2019. Plants emit informative airborne sounds under stress. bioRxiv 507590.

Krause, B., Farina, A., 2016. Using ecoacoustic methods to survey the impacts of climate change on biodiversity. Biol. Conserv. 195, 245–254. doi.org/10.1016/j.biocon.2016.01.013

Krause, B., Gage, S.H., Joo, W., 2011. Measuring and interpreting the temporal variability in the soundscape at four places in Sequoia National Park. Landsc. Ecol. 26, 1247–1256. doi.org/10.1007/s10980-011-9639-6

Kuhn, M., 2022. caret: Classification and Regression Training.

Lomolino, M. V., Pijanowski, B.C., Gasc, A., 2015. The silence of biogeography. J. Biogeogr. 42, 1187–1196. doi.org/10.1111/jbi.12525

Lorena, A.C., Jacintho, L.F.O., Siqueira, M.F., Giovanni, R. De, Lohmann, L.G., De Carvalho, A.C.P.L.F., Yamamoto, M., 2011. Comparing machine learning classifiers in potential distribution modelling. Expert Syst. Appl. 38, 5268–5275. doi.org/10.1016/j.eswa.2010.10.031

Machado, R.B., Aguiar, L., Jones, G., 2017. Do acoustic indices reflect the characteristics of bird communities in the savannas of Central Brazil ? Landsc. Urban Plan. 162, 36–43. doi.org/10.1016/j.landurbplan.2017.01.014

Matsinos, Y.G., Mazaris, A.D., Papadimitriou, K.D., Mniestris, A., Hatzigiannidis, G., Maioglou, D., Pantis, J.D., 2008. Spatio-temporal variability in human and natural sounds in a rural landscape. Landsc. Ecol. 23, 945–959. doi.org/10.1007/s10980-008-9250-7

Matsinos, Y.G., Tsaligopoulos, A., 2018. Hot spots of ecoacoustics in Greece and the issue of background noise. J. Ecoacoustics 2, 1–1. doi.org/10.22261/jea.u3xbiy

Maxwell, A.E., Warner, T.A., Fang, F., 2018. Implementation of machine-learning classification in remote sensing: An applied review. Int. J. Remote Sens. 39, 2784–2817. doi.org/10.1080/01431161.2018.1433343

Mennitt, D., Sherrill, K., Fristrup, K., 2014. A geospatial model of ambient sound pressure levels in the contiguous United States. J. Acoust. Soc. Am. 135, 2746–2764. doi.org/10.1121/1.4870481

Mennitt, D.J., Fristrup, K.M., 2016. Influence factors and spatiotemporal patterns of environmental sound levels in the contiguous United States In fl uential factors and spatiotemporal patterns of environmental sound levels in the contiguous United States. Noise Control Eng. 64, 342–353. doi.org/10.3397/1/376384

Mills, S., Weiss, S., Liang, C., 2013. VIIRS day/night band (DNB) stray light characterization and correction. Proc. SPIE 8866, 549–566. doi.org/10.1117/12.2023107

Müller, S., Mitesser, O., Oschwald, L., Scherer-Lorenzen, M., Potvin, C., 2022. Temporal Soundscape Patterns in a Panamanian Tree Diversity Experiment: Polycultures Show an Increase in High Frequency Cover. Front. Ecol. Evol. 10. doi.org/10.3389/fevo.2022.808589

Mullet, T.C., Gage, S.H., Morton, J.M., Huettmann, F., 2016. Temporal and spatial variation of a winter soundscape in south-central Alaska. Landsc. Ecol. 31, 1117–1137. doi.org/10.1007/s10980-015-0323-0

Ng, M. Le, Butler, N., Woods, N., 2018. Soundscapes as a surrogate measure of vegetation condition for biodiversity values: A pilot study. Ecol. Indic. 93, 1070–1080. doi.org/10.1016/j.ecolind.2018.06.003

Pekin, B.K., Jung, J., Villanueva-Rivera, L.J., Pijanowski, B.C., Ahumada, J.A., 2012. Modeling acoustic diversity using soundscape recordings and LIDAR-derived metrics of vertical forest structure in a neotropical rainforest. Landsc. Ecol. 27, 1513–1522. doi.org/10.1007/s10980-012-9806-4

Pieretti, N., Farina, A., Morri, D., 2011. A new methodology to infer the singing activity of an avian community: The Acoustic Complexity Index (ACI). Ecol. Indic. 11, 868–873. doi.org/10.1016/j.ecolind.2010.11.005

