Automated Recording Unit Pilot Project

SUMMARY

We tested the effectiveness of Automated Recording Units (ARUs) – devices which record audio at a preset schedule – for monitoring birds within Carnegie SVRA. We deployed ARUs at 44 distinct sampling locations during spring 2018 and also surveyed points using traditional in-person point counts. We annotated bird species in the recordings collected during the 10 minute surveys manually and also using automated software known as BirdNET. We found that under low and medium wind conditions ARU collected data was comparable to human collected point count data. In fact, slightly more species were detected during the manual annotation of recordings compared to human point counts, perhaps because recordings could be listened to multiple times as opposed to human point counts which require observers to record all species that are singing simultaneously. The BirdNET software performed relatively well at correctly identifying species when it detected a species on the recording, however it often missed vocalizing bird species completely when we set a stringent criteria for positively identifying a species. It seems likely that ARUs could be useful for monitoring birds at SVRAs, however there is still room for improvement for the automated bird annotation software. We look forward to continuing to test ARU technology at more SVRAs in the future

INTRODUCTION

California State Parks, Off-Highway Motor Vehicle Recreation Division (OHMVRD) administers nine parks known as State Vehicular Recreation Areas (SVRAs) in California that are open to off-highway vehicular recreation. The OHMVRD is required to monitor wildlife populations within each SVRA to ensure that habitat for wildlife is maintained or improved. As part of a larger project analyzing bird monitoring data collected by SVRAs we were tasked with examining new approaches for measuring disturbance from off-highway vehicles and the response of birds to that disturbance.

As part of our approach to evaluating bird response to off-highway vehicle use, we implemented a pilot project assessing the practicality of using automated recording units (ARUs). ARUs are devices that can be programmed to record sounds during pre-defined time periods, and as a result show potential for semi-automated monitoring of birds within an SVRA. Successful ARU deployment could address the difficulty SVRAs have had in finding skilled observers for their wildlife monitoring, as many parks are located in areas that are distant from large cities where skilled birders may reside. With ARUs, staff that are not expert bird observers can deploy the units and send the recordings to be annotated remotely, thus removing the need for an observer to be physically present at the park. Additionally, ARUs could potentially be used for recording the number of OHVs passing each sampling point and the degree of sound disturbance birds experience at a given sampling point (e.g, as measured by decibels of sound).

We implemented a pilot program at Carnegie SVRA that paired in-person point counts with counts of birds derived from human-annotated recordings collected simultaneously by ARUs deployed at the same points. The objective was to determine if ARUs could provide bird survey data that was comparable to traditional in-person point counts. Additionally, software for automatic annotation of bird recordings (BirdNET; https://birdnet.cornell.edu/) very recently became available. We also evaluated the effectiveness of BirdNET at correctly identifying bird species on the annotated recordings.

METHODS

Study Area

Carnegie SVRA is comprised of 5,015 acres within San Joaquin and Alameda counties of California. The habitat within the park is varied and consists of blue oak (*Quercus douglasii*)

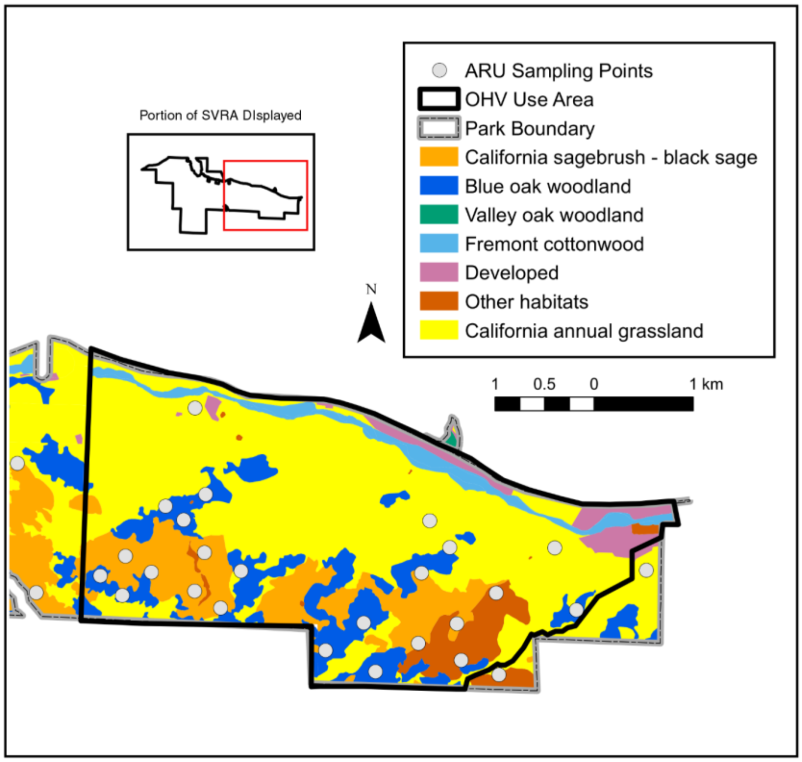

Figure 1. Map of habitat classes within the OHV riding portion of Carnegie SVRA. Sampling points where ARUs were deployed are gray filled circles. The inset in the top left uses a red rectangle to denote which portion of the park is displayed in the main figure.

