Received: 12 November 2021 | Revised: 15 September 2022 | Accepted: 15 September 2022

DOI: 10.1002/wsb.1395

TOOLS AND TECHNOLOGY

Automated recognition of ruffed grouse drumming in field recordings

Samuel Lapp[1] | Jeffery L. Larkin[2,3] | Halie A. Parker[2] | Jeffery T. Larkin[4] | Dakotah R. Shaffer[2] | Carolyn Tett[1] | Darin J. McNeil[5] | Cameron J. Fiss[6] | Justin Kitzes[1]

[1]University of Pittsburgh, 103 Clapp Hall, Fifth and Ruskin Avenues, Pittsburgh, PA 15260, USA

[2]Department of Biology, Indiana University of Pennsylvania, 1011 South Drive, Indiana, PA 15701, USA

[3]American Bird Conservancy, The Plains, VA 20198, USA

[4]Department of Environmental Conservation, University of Massachusetts-Amherst, 160 Holdsworth Way, Amherst, MA 01003-9285, USA

[5]Department of Forestry and Natural Resources, University of Kentucky, Lexington, KY 40546, USA

[6]Department of Environmental Biology, State University of New York College of Environmental Science and Forestry, 1 Forestry Dr. Syracuse, NY 13210, USA

Correspondence

Samuel Lapp, University of Pittsburgh, 103 Clapp Hall, Fifth and Ruskin Avenues, Pittsburgh, PA 15260, USA.
Email: sam.lapp@pitt.edu

Funding information

National Science Foundation, Grant/Award Numbers: 1935507, 2120084; Gordon and Betty Moore Foundation; Indiana University of Pennsylvania School of Graduate Studies and Research; National Resource Conservation Service: Conservation Effects Assessment Project, Grant/Award Number:

Abstract

Ruffed grouse (*Bonasa umbellus*) populations are declining throughout their range, which has prompted efforts to understand drivers of the decline. Ruffed grouse monitoring efforts often rely on acoustic drumming surveys, in which a surveyor listens for the distinctive drumming sound that male ruffed grouse produce during the breeding season. Field-based drumming surveys can fail to detect ruffed grouse when the birds drum infrequently or irregularly, making this species an excellent candidate for remote acoustic sensing with automated recording units (ARUs). An accurate automated recognition method for ruffed grouse drumming could enable effective and efficient use of ARU data for monitoring efforts; however, no such tool is currently available. Here we develop an automated method for detecting ruffed grouse drumming in audio recordings. Our detector uses a signal processing pipeline designed to recognize the accelerating pattern of drumming. We show that the automated recognition method accurately and efficiently detects drumming events in a set of labeled ARU field recordings. In a case study with 56 locations in Central Pennsylvania, we compared detections of ruffed grouse from 4 survey methods: field-based acoustic drumming surveys, surveys conducted by humans listening to ARU recordings, and automated recognition for both a 1-day and

68-7482-15-501; National Fish and Wildlife Foundation, Grant/Award Number: 0407.18.059680; Department of Biological Sciences at the University of Pittsburgh

a 28-day period. Field-based surveys detected drumming at 9 of 56 locations (16%), while surveys conducted by humans listening to ARU recordings detected drumming at 8 locations (14%). Using automated recognition, the 1-day recording period produced detections at 17 locations (30%) and the 28-day recording period produced detections at 34 locations (61%). Our case study supports the idea that automated recognition can unlock the value of ARU datasets by temporally expanding the survey period. We provide an open-source Python implementation of the recognition method to support further use in ruffed grouse monitoring efforts.

KEYWORDS

acoustic monitoring, ARU, automated recognition, *Bonasa umbellus*, drumming, machine learning, ruffed grouse, signal processing

The ruffed grouse (*Bonasa umbellus*) is a popular and declining game bird species with a wide range in North America (Rusch et al. 2020). Ruffed grouse inhabit deciduous and coniferous forests with dense understories, especially those dominated by aspen (*Populus* spp.; Rusch et al. 2020). Ruffed grouse select open, early successional forests for nesting and brood rearing (Turnbull et al. 2004). Ruffed grouse typically nest in regenerating forest 6–15 years post-disturbance until the canopy closes and wood stem density and ground vegetation diminishes (Sharp 1963, Dessecker and McAuley 2001). Visual detection of ruffed grouse in dense understories is difficult, but acoustic detection of males during the breeding season is comparatively easy. Breeding males advertise their presence to females and other males by beating their wings to produce an accelerating sequence of low-pitched thumps known as drumming (Rusch et al. 2020).

Ruffed grouse populations have declined throughout the Appalachian Mountains region in recent decades (−1.5%/yr; 95% CI = −3.1 – 0; Stauffer et al. 2018). Pennsylvania's second Breeding Bird Atlas similarly reported a reduction in the species' range, especially in the western portion of the Commonwealth (Wilson et al. 2012). Despite its popularity as a game bird, hunting is not thought to play a dominant role in population declines (Rusch et al. 2020), and annual hunter harvests in Pennsylvania have decreased steadily from 180,000 in 1998 to 10,000 in 2019 (Johnson 2020). Rather, ruffed grouse population declines have been linked to 3 factors; habitat loss via forest maturation, habitat loss or degradation due to fragmentation and land conversion, and the introduction of West Nile Virus (Porter and Jarzyna 2013, Stauffer et al. 2018).

The ruffed grouse's popularity as a game species and its apparent population decline have spurred conservation agencies and their partners to implement habitat management programs. In many cases, habitat management takes the form of timber harvesting that allows the regeneration of understory, thus providing ephemeral nesting habitat for female ruffed grouse (Rusch et al. 2020). In some regions that lack appropriate timber markets, habitat management programs employ alternatives for creating early successional forest communities such as non-commercial practices and private lands incentive programs (Litvaitis 2003, Litvaitis et al. 2021). Given the potentially confounding impacts of West Nile Virus and limited habitat availability, identifying where such habitat management efforts have the greatest potential for success will require monitoring ruffed grouse at both stand and regional scales.

Acoustic drumming surveys are commonly used to monitor for ruffed grouse because, unlike alternative survey methods (Hansen et al. 2015), they are rapid and non-consumptive. Indeed, drumming surveys have become the standard survey method for ruffed grouse employed by a variety of agencies and organizations across many parts of the species' range (Zimmerman 2006, Dougherty 2008, Hansen et al. 2011, Stewart and Sargent 2017, Witecha et al. 2019). Due to the practical constraints of surveying large areas, drumming surveys are conducted during peak breeding periods and typically consist of either (1) short listening bouts (i.e., a few minutes) at predetermined road-side locations (Petraborg et al. 1953) such as those conducted during the North American Breeding Bird Survey (Sauer et al. 2017), or (2) walking a transect along a road while listening for drumming (Jones et al. 2005). Both survey designs can fail to detect ruffed grouse presence because individual drumming activity can vary from day to day and with time of day (i.e., variable availability for detection). Additionally, an observer's ability to perceive drumming can be influenced by wind and rain conditions (Zimmerman and Gutiérrez 2007). Finally, many drumming surveys occur along roads and may provide biased estimates of ruffed grouse presence across a landscape (Betts et al. 2007). The shortcomings of field based surveys combined with the relatively distinct spectral characteristics of ruffed grouse drumming make them an excellent candidate for automated acoustic monitoring.

Automated recording units (ARUs) may overcome challenges associated with surveying for ruffed grouse by increasing temporal sampling effort (Borker et al. 2015, Colbert et al. 2015, Brauer et al. 2016, Darras et al. 2019, Gibb et al. 2019). Recently, the availability of low-cost ARUs such as the AudioMoth (Hill et al. 2019) has lowered the barrier to performing acoustic surveys at large spatial scales. In comparison to a typical ruffed grouse drumming survey effort of 2 5- to 6-min point counts per location per season (Petraborg et al. 1953), an AudioMoth might record for 90 mins per day on 60 consecutive days, increasing the total sampling time by more than 2 orders of magnitude. Intuitively, the probability of acoustically detecting target species is likely to increase both with the amount of time surveyed each day and with the number of survey dates (McNeil et al. 2014, Wood et al. 2021). For ruffed grouse in particular, increased survey effort could overcome issues of infrequent or inconsistent drumming and of suboptimal weather conditions for detecting drumming. Moreover, deploying ARUs across landscapes could mitigate potential survey bias associated with road-based monitoring, as long as survey points are chosen with a rigorous, randomized sampling design (Smith et al. 2017).