Pimm, S.L., Jenkins, C.N., Abell, R., Brooks, T.M., Gittleman, J.L., Joppa, L.N., Raven, P.H., Roberts, C.M., Sexton, J.O., 2014. The biodiversity of species and their rates of extinction, distribution, and protection. Science. 344, 1246752. doi.org/10.1126/science.1246752

Quinn, C.A., Burns, P., Gill, G., Baligar, S., Snyder, R.L., Salas, L., Goetz, S.J., Clark, M.L., 2022. Soundscape classification with convolutional neural networks reveals temporal and geographic patterns in ecoacoustic data. Ecol. Indic. 138, 108831. doi.org/10.1016/j.ecolind.2022.108831

Quinn, C.A., Burns, P., Hakkenberg, C.R., Salas, L., Pasch, B., Goetz, S.J., Clark, M.L., [in review] Soundscape components inform acoustic index patterns and refine estimates of bird species richness.

R Core Team, 2022. R: A language and environment for statistical computing.

Radeloff, V.C., Dubinin, M., Coops, N.C., Allen, A.M., Brooks, T.M., Clayton, M.K., Costa, G.C., Graham, C.H., Helmers, D.P., Ives, A.R., Kolesov, D., Pidgeon, A.M., Rapacciuolo, G., Razenkova, E., Suttidate, N., Young, B.E., Zhu, L., Hobi, M.L., 2019. The Dynamic Habitat Indices (DHIs) from MODIS and global biodiversity. Remote Sens. Environ. 222, 204–214. doi.org/10.1016/j.rse.2018.12.009

Rappaport, D.I., Royle, J.A., Morton, D.C., 2020. Acoustic space occupancy: Combining ecoacoustics and lidar to model biodiversity variation and detection bias across heterogeneous landscapes. Ecol. Indic. 113, 106172. doi.org/10.1016/j.ecolind.2020.106172

Roberts, D.R., Bahn, V., Ciuti, S., Boyce, M.S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J.J., Schröder, B., Thuiller, W., Warton, D.I., Wintle, B.A., Hartig, F., Dormann, C.F., 2017. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography (Cop.). 40, 913–929. doi.org/10.1111/ecog.02881

Rodriguez, A., Gasc, A., Pavoine, S., Grandcolas, P., Gaucher, P., Sueur, J., 2014. Temporal and spatial variability of animal sound within a neotropical forest. Ecol. Inform. 21, 133–143. doi.org/10.1016/j.ecoinf.2013.12.006

Schmitz, O.J., Sylvén, M., Atwood, T.B., Bakker, E.S., Berzaghi, F., Brodie, J.F., Cromsigt, J.P.G.M., Davies, A.B., Leroux, S.J., Schepers, F.J., Smith, F.A., Stark, S., Svenning, J.C.,

Tilker, A., Ylänne, H., 2023. Trophic rewilding can expand natural climate solutions. Nat. Clim. Chang. doi.org/10.1038/s41558-023-01631-6

Sethi, S.S., Jones, N.S., Fulcher, B.D., Picinali, L., Clink, D.J., Klinck, H., Orme, C.D.L., Wrege, P.H., Ewers, R.M., 2020. Characterizing soundscapes across diverse ecosystems using a universal acoustic feature set. Proc. Natl. Acad. Sci. U. S. A. doi.org/10.1073/pnas.2004702117

Shaw, T., Müller, S., Scherer-Lorenzen, M., 2022. Slope does not affect autonomous recorder detection shape: considerations for acoustic monitoring in forested landscapes. Bioacoustics 31, 261–282. doi.org/10.1080/09524622.2021.1925590

Speiser, J.L., Miller, M.E., Tooze, J., Ip, E., 2019. A comparison of random forest variable selection methods for classification prediction modeling. Expert Syst. Appl. 134, 93–101. doi.org/10.1016/j.eswa.2019.05.028

Sueur, J., Farina, A., 2015. Ecoacoustics: the Ecological Investigation and Interpretation of Environmental Sound. Biosemiotics 8, 493–502. doi.org/10.1007/s12304-015-9248-x

Sueur, J., Farina, A., Gasc, A., Pieretti, N., Pavoine, S., 2014. Acoustic indices for biodiversity assessment and landscape investigation. Acta Acust. united with Acust. 100, 772–781. doi.org/10.3813/AAA.918757

Sueur, J., Pavoine, S., Hamerlynck, O., Duvail, S., 2008. Rapid acoustic survey for biodiversity appraisal. PLoS One 3, 1–10. doi.org/10.1371/journal.pone.0004065

Thieurmel, B., Elmarhraoui, A., 2022. suncalc: Compute Sun Position, Sunlight Phases, Moon Position and Lunar Phase.