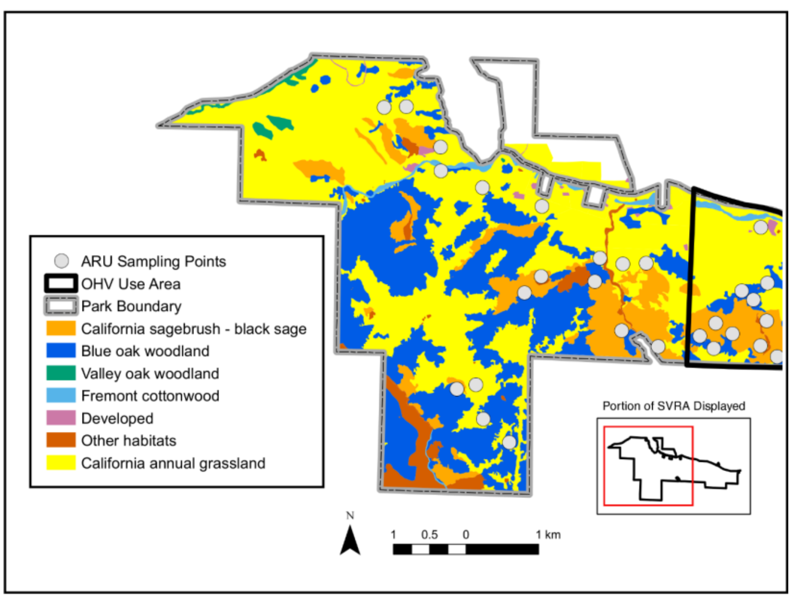

Figure 2. Map of habitat classes within the portion of Carnegie SVRA that is closed to OHV use (enclosed within the gray polygon. Sampling points where ARUs were deployed are gray filled circles. The bottom right inset uses a red rectangle to denote which portion of the park is displayed in the main figure. Habitat classes that were not delineated for the northern portion of the park are left blank.

and valley oak (*Quercus lobate*) woodland, California annual grassland, California sagebrush (*Artemisia californica*) - black sage (*Salvia mellifera*) mixed habitat, mule fat (*Baccharis salicifolia*) thickets, and riparian habitat primarily comprised of Fremont cottonwood (*Populus fremontii*). The majority of the park (~3,500 acres) is closed to OHV use (Figs. 1 and 2).

Point Count and ARU Surveys

In 2010, the staff of Carnegie SVRA used random stratified sampling to establish 114 Habitat Monitoring System (HMS) sampling points within OHV and non-OHV use areas. Points were stratified by four habitat types: oak woodland (blue and valley oak pooled), California annual grassland, California sagebrush – black sage, and riparian. We randomly selected a subset of 72 of the 114 established sampling points, stratified by OHV treatment to produce 36 in OHV and 36 in non-OHV use areas. Due to logistical constraints we were only able to visit 46 of those 72 points, of which we deployed recording units at 44 points (Figs. 1 and 2). Birds were surveyed using a distance sampling point count protocol (Buckland et al. 2015). All point counts were conducted by the same experienced observer (Jerry Cole), who recorded distances to all birds seen or heard during a 10 minute period. Each sampling point was visited once during the first three hours after sunrise (between 5:55 and 8:30 AM), and spanned the dates May 14 to May 24, 2018.