The value of an ARU sampling approach, however, is limited by the ability to efficiently extract species detections from audio data. Since deployments of ARUs can easily accumulate thousands of hours of audio recordings in a season, manual review of audio data from such deployments is seldom practical. Instead, researchers must either select relatively small portions of audio to manually review or use automated recognition methods to more efficiently analyze large datasets (Knight et al. 2017, Gibb et al. 2019). In general, open-source automated recognition methods for automated acoustic monitoring are not currently available for most taxa, and face many challenges such as the lack of labeled audio recordings and lack of generalizability (Knight et al. 2017, Stowell 2022). As a result, most ecological studies still rely on manual review of a limited subset of ARU recordings to extract species presence information (Venier et al. 2012, Sidie-Slettedahl et al. 2015, Vold et al. 2017, Sugai et al. 2019, Wilhite et al. 2020). The use of automated recognition methods to process large quantities of ARU data can greatly increase the number of detections relative to annotation effort even when recognizer precision is poor (Shonfield 2018, Knight et al. 2022). Thus, while manual review of ARU recordings can provide useful information to complement data from field-based surveys, the expanded temporal survey periods of ARUs become most valuable when used in combination with an automated recognition tool.

As true for most taxa, none of the currently available methods for automated recognition of ruffed grouse have been validated on ARU datasets. Template-based cross-correlation (also called template matching) of spectrograms (Goëau et al. 2014) could be used to automate detection of drumming in audio, but this approach has been replaced by machine learning approaches (Stowell 2022) because it is computationally expensive and generalizes poorly across individual and geographic variation. Machine learning methods for bird sounds have improved in recent years thanks to the use of deep learning algorithms, but still perform with low accuracy on soundscape recordings where vocalizations may be distant and overlapping (Kahl et al. 2020, Stowell 2022). The BirdNET classifier (Kahl et al. 2021) is the only published machine learning method we are aware of that automates the recognition of ruffed grouse presence in audio recordings. BirdNET is a supervised deep learning classifier that uses methods adapted from image recognition to detect bird vocalizations in a spectrogram, a visual frequency versus time representation of audio (Figure 1).

We developed an open-source automated recognition method for identifying ruffed grouse drumming events in audio recordings. Our automated recognition method recognizes drumming events through signal processing, rather than with template-based cross-correlation or deep learning. We test the accuracy of the detector on a labeled set of audio recordings and compare its performance to the BirdNET classifier. We also conduct a pilot study with ARUs at 56 locations in Pennsylvania and compare detections across 4 survey methods: field-based human point counts, ARU-based human point counts, and automated recognition on ARU data from 1-day and 28-day periods.

STUDY AREA

Our study area consisted of 3 State Game Lands managed by the Pennsylvania Game Commission (PGC) in north central Pennsylvania, USA (see Parker et al. 2020 for a detailed description of study area). Each State Game Land (State Game Lands 033, 060, and 100, see Pennsylvania Game Commission 2022) was dominated by forest (88–94%), comprised of mostly dry-oak (*Quercus* spp.) heath, dry-oak mixed-hardwood, and red maple (*Acer rubrum*) cover types. Elevations for the study sites ranged from 530 to 762 meters. We assessed ruffed grouse drumming in 10 regenerating stands where overstory removal timber harvests occurred between 2009 and 2012 (meaning that 8 to 11 years had elapsed since harvest at the time of our pilot study).

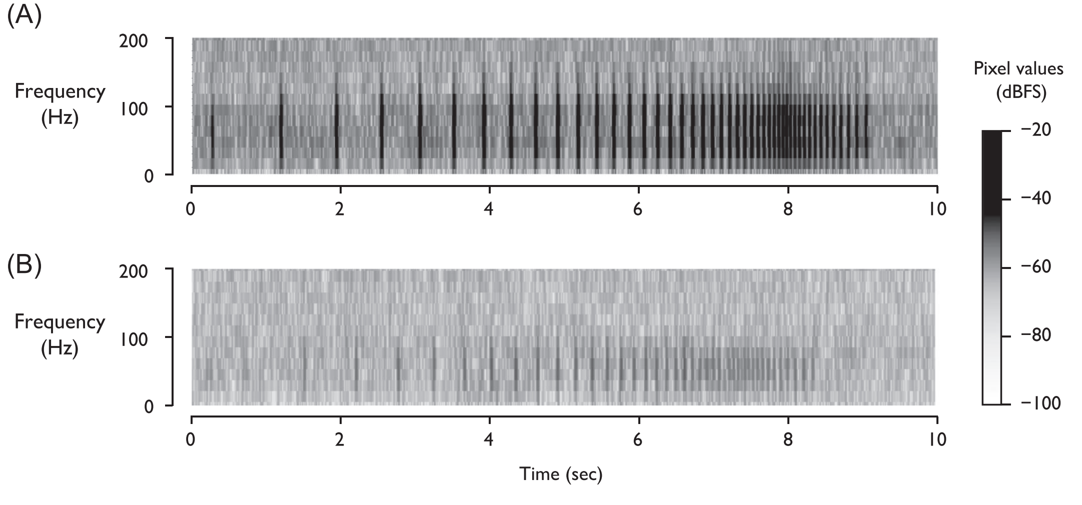

FIGURE 1 Spectrograms of ruffed grouse drumming events detected by the automated method: (A) close event, audible on headphones; (B) distant event, not audible but visible on spectrogram. Darker pixels indicate more energy at a given frequency and time.

METHODS

We deployed 60 AudioMoth acoustic recorders within the 10 regenerating timber harvests selected for our case study. Survey locations were spaced at least 200 meters apart. AudioMoths were programmed to record audio at a 32 kHz sampling rate for 3 30 min periods each morning and recorded for 4 weeks during peak ruffed grouse courtship, from 19 April to 16 May 2020. Four devices malfunctioned or were lost during the deployment (e.g., water or black bear [*Ursus americanus*] damage), resulting in 56 locations with audio recordings distributed across the 3 State Game Lands (Figure 2).

We surveyed for ruffed grouse at each of the 56 locations using 4 methods which reflect realistic alternatives for a monitoring effort:

1. Field-based human point counts: We conducted a single 10-min acoustic drumming survey at each ARU location between 15–17 May 2020 (Ralph et al. 1995, Bibby et al. 2000). We will refer to these acoustic drumming surveys as field-based human point counts to distinguish them from the alternate methods described below. We completed all surveys between 15 minutes and 4 hours post-sunrise. Each survey was conducted by a single observer. Upon arriving at each point, observers waited 1 minute before initiating the survey to allow the birds to settle down after potentially being disturbed (Bibby et al. 2000). Surveys were not conducted in unfavorable weather conditions, such as strong wind (>4 on Beaufort wind index) or rain.
2. Automated recording unit-based human point counts: We performed 5 min acoustic drumming surveys on 4 dates at each location by listening to the audio recorded by ARUs at each location and annotating all drumming events using Raven Pro (Bioacoustics Research Program 2019). The 4 ARU-based human point count dates were spaced across the ruffed grouse monitoring season (25 April, 1 May, 7 May, and 16 May) and contained audio from 07:40–07:45 Eastern Daylight Time. This approach improves consistency of monitoring by using ARU recordings to survey each location simultaneously. All annotations of ruffed grouse drumming were cross evaluated by a second reviewer for accuracy. The ARU-based human point count effort resulted in a total of 224 annotated 5-minute audio clips.
3. Automated recognition for a 1-day period: We developed an automated recognition method for recognizing ruffed grouse drumming in ARU recordings. We then ran the automated procedure on AudioMoth recordings acquired on 16 May 2020 from each survey location and performed a manual verification process which we also describe below. Because we annotated the period from 07:40–07:45 on this date during ARU-based human point counts and used the labels to tune model parameters, we excluded those 5-minute segments from this analysis. We analyzed the remaining 85 mins of audio for each location. To verify drumming detections from the automated method, we considered all sequences detected by our method candidate detections and extracted