Tingley, M.W., Ruiz-Gutiérrez, V., Wilkerson, R.L., Howell, C.A., Siegel, R.B., 2016. Pyrodiversity promotes avian diversity over the decade following forest fire. Proc. R. Soc. B Biol. Sci. 283. doi.org/10.1098/rspb.2016.1703

Towsey, M., Wimmer, J., Williamson, I., Roe, P., 2014. The use of acoustic indices to determine avian species richness in audio-recordings of the environment. Ecol. Inform. 21, 110–119. doi.org/10.1016/j.ecoinf.2013.11.007

Tucker, D., Gage, S.H., Williamson, I., Fuller, S., 2014. Linking ecological condition and the soundscape in fragmented Australian forests. Landsc. Ecol. 29, 745–758. doi.org/10.1007/s10980-014-0015-1

Turner, A., Fischer, M., Tzanopoulos, J., 2018. Sound-mapping a coniferous forest—Perspectives for biodiversity monitoring and noise mitigation. PLoS One 13, 1–21. doi.org/10.1371/journal.pone.0189843

Turner, M.G., 1989. Landscape Ecology : The Effect of Pattern on Process. Annu. Rev. Ecol. Syst. 20, 171–197.

Vierling, K.T., Vierling, L.A., Gould, W.A., Martinuzzi, S., Clawges, R.M., 2008. Lidar: Shedding new light on habitat characterization and modeling. Front. Ecol. Environ. 6, 90–98. doi.org/10.1890/070001

Villanueva-Rivera, L.J., Pijanowski, B.C., Doucette, J., Pekin, B., 2011. A primer of acoustic analysis for landscape ecologists. Landsc. Ecol. 26, 1233–1246. doi.org/10.1007/s10980-011-9636-9

Ware, H.E., McClure, C.J.W., Carlisle, J.D., Barber, J.R., Daily, G.C., 2015. A phantom road experiment reveals traffic noise is an invisible source of habitat degradation. Proc. Natl. Acad. Sci. U. S. A. 112, 12105–12109. doi.org/10.1073/pnas.1504710112

Wright, M.N., Ziegler, A., 2017. ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. J. Stat. Softw. 77, 1–17. doi.org/10.18637/jss.v077.i01

Yip, D.A., Bayne, E.M., Sólymos, P., Campbell, J., Proppe, D., 2017. Sound attenuation in forest and roadside environments: Implications for avian point-count surveys. Condor 119, 73–84. doi.org/10.1650/CONDOR-16-93.1

Zhu, Z., Woodcock, C.E., 2014. Continuous change detection and classification of land cover using all available Landsat data. Remote Sens. Environ. 144, 152–171. doi.org/10.1016/j.rse.2014.01.011

Zuur, A.F., Ieno, E.N., Elphick, C.S., 2010. A protocol for data exploration to avoid common statistical problems. Methods Ecol. Evol. 1, 3–14. doi.org/10.1111/j.2041-210x.2009.00001.x

CHAPTER 5: CONCLUSION

The increased need to expand biomonitoring efforts has introduced a range of novel tools and methods. Monitoring biodiversity enables scientists, conservationists, and land managers to understand the current community assemblages across a landscape and then track how these assemblages change both naturally and in relation to disparate disturbances like human development and changing climatic conditions. Ecoacoustics has emerged as a promising new field focused on developing and applying biomonitoring methods that combine remote sensing, landscape ecology, and bioacoustics. The acoustic index is the primary tool in the ecoacoustic toolkit, which is an efficiently derived numerical summary of passively recorded acoustic activity. Acoustic indices have been applied across various habitat types to study different aspects of community biodiversity and have been related to a range of environmental characteristics. However, because approximately 70 indices exist and ecoacoustics as a field is still nascent, work that investigates acoustic index generalizability and their assumptions is lacking. The three chapters in this dissertation address the uncertainty in applying acoustic indices as metrics for biodiversity and apply these metrics in a species distribution modeling framework to understand environmental drivers and spatiotemporal patterns of soundscapes.