Automated Recording Units were set up at a sampling point the day prior to the in-person point count sampling and were programmed to begin recording 15 minutes prior to sunrise (i.e., at 5:40 AM) and cease recording at 11 AM. We used Song Meter 4 (SM4: Wildlife Acoustics, Inc., Maynard, MA, USA) units for this study. ARUs were secured to a steel t-post within a wooden housing that protected them from theft, and had openings so that microphones were not obstructed (Fig. 3). Units were installed approximately 1 m above the ground. We set the units to record at a sampling rate of 44.1 kHz because we wanted to capture the full range of bird sound (~ 22 kHz) and because the Nyquist Theorem states that audio must be captured at 2x the rate of the highest frequency of target sound to be accurately reproduced. All other ARU settings were kept at their default values. The point count observer stood near the ARU during the paired in-person survey so that the unit would record and the observer would hear the same bird vocalizations. The observer audibly announced the beginning of the 10 minute survey so that during annotation of the ARU

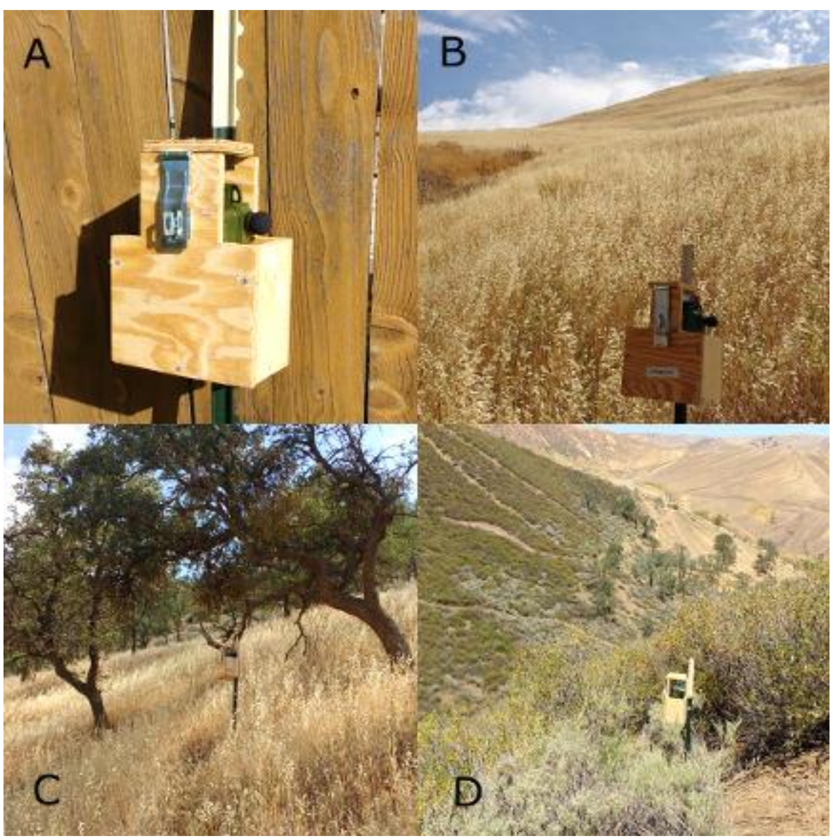

Figure 3. Automated recording unit (ARU) in security box attached to steel t-post (A). ARUs deployed at California annual grassland (B), oak woodland (C), and California sagebrush – black sage (D) sampling points.

recording the point count and the recording could be precisely matched.

ARU Recording Annotation

We annotated the 10 minute section of recordings that corresponded to the in-person point count survey of each point. We used a modified version of the annotation protocol developed by researchers at the Alberta Biodiversity Monitoring Institute (Lankau et al. 2015). The annotator listened to the 10 minute recording in 1 minute segments using the audio software Audacity (www.audacityteam.org) while also reviewing the visual representation of the sound, known as the sonogram (Fig. 4). The annotator could pause, review segments multiple times, and use reference recordings from eBird (https://ebird.org/media/catalog?mediaType=a) and Xeno-Canto (https://www.xeno-canto.org/) to confirm species' identities. Birds were marked as being present on a given 1 minute section of a recording if they vocalized any time during that period. If the observer was confident that multiple individuals of a species were present during a recording (e.g., two birds vocalizing nearly simultaneously or distant from one another in space) then those individuals were marked as separate individuals. When the observer progressed to the next segment of the recording (e.g., listening to minutes 1 - 2 of the recording after minutes 0 - 1 of the recording), only birds species that vocalized again were marked as present.

We also estimated the noise level of wind, rain, and other ambient sounds that may limit an observer's ability to hear bird vocalizations (e.g., highway traffic hum, rushing water, electric hum of high-voltage transmission lines) on a scale of 0 - 3 (Table 1). This system followed the categories developed by Lankau et al. (2015), with the exception that we lumped any noise other than wind or rain into a single category.