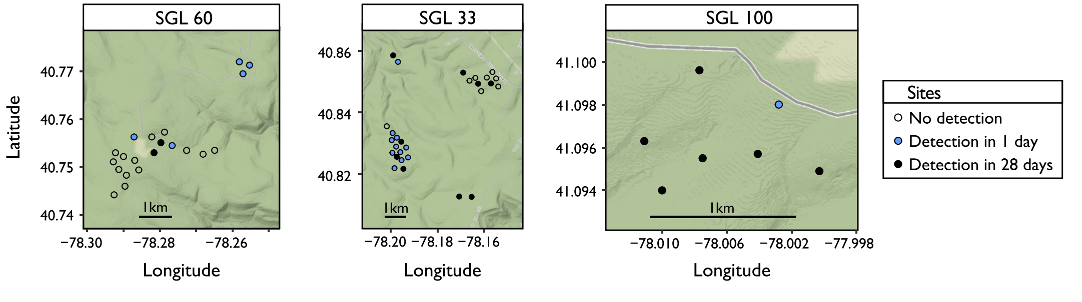

FIGURE 2 The 56 survey locations were distributed across 3 State Game Lands (SGLs). Survey locations are marked in blue if we verified an automated detection during the 1-day survey period and in black if we verified a detection during the 28-day survey but not during the 1-day survey. Points without a detection are marked with an empty circle.

10 second audio clips centered on each candidate detection. We manually reviewed the 10 second audio clips by listening to the audio and visually inspecting a spectrogram of the low frequencies (0–200 Hz). For each location, we sorted the candidate detections by the number of points in the detected sequence (from most to least) and reviewed the audio clips until (a) the location had a human-verified detection, or (b) we had reviewed all candidate detections for the location. All verified detections were additionally reviewed by a second annotator for accuracy. This process resulted in a set of manually verified ruffed grouse detections across the 56 locations.

4. Automated recognition for a 28-day period: We repeated the automated recognition procedure (3), but included recordings from the entire 28-day ARU survey period (19 April to 16 May 2020). Again, audio segments that were part of the validation set were excluded from this analysis. This resulted in a set of manually verified ruffed grouse detections from automated detections on 2,515 minutes of audio (90*28 – 5) for each site.

Developing an automated recognition procedure

We developed an automated recognition method to detect ruffed grouse drumming events in audio recordings using a series of signal processing operations. Our automated recognition method searches for the distinctive structure of low-pitched, accelerating thumps (each of which is a wingbeat) that make up a drumming display (Figure 1). During analysis of audio data, recordings were split into 1-minute clips, and each clip was analyzed using this recognition method. Here, we describe the 4 signal processing steps used in the automated recognition method.

First, we used a continuous wavelet transform (Jordan et al. 1997) with a Morlet wavelet tuned to 50 Hz to extract the low frequency band where ruffed grouse drumming is strongest while removing most background noise (Figure 3A,B). The continuous wavelet transform retains high time resolution and isolates a specific frequency of interest, resulting in a one-dimensional signal (Addison 2018). Because we were only interested in low-pitched sounds, we downsampled the audio sampling rate from 32,000 Hz to 400 Hz before performing the continuous wavelet transform. Downsampling the audio greatly reduced the computational cost of the continuous wavelet transform operation. When we applied the continuous wavelet transform to an audio signal containing a drumming event, the spikes in the resulting signal corresponded to individual thumps from the bird's wingbeats (Figure 3B). Next, we processed the resulting signal to better isolate the spikes from ruffed grouse drumming (Figure 3C). This involved normalizing the signal (dividing by the maximum value), squaring the signal to create positive values, taking the magnitude of the Hilbert transform to trace the amplitude modulation envelope (as demonstrated in Smith 2021), and normalizing the signal once again. Next, we located the temporal positions of strong spikes in the signal (Figure 3C, blue dots) using the peak-finding algorithm scipy.signal.find_peaks from the Python package SciPy (version 1.5.1, Virtanen et al. 2020). We generated a 2-dimensional feature y(t) consisting of the times of detected peaks (t) and the forward difference of time between the same detected peaks (y) (Figure 3D). Because the individual thumps of a ruffed grouse drum accelerate in time, peaks extracted from a drumming event created the smooth downward trend in a plot of y versus t (Figure 3D). We observed that this downward trend in a plot of forward time difference between detected peaks (y) versus time (t) was distinctive to ruffed grouse drumming events and was unlikely to occur by chance in low frequency background noise (see examples in Supporting Information). Then, we applied an algorithm to detect drumming events in the y(t) feature. The algorithm, inspired by bat detection filters (Britzke and Murray 2000), searches for a consistent downward trend in the y versus t plot that matches the shape of the feature we observed for drumming events (Figure 3D). Progressing iteratively through the points in the y(t) feature, the algorithm adds points to a candidate drumming sequence if they meet a set of conditions relative to the preceding points in the candidate sequence (Table 1, Conditions A-C). Points that do not meet these conditions are considered invalid and are not added to the candidate sequence. The candidate sequence is terminated when there are over 3 consecutive invalid points (Table 1, Condition D) or over 0.8 seconds have elapsed since the previous point in the sequence (Table 1, Condition E). The candidate sequence is considered