Chapter 2 presents a method to classify coarse-scale soundscape sources (Anthropophony, Biophony, Geophony, periods of Quiet, and ARU Interference) to understand the relationship between sounds and acoustic index patterns in Chapter 3. Chapter 2 uses methods from computer vision to learn the patterns of soundscape components, and I provide a code repository to allow for generalization to other studies. I demonstrate that a pre-trained convolutional neural network can be fine-tuned for soundscape classification with a relatively small dataset. The classifier has

a high overall accuracy across soundscape component classes (F0.75-score = 0.88), and I establish that soundscape components reflect anticipated patterns across time of day and land use/land cover (LULC) as well as novel emergent patterns. For example, two prominent studies that quantified Anthropophony and Biophony were set in relatively homogeneous rural or urban settings, while the Sonoma County data used here spans a highly heterogeneous landscape. I could compare rural and urban patterns with these studies while exploring how patterns behave in transitional and edge zones. I also include Quiet, times in which no emergent sounds were recorded, and use this class to quantify naturally-quiet landscapes and times of day, critical dimensions for animal and human wellbeing. Overall, analyses in this chapter demonstrate the ability to use a small acoustic training set of recordings to extract ecologically meaningful soundscape dynamics from a sizeable acoustic dataset.

Chapter 3 focuses on quantifying the dimensions of soundscapes that most strongly drive the values of fifteen commonly applied acoustic indices. This analysis represents one of two studies that attempt to untangle how sonic conditions relate to describe acoustic index values. The other study used approximately 13,000 minutes of data to model acoustic indices using biophinic diversity. They then explored how different soundscape components altered this relationship and established a sensitivity of eleven indices to sound types and recording settings. This was the first holistic analysis to investigate the pattern of multiple indices concerning their underlying soundscapes. I expanded concepts from this analysis by using the classifications in Chapter 2 and a dataset of almost two orders of magnitude larger and five additional acoustic indices to contextualize sonic conditions in relation to acoustic indices and bird species richness. This analysis is two-fold. First, in further investigating the effects of sonic conditions on acoustic

indices and second, contrasting these patterns with another study domain to assess the level of acoustic index generalizability. I found that Biophony alone could capture the same level of variation in modeling bird species richness as fifteen acoustic indices during morning hours. Combining acoustic indices and Biophony enabled 66% of the variance in bird species richness to be modeled, further demonstrating the utility of recent trends utilizing multi-index approaches to represent biodiversity. However, study-specific aural validation is still required to interpret broad patterns in acoustic index reliably, and I believe indices require further analysis before being used as biomonitoring tools due to nontrivial changes in values in different acoustic settings. From this analysis, I provide recommendations to consider sonic conditions for the future application of acoustic indices.

After establishing soundscape components as informative ecoacoustic metrics in Sonoma County and investigating patterns in acoustic indices, I use these metrics to understand the spatiotemporal patterns of soundscapes across Sonoma. Mapping soundscapes has been done since the early 2000s, primarily for urban planning, but biomonitoring-oriented soundscape mapping only emerged in the past decade. These maps tend to focus on a limited relationship with a single dimension of the landscape (e.g., with only human impact, vegetation condition, or forest structure) but rarely integrate multiple datasets that capture a more holistic representation of the environment. Chapter 4 demonstrates the ability to use a large set of environmental variables to predict soundscape metrics. I found evidence that agrees with more targeted studies that 1) animal patterns (Biophony) are strongly related to climate conditions, 2) forest structure is an essential variable in predicting soundscape metrics, particularly Biophony at dawn, and 3) we can use prediction surfaces to understand complex soundscape patterns across heterogeneous

landscapes at different times of the day. I use these maps of Biophony to capture post-wildfire community reorganization related to burn severity. Together, the analyses in this chapter demonstrate our ability to quantify hotspots and hot moments of soundscape activity for biomonitoring efforts.

The analyses in this dissertation focus fully on Sonoma County, California and there remain many opportunities to test the generalizability of methods and findings to other study domains. I believe ecoacoustic biodiversity is most effective for local to regional scale analyses like those presented in Chapter 4, and maybe someday, global analyses, as field data collection requires physically deployed recorders across the landscape. Furthermore, to generate regional-scale mapping products, it is crucial to ensure that endpoints of the system are captured in recordings and as much diversity within the system is measured as there is a large amount of variability in soundscape data. I envision future ecoacoustic analyses utilizing machine and deep learning methods more frequently, particularly in generating annotated acoustic data that can improve the generalizability of acoustic classifiers (Chapter 2). The application of ecoacoustics is rapidly expanding, mainly focusing on work to ensure the proper application of methods so these tools can be helpful for biomonitoring, conservation efforts, and as an aid to urban planners to think of landscape biodiversity holistically and enable more accessible field-based biodiversity estimates.