ARU Automated Recording Annotation

We used the recently developed BirdNET artificial neural network software (https://github.com/kahst/BirdNET) to automatically annotate the bird species vocalizing on 1 minute segments of the 10 minute human-annotated recordings. BirdNET is a prototype software for classifying 984 species of birds within North America and Europe. The software divides audio files into small time segments (default of 3 s), then generates a text file output of the species detected and the confidence in those

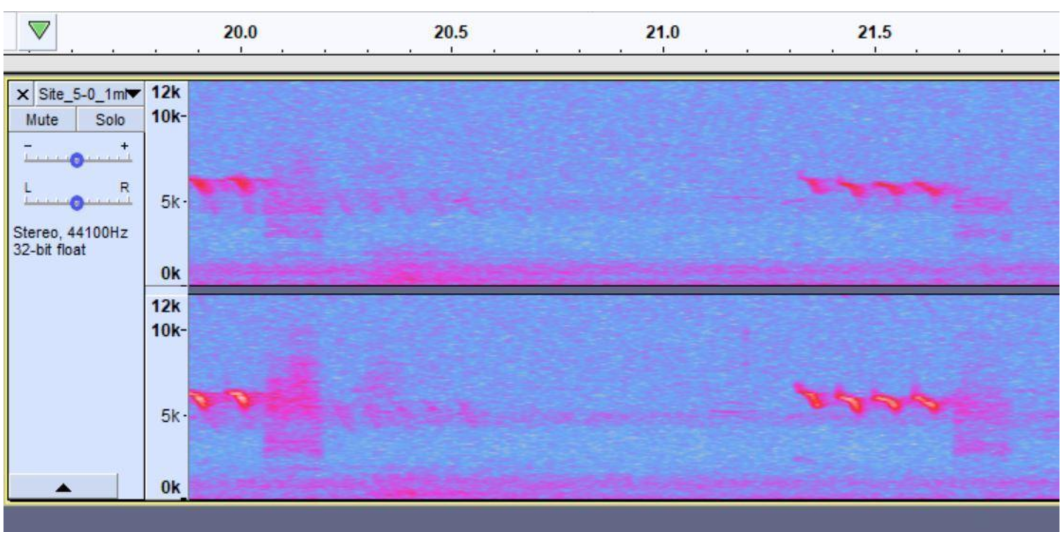

Figure 4. Screen capture of sound visual representation of Oat Titmouse vocalization in Audacity sound editing software from an ARU recording at Carnegie SVRA. Outputs from left and right microphones are represented in the top and bottom panels. Frequency of sound in kHz is represented on the y-axis in each panel, time elapsed in seconds is on the x-axis, and brightness of color indicates the intensity of sound in a given frequency.

Table 1. Categories of rain, wind, and other noise heard on ARU recordings. Modified from Lankau et al. (2015).

Noise Type	Code	Description
Rain	0	no rain
	1	affects the ability to hear distant/faint species, drops seldom hit microphones
	2	affects the ability to hear nearby species, drops often hit microphones
	3	significantly affects the ability to detect species, drops almost always hit microphones
Wind	0	no wind
	1	rustling leaves/trees creaking (background noise), affects ability to detect distant/faint species
	2	begins to muffle microphones (frequency and decibel rates begin to spike), occasionally affects ability to detect nearby species
	3	always muffles microphones, frequency and decibel graphs spike constantly (sometimes cuts out due to noise threshold)
Other	0	No noise
	1	affects detection of distant/faint species
	2	begins to affect detection of closer species
	3	species difficult to detect even if present

identifications ranging from 0 to 1. These annotations can be loaded into Audacity software to visualize when species were detected on each recording for further inspection (Fig. 5) or loaded as text files for analysis in the R statistical analysis environment (R Core Team 2017).

We ran BirdNET using all default settings, except we allowed time segments to overlap by 1.5 s (i.e., for a 9 s audio segment this means BirdNET will classify five 3 s segments, rather than three 3 s segments). We hypothesized that greater segment overlap would result in a higher probability of capturing a full vocalization of an individual bird. The birds detected by BirdNET were filtered post-hoc to include only species that were ever detected at Carnegie during spring human point counts from 2010-2019. We simplified the model output to a record of species' detection or non-detection for each 1 minute segment. The model output was collapsed this way to match the human-annotated recordings (e.g., noting a species presence during 1 minute periods).

Data Analysis

Comparison of point counts to human-annotated recordings

We calculated the correlation between point-level species richness derived from point count surveys and human-annotated recordings using a Spearman's rank correlation. We then compared point count and ARU-derived richness at sites with low (codes 0 and 1), medium (code 2), and high (code 3) wind using a two-tailed Wilcoxon signed rank test to determine if observed species richness was affected by wind noise. The Wilcoxon test only compares the differences in counts of species, and ignores sites with a difference of 0. We also compared counts of individual bird species that were detected at >15 points from point count surveys and ARU annotation using a two-tailed Wilcoxon signed rank test.