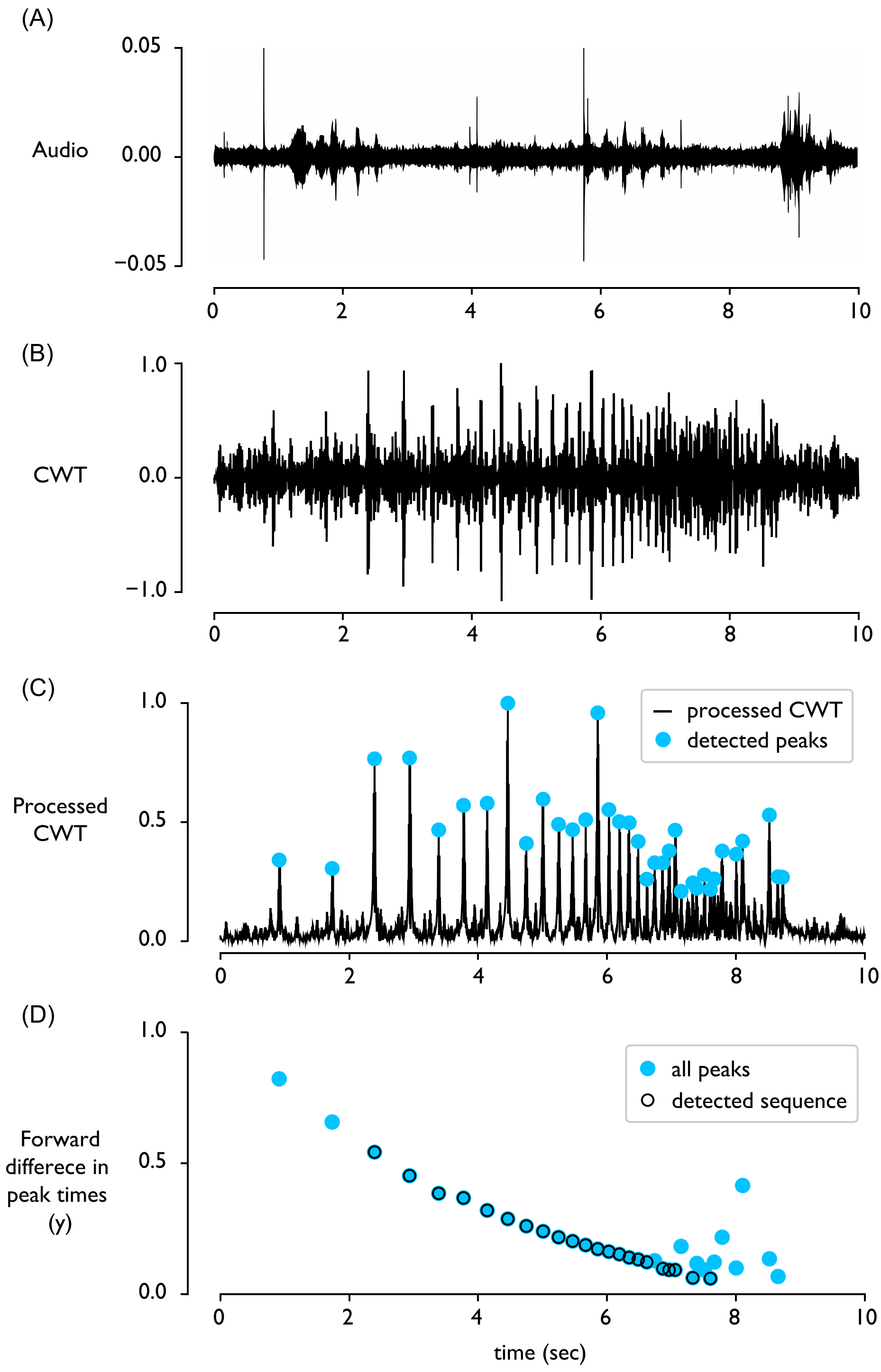

FIGURE 3 Drumming events in an audio signal were detected in 4 steps: a continuous wavelet transform (CWT) isolated the 50 Hz content from an audio signal (A, B); normalizing, squaring, and smoothing the signal isolated strong peaks (C); peak-detection localized peaks in time (c, blue dots); drumming events were detected based on their smooth downward curve in a plot of the forward difference of peak times, y, versus peak time (D).

TABLE 1 Conditions used in ruffed grouse drum detection algorithm.

Condition	Description	Value(s)	Purpose
A	Time elapsed from current point to most recent point included in candidate sequence (i.e., Δt)	0.05 to 0.8 sec	Avoid very short or very long intervals between peaks
B	Difference in y between current point and most recent point in candidate sequence (i.e., Δ*y*)	−0.2 to 0.0 sec	Sequence should accelerate
C	Difference in Δ*y* between current and most recent point in candidate sequence	−0.05 to 0.15 sec	Sequence should not suddenly change acceleration
D	Number of consecutive points that can be skipped within a sequence	3 points	Allow sequence to continue past limited number of invalid points
E	Maximum time between consecutive points in a sequence	0.8 sec	After a long pause, begin a new sequence
F	Duration of entire sequence	1 to 15 sec	Match fragment or entire drum
G	Number of points in entire sequence	9 to 100 points	Match fragment or entire drum

a detection if it is of an appropriate length to represent an entire drumming event or a substantial fraction of a drumming event (Table 1, Conditions F-G), and is otherwise discarded. The process of searching for drumming sequences then continues with the subsequent point and an empty candidate sequence. The algorithm successfully detects drumming events in the y(t) feature, even when the original audio signal has significantly louder concurrent events (Figure 3D).

We chose the parameters used in the algorithm to reflect observed patterns in ruffed grouse drumming events. We refined parameter values by iteratively applying the algorithm to a validation dataset (see below), inspecting examples of false positives and false negatives, and modifying parameters to improve performance. We discuss parameter tuning in detail in the Appendix and provide a Python implementation of the algorithm in the Supporting Information. The algorithm is also implemented in the open-source bioacoustics package OpenSoundscape (Lapp et al. 2022) as find_accel_sequences.

Evaluating detection methods

We used the annotated audio data from the ARU-based human point counts as a validation set to measure the accuracy of our automated recognition method and of BirdNET. We split the 224 labeled 5-minute audio files into 1,120 1-min clips and labeled ruffed grouse present in all 1-minute clips that contained an annotation for ruffed grouse drumming. We refer to the created set of labeled clips as the validation set. Our automated recognition method predicted that ruffed grouse was present in a 1-minute audio clip if it detected at least one drumming event in the clip. We evaluated the performance of BirdNET and our automated recognition method on the presence or absence of ruffed grouse in 1-minute audio clips in the validation set using the metrics recommended by Knight et al. (2017): precision (fraction of detections that truly contained drumming), recall (fraction of all drumming events that were captured), and F1 score (harmonic mean of precision and recall). We could not calculate areas under the receiver-operating curve or precision-recall curve for our method because it does not output continuous scores.

The BirdNET classifier uses a convolutional neural network to detect the presence of 500 North American bird species (Kahl et al. 2021). To evaluate the performance of BirdNET for detecting ruffed grouse, we downloaded and installed the BirdNET model and ran predictions on our labeled set of audio files according to instructions on the GitHub repository (Kahl 2022). By default, BirdNET produces confidence scores for the presence of a species on 3-second audio clips. To generate predictions for 1-minute clips with BirdNET, we took the maximum score of any 3-second clip across each 1-minute file. We chose the prediction threshold for BirdNET that maximized the F1 score on the validation set. We then used the case study as a measure of our method's utility for real-world monitoring efforts. We compared the number of locations with ruffed grouse detections for each of the 4 survey methods: (1) field-based human point counts, (2) ARU-based human point counts, (3) automated recognition on a 1-day period, and (4) automated recognition on a 28-day period. We note that ruffed grouse drumming may be detectable at distances greater than 100 meters, and therefore a single drumming event could potentially be detected at multiple survey locations. Furthermore, during the 28-day period ruffed grouse may have moved between the survey locations. Thus, as movements would violate the assumptions of a single-season occupancy model (MacKenzie et al. 2018), the number of detections in our case study is not an estimate of occupancy.

RESULTS

Our automated recognition method accurately detected drumming in the validation set. In the validation set, manual labeling determined that a total of 37 out of 1,120 1-minute clips contained ruffed grouse drumming. Our method successfully detected drumming in 29 of these 37 clips (78% recall) and returned 12 false-positive detections (71% precision, F1 score of 74%). We note that every 5-minute continuous segment of annotated data in the validation set that had an annotation also had a verified automated detection; in other words, recall at the 5-minute level was 100%. The BirdNET classifier (Kahl et al. 2021) performed poorly on the validation set, achieving a maximum F1 score of 21% using a threshold score of 0.43. At a threshold score of 0.43, BirdNET detected ruffed grouse in 5 of 37 clips with drumming (13% recall) and produced 5 false-positive detections (50% precision).

In our case study, the automated recognition method detected drumming at more locations than both field- and ARU-based human point counts. Field-based human point counts detected ruffed grouse at 9 of 56 locations (16%), while ARU-based human point counts detected ruffed grouse at 8 locations (14%). On 16 May 2020 our automated recognition procedure resulted in human-verified detections at 17 locations (30%), including 8 of the 9 locations with detections during field-based human point counts. The remaining location produced 2 possible detections that contained low frequency drum-like patterns, but we could not verify that they were ruffed grouse drumming events. Automated recognition on ARU data from the 28-day survey period resulted in human-verified detections at 34 of 56 locations (61%), including all 9 locations with detections during field-based human point counts.

Analyzing 4,703 30-minute audio clips (2,351 hours) from the ARU data with the automated recognition method took 75 minutes on a parallelized computing system with 28 parallelized CPU tasks and generated 3,418 candidate detections. By following our manual review procedure for candidate detections starting with the longest sequences, we only needed to manually inspect 97 10-second clips (0.01% of audio data). After reviewing this small subset of candidate detections, every location had either (1) a human-verified detection or (2) no further unreviewed detections.

DISCUSSION

The accuracy of our method, combined with the results of the case study, suggests that the ruffed grouse recognition method presented here could serve as an effective tool for detecting drumming in temporally expansive ARU datasets. Using automated recognition on 1 day of recordings, which increased the survey effort from

10 minutes of field-based detection to 85 minutes of detection on recorded audio, nearly doubled the number of locations with detections from 9 to 17. Furthermore, expanding the survey period from 1 day to 28 days doubled the number of sites with detections again, for a total of 34 locations with detections. Our results support the idea that using ARU data can increase detections of a target species by expanding the survey period on 2 temporal scales: the amount of time surveyed on a single day and the number of days surveyed. While achieving 1 of these 2 temporal expansions may be feasible using manual review of audio, achieving both without the help of automated recognition would be cost ineffective, requiring thousands of hours of manual annotation. Neither temporal expansion during a day nor temporal expansion across multiple days are particularly feasible with traditional point count sampling. Therefore, in scenarios where a focal species produces temporally scarce audible events, automated recognition may be essential for extracting species detections from large ARU datasets.

We manually inspected false negatives in the validation set to investigate why our automated recognition method failed to recognize some drumming events. Out of 10 1-minute clips annotated with ruffed grouse that did not have an automated detection, 2 had drumming events partially cut off at the beginning or end of the clip. The issue of missing drumming at the ends could be resolved by using overlapping analysis windows, though the results would need to be interpreted with caution because single drumming events could be detected multiple times. The other 8 false-negative clips contained faint drumming events, often with substantial background noise. Four of the 8 had detected series of 6 or 7 peaks, slightly short of the threshold of 9 peaks for a detection. The false negatives represent the outer limit of the radius of detection for this study's combination of ARUs, environmental conditions, and automated recognition method.

To investigate whether the same drumming events were detected at multiple recorders, we examined annotations in the validation set for detections that may have been the same event. Because ARU clocks can drift by an estimated several seconds per month, we examined all annotations whose start times were within 10 seconds of each other. We found 4 instances in the validation set where annotated grouse drumming events occurred at approximately the same time. Two of the events came from a set of 3 recorders, each spaced less than 260 meters apart from the others, and could represent multiple detections of the same event or counter-drumming between individuals. The other 2 instances were from recorders spaced over 5 kilometers apart and presumably occurred due to random chance rather than 2 recorders detecting the same event. Notably, we did not detect simultaneous drumming events on other combinations of recorders despite there being 33 pairs of recorders with spacing between 200 and 300 meters for which at least 1 of the 2 recorders had a drumming detection.

Measuring the area surveyed by an ARU is important for calibrating results of monitoring between ARUs and human surveyors (Van Wilgenburg et al. 2017, Yip et al. 2017, Darras et al. 2018) and can aid in calculating species distribution (Leseberg et al. 2022) or density if site-level heterogeneity is modeled (Darras et al. 2016). Future work should experimentally quantify the sampling area covered by ARU surveys and field-based human point counts by measuring the effective detection radius of ruffed grouse drumming for each detection method (Darras et al. 2018), potentially also modeling detection as a function of covariates. In contrast to most avian vocalizations (Shonfield and Bayne 2017), the effective detection radius for ruffed grouse drumming may be larger for ARU recordings than for field-based human point counts because it contains energy at very low frequencies. The peak frequency in ruffed grouse drumming (45 Hz) is near the lower bound of human hearing, and human hearing progressively decreases in sensitivity for low frequencies (Suzuki and Takeshima 2004). Additionally, higher frequencies are attenuated more than lower frequencies as sound travels through air (Stokes' law of sound attenuation, Stokes 1880) and as a result, as a sound source becomes progressively farther away, only the lower frequencies reach the listener. For ruffed grouse drumming, the combination of these 2 phenomena can mean that very distant drums (with only the low frequency content remaining) are inaudible to the human ear but can still be captured by an ARU and seen on a spectrogram. We noted that during manual review of audio for the validation set, many of the drumming events detected by the automated recognition method were inaudible when played back over closed-back headphones but could be seen in the lowest frequencies of a spectrogram (Figure 1B, for example).

While many contemporary automated recognition approaches in bioacoustics use convolutional neural networks or other supervised deep learning approaches (Stowell et al. 2019, Kahl et al. 2020), developers of a new bioacoustic recognition method might choose to use signal processing rather than supervised deep learning for 2 reasons. First, supervised learning requires large amounts of labeled audio data for training, which is not available for most taxa besides North American and European birds. Second, the features of a signal processing detector are hand-built and therefore easily interpretable, while features synthesized by deep learning approaches can be difficult to interpret. For example, detected sequences from our method that contain 15 versus 30 points are directly interpretable as containing fewer or more wingbeats from a drumming event, while a low or high score produced by a deep learning method cannot be related directly to any specific quality of the original audio event. As a result, though both deep learning and signal processing approaches can result in false-positive and false-negative detections, the reason for misdetections may be easier to interpret when using a signal processing approach. Additionally, because signal processing approaches often home in on specific distinguishing characteristics of a signal, such as the low frequency accelerating pattern of drumming, they can sometimes lead to highly accurate detection methods that outperform deep learning approaches (Lapp et al. 2021).

Signal processing methods have some disadvantages compared to deep learning approaches (reviewed by Stowell 2022). Deep learning algorithms can be trained to detect many classes at a time rather than only a single class, enabling the possibility of community-level analyses (Stowell et al. 2019, Kahl et al. 2020). Additionally, a rule-based approach like ours enforces a specific structure for the target vocalization, which could be problematic if the pattern of drumming differs substantially across individuals or across regions (Garcia et al. 2012). Deep learning algorithms such as convolutional neural networks can learn high-level representations of audio that incorporate variability, potentially allowing them to generalize better to new scenarios (Stowell et al. 2019). We note, however, that both signal processing and deep learning approaches are likely to decrease in accuracy or fail completely when the data analyzed is substantially different from the data used for training and validation. Finally, deep learning approaches can provide continuous outputs rather than discretized or binary detections, which could be useful for interpreting and statistically analyzing the outputs of a model (Knight et al. 2017, Rhinehart et al. 2020, Leseberg et al. 2022). Alternatives to supervised deep learning and signal processing include supervised shallow learning (relying on hand-crafted features; Priyadarshani et al. 2018) and unsupervised learning (Stowell and Plumbley 2014) which is often used in tandem with supervised learning.

We believe that effectively isolating the signal of ruffed grouse drumming from low-frequency noise was critical to the success of our detection method. Because the BirdNET classifier is designed to simultaneously detect several hundred bird species (including passerines with high-frequency calls), it is not optimized for extracting low frequencies. The BirdNET classifier created a spectrogram that spanned a range of 150 Hz to 15 kHz. The peak frequencies of drumming (around 45 Hz; Garcia et al. 2012) were excluded and low frequencies were combined into the lowest pixels of the spectrogram, which likely made it difficult for the classifier to detect distant drumming events in the presence of other low-frequency sounds. In contrast, by using a continuous wavelet transform, our method isolated the frequency band of highest energy for ruffed grouse drumming and generally achieved a high signal to noise ratio even in the presence of low-frequency noise (see Figure 3A–C). This difference in feature representation helps to explain why BirdNET performed poorly on the validation set while our automated recognition method performed well.

Wavelets are useful feature extraction tools that have been widely adopted in various fields from physics to economics (Addison 2018) and have also been applied to bioacoustics problems such as denoising (Xie et al. 2021), pattern recognition (Bardeli et al. 2010), and feature extraction (Colonna et al. 2012, Kiskin et al. 2020). Sibul (1999) reviews wavelet transforms and their application to bioacoustics signal processing. The continuous wavelet transform is computationally expensive compared to the often-used Fourier transform, which can often extract similar information from a signal. However, wavelets can achieve higher time resolution than the Fourier transform and are useful when trying to isolate a specific frequency from a noisy signal (Addison 2018). In our case, we greatly reduced the computational cost of continuous wavelet analysis by downsampling audio from 32,000 Hz to 400 Hz, which was possible because we were only interested in low frequency sounds. Additionally, we only performed a single continuous wavelet transform tuned to 50 Hz, whereas creating wavelet features for multiple frequencies requires one continuous wavelet transform per frequency. We recommend considering continuous wavelet transforms as bioacoustic feature extraction tools when the computational costs can be avoided or when spectrograms cannot provide sufficient time resolution at a desired frequency resolution.

In the broader context of automated acoustic monitoring, the detection method we presented here could be adapted with minor changes to detect acoustic signals from other species. For instance, the algorithm could be easily modified to detect decelerating rather than accelerating patterns. Notably, the method will increase linearly in computational cost with the frequency of the target vocalization because the continuous wavelet transform operation will need to be performed on audio with a sampling rate of at least twice the target frequency. For vocalizations that have a steady periodic structure that does not speed up or slow down, the RIBBIT method (Lapp et al. 2021) of detecting sounds with periodic amplitude modulation may be more effective, though experimentation is warranted on a species-by-species basis. Both the RIBBIT method and the method presented here are available in OpenSoundscape (Lapp et al. 2022*b*) for open-source use and modification.

We highlight 3 considerations for the application of automated acoustic monitoring and automated recognition to ruffed grouse monitoring efforts. First, the entire process from planning a study to deployment, pick up, data management, analysis, and reporting should be considered before beginning an automated acoustic monitoring project. The manual effort required to analyze or verify the results of a recognizer will depend on the type of analysis required. For occupancy analyses, it is sufficient to verify just one detection per site and sampling period, while other analyses may require more manual effort. Second, and relatedly, the assumptions of occupancy modeling (or other models that will be used to analyze resulting data) should inform the choice of survey locations. Specifically, survey locations should be distributed with a minimum spacing of twice the maximum estimated radius of detection to minimize the chances of detecting a single individual at multiple points. Even with sufficient spacing to avoid detecting the same drumming event at multiple recorders, the assumptions of an occupancy model may be violated during a season; in particular, the assumption of closure may not hold if individual ruffed grouse move across the landscape. Third, the total cost of a monitoring effort should be considered when choosing between in-person point counts, automated acoustic monitoring, or a combination of the 2. Deploying and picking up automated recording equipment requires 2 visits to each survey point per season, but unlike acoustic drumming surveys, these visits can be conducted by untrained personnel, for instance by mailing ARUs to private landowners or agency land managers.

RESEARCH IMPLICATIONS

We developed an automated ruffed grouse recognition method that accurately extracts drumming events from ARU audio recordings. Our results highlight the value of automated recognition methods combined with automated acoustic monitoring for species that produce temporally scarce audible events. We provide open-source code for the ruffed grouse recognition method in the bioacoustics Python package OpenSoundscape (Lapp et al. 2022*b*) to support its use in research and monitoring programs. Monitoring how ruffed grouse populations react to various management approaches will be critical to the success of conservation efforts. By deploying ARUs across a landscape and using this automated method to detect drumming events, conservation practitioners can implement cost-effective and large-scale ruffed grouse monitor programs. Considering the difficulty of monitoring this species by traditional means, the opportunity to expand survey efforts temporally and spatially using this tool could be central to the success of ruffed grouse conservation agendas.

ACKNOWLEDGMENTS

We thank T. Rhinehart, L. Freeland-Haynes, R. P. Lyon, L. Katsis, and L. Chronister for comments on earlier drafts of this manuscript. This material is based upon work supported by the National Science Foundation under Grants 1935507 and 2120084. This research was also supported in part by the University of Pittsburgh Center for Research Computing through the resources provided. Finally, we thank the Forestry and Land Management staff of the Pennsylvania Game Commission's Northcentral Region for allowing access to study sites and assisting with project logistics (PGC permit 36850). Our work was financially supported by the Department of Biological Sciences at the University of Pittsburgh, the Gordon and Betty Moore Foundation, and the National Fish and Wildlife Foundation's Central Appalachia Habitat Stewardship Program. This research was funded by Natural Resource Conservation Service's Conservation Effects Assessment Project (No. 68-7482-15-501), National Fish and Wildlife Foundation (No. 0407.18.059680), and Indiana University of Pennsylvania's School of Graduate Studies and Research. We thank D. Dahlgren (Associate Editor), A. Knipps (Editorial Assistant), A. Tunstall (Copy Editor) and J. Levengood (Content Editor) and 2 anonymous reviewers for constructive comments that improved our manuscript.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

ETHICS STATEMENT

Institutional approval was not required for avian point count surveys or automated recording unit deployments. No animals were captured, handled, or harmed in this observational study.

DATA AVAILABILITY STATEMENT

The 5-minute audio clips and corresponding annotations for the validation set are available in an online repository (Lapp et al. 2022*a*). An online GitHub repository (Lapp 2022) provides Python scripts to reproduce the analyses in this manuscript using the open-source library OpenSoundscape (Lapp et al. 2022*b*). The complete audio dataset from the study is available from the corresponding author upon reasonable request.

ORCID

Samuel Lapp http://orcid.org/0000-0003-1637-6822

REFERENCES

Addison, P. S. 2018. Introduction to redundancy rules: the continuous wavelet transform comes of age. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 376:20170258.

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

Betts, M. G., D. Mitchell, A. W. Diamond, and J. Béty. 2007. Uneven rates of landscape change as a source of bias in roadside wildlife surveys. Journal of Wildlife Management 71:2266–2273.

Bibby, C., N. Burgess, D. Hill, and S. Mustoe. 2000. Bird census techniques. Second. Academic Press, London.

Bioacoustics Research Program. 2019. Raven Pro: Interactive Sound Analysis Software. Center for Conservation Bioacoustics, Cornell Lab of Ornithology, Ithaca, NY, USA. <https://ravensoundsoftware.com/software/raven-pro/>. Accessed 13 Jul 2020.

Borker, A. L., P. Halbert, M. W. Mckown, B. R. Tershy, and D. A. Croll. 2015. A comparison of automated and traditional monitoring techniques for marbled murrelets using passive acoustic sensors. Wildlife Society Bulletin 39:813–818.

Brauer, C. L., T. M. Donovan, R. M. Mickey, J. Katz, and B. R. Mitchell. 2016. A comparison of acoustic monitoring methods for common anurans of the northeastern United States. Wildlife Society Bulletin 40:140–149.

Britzke, E., and K. Murray. 2000. A quantitative method for selection of identifiable search-phase calls using the Anabat system. Bat Research News 41.

Colbert, D. S., J. A. Ruttinger, M. Streich, M. Chamberlain, L. M. Conner, and R. J. Warren. 2015. Application of autonomous recording units to monitor gobbling activity by wild turkey. Wildlife Society Bulletin 39:757–763.

Colonna, J. G., A. D. Ribas, E. M. dos Santos, and E. F. Nakamura. 2012. Feature subset selection for automatically classifying anuran calls using sensor networks. Pages 1–8 *in* Proceedings of the 2012 International Joint Conference on Neural Networks. IEEE, Brisbane, Australia.

Darras, K., P. Batáry, B. J. Furnas, I. Grass, Y. A. Mulyani, and T. Tscharntke. 2019. Autonomous sound recording outperforms human observation for sampling birds: a systematic map and user guide. Ecological Applications e01954.

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

Darras, K., P. Pütz, Fahrurrozi, K. Rembold, and T. Tscharntke. 2016. Measuring sound detection spaces for acoustic animal sampling and monitoring. Biological Conservation 201:29–37.

Dessecker, D. R., and D. G. McAuley. 2001. Importance of early successional habitat to ruffed grouse and American woodcock. Wildlife Society Bulletin 29:456–465.

Dougherty, E. M. 2008. Ruffed grouse drumming counts: an examination of observer and roadside effects. Dissertation, The Ohio State University.

Garcia, M., I. Charrier, D. Rendall, and A. N. Iwaniuk. 2012. Temporal and spectral analyses reveal individual variation in a non-vocal acoustic display: the drumming display of the ruffed grouse (*Bonasa umbellus*). Ethology 118:292–301.

Gibb, R., E. Browning, P. Glover-Kapfer, and K. E. Jones. 2019. Emerging opportunities and challenges for passive acoustics in ecological assessment and monitoring. Methods in Ecology and Evolution 10:169–185.

Goëau, H., H. Glotin, W.-P. Vellinga, R. Planqué, A. Rauber, and A. Joly. 2014. LifeCLEF bird identification task 2014. Pages 585–597 *in* CLEF: Conference and Labs of the Evaluation Forum. Volume CEUR Workshop Proceedings. Sheffield, United Kingdom.

Hansen, C. P., J. J. Millspaugh, and M. A. Rumble. 2011. Occupancy modeling of ruffed grouse in the Black Hills National Forest. Journal of Wildlife Management 75:71–77.

Hansen, M. C., C. A. Hagen, D. A. Budeau, V. L. Coggins, and B. S. Reishus. 2015. Comparison of three surveys for estimating forest grouse population trends. Wildlife Society Bulletin 39:197–202.

Hill, A. P., P. Prince, J. L. Snaddon, C. P. Doncaster, and A. Rogers. 2019. AudioMoth: A low-cost acoustic device for monitoring biodiversity and the environment. HardwareX 6:e00073.

Johnson, K. E. 2020. A multi-regional assessment of factors influencing American Woodcock use of managed early successional communities. Thesis, Indiana University of Pennsylvania, Indiana, PA, USA.

Jones, B., C. Harper, D. Buehler, and G. Warburton. 2005. Use of spring drumming counts to index ruffed grouse populations in the southern Appalachians. Pages 135–143 *in* The proceedings of the Annual Conference of the Southeastern Association of Fish and Wildlife Agencies. Volume 59. St. Louis, MO, USA.

Jordan, D., R. W. Miksad, and E. J. Powers. 1997. Implementation of the continuous wavelet transform for digital time series analysis. Review of Scientific Instruments 68:1484–1494.

Kahl, S. 2022. BirdNET. Python. <https://github.com/kahst/BirdNET>. Accessed 6 May 2022.

Kahl, S., M. Clapp, W. Hopping, H. Goëau, R. Planqué, W.-P. Vellinga, and A. Joly. 2020. Overview of BirdCLEF 2020: Bird Sound Recognition in Complex Acoustic Environments. CLEF: Conference and Labs of the Evaluation. Thessaloniki, Greece.

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

Kiskin, I., D. Zilli, Y. Li, M. Sinka, K. Willis, and S. Roberts. 2020. Bioacoustic detection with wavelet-conditioned convolutional neural networks. Neural Computing and Applications 32:915–927.

Knight, E. C., K. C. Hannah, G. J. Foley, C. D. Scott, R. M. Brigham, and E. Bayne. 2017. Recommendations for acoustic recognizer performance assessment with application to five common automated signal recognition programs. Avian Conservation and Ecology 12:14.

Knight, E., K. Hannah, and J. DeMoor. 2022. In the still of the night: revisiting Eastern Whip-poor-will surveys with passive acoustic monitoring. Avian Conservation and Ecology 17.

Lapp, S. 2022. GitHub: ruffed grouse manuscript 2022. Zenodo. <https://github.com/kitzeslab/ruffed_grouse_manuscript_2022>. Accessed 15 Sep 2022.

Lapp, S., H. Parker, and C. Tett. 2022*a*. ARU audio recordings with ruffed grouse annotations (Pennsylvania, 2020). Dryad. <http://datadryad.org/stash/dataset/doi:10.5061/dryad.hdr7sqvmc>. Accessed 15 Sep 2022.

Lapp, S., T. Rhinehart, L. Freeland-Haynes, and J. Kitzes. 2022*b*. OpenSoundscape version 0.5.1. <opensoundscape.org>.

Lapp, S., T. Wu, C. Richards-Zawacki, J. Voyles, K. M. Rodriguez, H. Shamon, and J. Kitzes. 2021. Automated detection of frog calls and choruses by pulse repetition rate. Conservation Biology 35:1659–1668.

Leseberg, N. P., W. N. Venables, S. A. Murphy, N. A. Jackett, and J. E. M. Watson. 2022. Accounting for both automated recording unit detection space and signal recognition performance in acoustic surveys: A protocol applied to the cryptic and critically endangered Night Parrot (Pezoporus occidentalis). Austral Ecology 47:440–455.

Litvaitis, J. A. 2003. Are pre-Columbian conditions relevant baselines for managed forests in the northeastern United States? Forest Ecology and Management 185:113–126.

Litvaitis, J. A., J. L. Larkin, D. J. McNeil, D. Keirstead, and B. Costanzo. 2021. Addressing the early-successional habitat needs of at-risk species on privately owned lands in the eastern United States. Land 10:1116.

MacKenzie, D. I., J. D. Nichols, J. A. Royle, K. H. Pollock, L. L. Bailey, and J. E. Hines. 2018. Occupancy estimation and modeling. Elsevier, San Diego, California, USA.

McNeil, D. J., Jr., C. R. V. Otto, and G. J. Roloff. 2014. Using audio lures to improve golden-winged warbler (Vermivora chrysoptera) detection during point-count surveys. Wildlife Society Bulletin 38:586–590.

Parker, H. A., J. T. Larkin, D. Heggenstaller, J. Duchamp, M. C. Tyree, C. S. Rushing, E. Just Domoto, and J. L. Larkin. 2020. Evaluating the impacts of white-tailed deer (Odocoileus virginianus) browsing on vegetation in fenced and unfenced timber harvests. Forest Ecology and Management 473:118326.

Pennsylvania Game Commission. 2022. State Game Lands PDF Maps. Pennsylvania Game Commission. <https://www.pgc.pa.gov:443/HuntTrap/StateGameLands/Pages/State-Game-Lands-Maps.aspx>. Accessed 18 Apr 2022.

Petraborg, W. H., E. G. Wellein, and V. E. Gunvalson. 1953. Roadside drumming counts as a spring census method for ruffed grouse. Journal of Wildlife Management 17:292–295.

Porter, W. F., and M. A. Jarzyna. 2013. Effects of landscape-scale forest change on the range contraction of ruffed grouse in New York State, USA. Wildlife Society Bulletin 37:198–208.

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

Ralph, C. J., J. R. Sauer, and S. Droege. 1995. Monitoring bird populations by point counts. Pacific Southwest Research Station.

Rhinehart, T. A., L. M. Chronister, T. Devlin, and J. Kitzes. 2020. Acoustic localization of terrestrial wildlife: current practices and future opportunities. Ecology and Evolution 10:6794–6818.

Rusch, D. H., S. Destefano, M. C. Reynolds, and D. Lauten. 2020. Ruffed Grouse (Bonasa umbellus). S. M. Billerman, B. K. Keeney, P. G. Rodewald, and T. S. Schulenberg, editors. Birds of the World. Cornell Lab of Ornithology. <https://birdsoftheworld.org/bow/species/rufgro/1.0/introduction>. Accessed 5 Oct 2021.

Sauer, J. R., K. L. Pardieck, D. J. Ziolkowski, A. C. Smith, M.-A. R. Hudson, V. Rodriguez, H. Berlanga, D. K. Niven, and W. A. Link. 2017. The first 50 years of the North American Breeding Bird Survey. The Condor 119:576–593.

Sharp, W. M. 1963. The effects of habitat manipulation and forest succession on ruffed grouse. Journal of Wildlife Management 27:664–671.

Shonfield, J. 2018. Using Bioacoustics to Examine the Effects of Industrial Disturbance on Owls and their Prey. Dissertation, University of Alberta.

Shonfield, J., and E. M. Bayne. 2017. Autonomous recording units in avian ecological research: current use and future applications. Avian Conservation and Ecology 12:42–54.

Sibul, L. H. 1999. Wavelet transforms for bioacoustic signal processing. Journal of the Acoustical Society of America 106:2129.

Sidie-Slettedahl, A. M., K. C. Jensen, R. R. Johnson, T. W. Arnold, J. E. Austin, and J. D. Stafford. 2015. Evaluation of autonomous recording units for detecting 3 species of secretive marsh birds: Marsh Bird Surveys. Wildlife Society Bulletin 39:626–634.

Smith, A. N. H., M. J. Anderson, and M. D. M. Pawley. 2017. Could ecologists be more random? Straightforward alternatives to haphazard spatial sampling. Ecography 40:1251–1255.

Smith, J. O. 2021. Analytic signals and Hilbert transform filters. Mathematics of the Discrete Fourier Transform (DFT). W3K Publishing. <https://ccrma.stanford.edu/~jos/r320/Analytic_Signals_Hilbert_Transform.html>

Stauffer, G. E., D. A. W. Miller, L. M. Williams, and J. Brown. 2018. Ruffed grouse population declines after introduction of West Nile virus. Journal of Wildlife Management 82:165–172.

Stewart, C. A., and L. Sargent. 2017. Ruffed grouse and american woodcock status in Michigan, 2017. 16.

Stokes, G. 1880. On the theories of the internal friction of fluids in motion, and of the equilibrium and motion of elastic solids. Pages 75–129 *in* Mathematical and Physical Papers vol. 1 Cambridge University Press, Cambridge, United Kingdom.

Stowell, D. 2022. Computational bioacoustics with deep learning: a review and roadmap. PeerJ 10:e13152.

Stowell, D., and M. D. Plumbley. 2014. Automatic large-scale classification of bird sounds is strongly improved by unsupervised feature learning. PeerJ 2:e488.

Stowell, D., M. D. Wood, H. Pamuła, Y. Stylianou, and H. Glotin. 2019. Automatic acoustic detection of birds through deep learning: The first Bird Audio Detection challenge. D. Orme, editor. Methods in Ecology and Evolution 10:368–380.

Sugai, L. S. M., T. S. F. Silva, J. W. Ribeiro, and D. Llusia. 2019. Terrestrial Passive Acoustic Monitoring: Review and Perspectives. BioScience 69:15–25.

Suzuki, Y., and H. Takeshima. 2004. Equal-loudness-level contours for pure tones. Journal of the Acoustical Society of America 116:918–933.

Turnbull, C., J. Clark, T. Shenk, T. Nicolette, and K. Hockenberry. 2004. Survey of Ruffed Grouse drumming in Central Pennsylvania. 4.

Van Wilgenburg, S. L., P. Sólymos, K. J. Kardynal, and M. D. Frey. 2017. Paired sampling standardizes point count data from humans and acoustic recorders. Avian Conservation and Ecology 12.

Venier, L. A., S. B. Holmes, G. W. Holborn, K. A. Mcilwrick, and G. Brown. 2012. Evaluation of an automated recording device for monitoring forest birds. Wildlife Society Bulletin 36:30–39.

Virtanen, P., R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, et al. 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nature Methods 17:261–272.

Vold, S. T., C. M. Handel, and L. B. McNew. 2017. Comparison of acoustic recorders and field observers for monitoring tundra bird communities. Wildlife Society Bulletin 41:566–576.

Wilhite, N. G., P. E. Howell, and J. A. Martin. 2020. Evaluation of Acoustic Recording Devices to Survey Northern Bobwhite Populations. Wildlife Society Bulletin 44:200–207.

Wilson, A. M., D. W. Brauning, and R. S. Mulvihill. 2012. Second atlas of breeding birds in Pennsylvania. Pennsylvania State University Press.

Witecha, M., B. Dhuey, A. Gerrits, C. Pollentier, A. Schnieder, and J. Steigerwaldt. 2019. Wisconsin ruffed grouse management plan 2020-2030. Wisconsin Department of Natural Resources.

Wood, C. M., S. Kahl, P. Chaon, M. Z. Peery, and H. Klinck. 2021. Survey coverage, recording duration and community composition affect observed species richness in passive acoustic surveys. Methods in Ecology and Evolution 12: 885–896.

Xie, J., J. G. Colonna, and J. Zhang. 2021. Bioacoustic signal denoising: a review. Artificial Intelligence Review 54: 3575–3597.

Yip, D. A., L. Leston, E. M. Bayne, P. Sólymos, and A. Grover. 2017. Experimentally derived detection distances from audio recordings and human observers enable integrated analysis of point count data. Avian Conservation and Ecology 12.

Zimmerman, G. S. 2006. Habitat and population regulation of ruffed grouse in Minnesota. Dissertation, University of Minnesota, Minneapolis, MN.

Zimmerman, G. S., and R. J. Gutiérrez. 2007. The influence of ecological factors on detecting drumming ruffed grouse. Journal of Wildlife Management 71:1765–1772.

Associate Editor: D. Dahlgren.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's website. Supporting information includes (1) a Python notebook with a demonstration of the recognition method and (2) an implementation of the drum detection algorithm in Python.

How to cite this article: Lapp, S., J. L. Larkin, H. A. Parker, J. T. Larkin, D. R. Shaffer, C. Tett, D. J. McNeil, C. J. Fiss, and J. Kitzes. 2023. Automated recognition of ruffed grouse drumming in field recordings. Wildlife Society Bulletin 47:e1395. https://doi.org/10.1002/wsb.1395

APPENDIX A: PARAMETER TUNING FOR THE RECOGNITION METHOD

We chose the parameters used in the drum detection algorithm to reflect observed patterns in ruffed grouse drumming events. We refined parameter values by iteratively applying the algorithm to a validation dataset (see below), inspecting examples of false positives and false negatives, and modifying parameters to improve performance. The reasoning behind each parameter choice is detailed below. Once we had tuned the parameters of the algorithm, it successfully detected drumming events in the y(t) feature (Figure 3D) while ignoring other low-frequency noises (see example analysis in Supporting Information).

The conditions and corresponding values used in the algorithm are listed in Table 1. Conditions A, B, and C evaluate whether each subsequent point in y(t) is valid. If the point is valid, it is added to the current candidate sequence. Condition A constrains the backwards difference (Δt) of time between consecutive points in a candidate series. In Condition A, the upper limit of Δt ensures that peaks far apart in time are not considered part of the same sequence, while the lower limit of Δt avoids counting very closely spaced double-peaks as independent. Note that Δt, the backwards difference between a candidate point and the latest point included in the candidate series, is distinct from y, the forward difference between all detected peaks. The algorithm uses y as a pre-generated feature vector but calculates Δt during the iterative search for sequences. Condition B constrains the backwards difference of y (Δy) between consecutive points in a candidate series. Condition B confines Δy to small negative values because each subsequent y value for the peaks in an accelerating drumming sequence will be slightly lower than the last, but not greatly lower. Condition C constrains the second backwards difference of y between points in a candidate series. We added Condition C in addition to Conditions A and B because we observed that some false detections showed sharp upward or downward curvature on a plot of y versus t, while drumming events consistently show a slightly positive second backward difference of y (e.g., Figure 3D).

Conditions D and E determine when a candidate sequence is terminated and a new sequence is started. Condition D, which limits the number of points that can be skipped over between 2 consecutive points, was set to 3 to allow some noisy intermediate peaks to be skipped (for instance, see Figure 3D at 7.0 seconds) without allowing excessive noise. Similarly, Condition E limits the amount of time between peaks in a sequence because after a long pause, peaks are unlikely to belong to the same sequence.

Conditions F and G constrain the sequence to lengths matching all or at least a significant fragment of a ruffed grouse drumming event. Drumming events typically last 9–10 seconds and contain 30–40 accelerating wingbeats. After the accelerating portion of the drumming display there is an additional sequence of about 10 wingbeats that slow down (Figure 3C, 7–9 seconds), but we designed our detector to ignore this part of the drumming event. Sequences lasting longer than 15 seconds or with more than 100 peaks are likely to be false positives caused by machinery or other periodic low-frequency noise, while sequences lasting under 1 second or with less than 9 peaks could be caused by any number of low frequency events.

We also chose the 1-minute clip length as input to the algorithm by experimenting with different clip lengths and testing performance on the validation set. Because the continuous wavelet transform signal is normalized during our post-processing steps, changing the length of the audio clips fed into the detector could affect the performance. Specifically, longer audio clips are more likely to contain a loud event, which would reduce the levels of other events in the normalized signal. Therefore, we would expect the recognition algorithm's recall to decrease with increasing audio clip length and we would expect its precision to decrease with decreasing audio clip length.