Comparison of human-annotated recordings to BirdNET automatic annotation

We summarized the number of species detected across all sites that had low and medium wind noise by both human

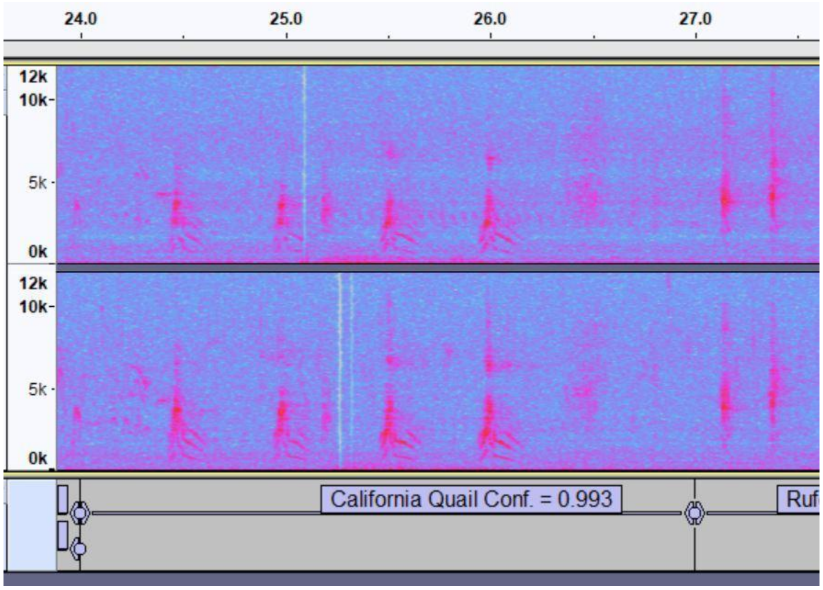

Figure 5. Screen capture of sound visual representation of California Quail vocalizations in Audacity sound editing software. Left and right audio channels are represented in the top and bottom panels. Bird species identity output from BirdNET for the 3 s segment is displayed below the audio, along with the model confidence in the species identity. Frequency of sound in kHz is represented on the y-axis in each panel, time elapsed in seconds is represented on the x-axis, and brightness of color indicates the intensity of sound in that frequency.

annotation and BirdNET annotation. We reduced the comparison sites to this smaller subset because excessive wind noise makes automated species identification difficult by obscuring bird vocalizations. We used two confidence level thresholds for filtering the BirdNET automatic annotation output: medium (> 0.65 model confidence) and high (> 0.85 model confidence) because we were interested in determining how model confidence level may affect the performance of BirdNET as a classifier. We used the industry standard metrics of precision and recall to assess the capacity of BirdNET to function as a practical bird sound identification system. We calculated precision as:

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives}$$

Where true positives are counts of instances when BirdNET identified a bird species as present on a 1 minute recording segment and the same bird species was also detected during human annotation. False positives are counts of instances where BirdNET identified a species as present during a 1 minute recording segment but the species was not detected during human annotation. We calculated recall as:

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives}$$

Where false negatives are counts of 1 minute segments that BirdNET did not detect a bird species but the species was detected during human annotation. It is possible that the observer annotating the recordings may have missed annotating a species or misidentified a species, but for this analysis we assumed that human annotations were free of errors. Precision and recall were calculated for each species individually.

RESULTS

We annotated 7 hrs and 40 minutes of recordings from ARU units deployed at 44 sampling points within Carnegie SVRA. The human annotation process yielded detections of 53 bird species on the ARU recordings, while the simultaneous in-person point counts detected 61 bird species. Combining results from both methods, we detected 66 bird species (see Appendix Table A1). Fifty bird species were detected with both survey methods (Appendix Table A1). Eleven species were detected only during in-person point counts and 5 species only during human annotation of ARU recordings (Table 2).

Table 2. Bird species detected only during in-person point count surveys or only during human annotation of ARU recordings.

Common Name	Scientific Name	Detected during point count	Annotated from ARU recording
Black-throated Gray Warbler	*Setophaga nigrescens*	X	
Brewer's Blackbird	*Euphagus cyanocephalus*	X	
Brown-headed Cowbird	*Molothrus ater*		X
Cassin's Vireo	*Vireo cassinii*	X	
Cliff Swallow	*Petrochelidon pyrrhonota*	X	
Hermit Warbler	*Setophaga occidentalis*	X	
Horned Lark	*Eremophila alpestris*	X	
Hutton's Vireo	*Vireo huttoni*	X	
Northern Mockingbird	*Mimus polyglottos*	X	
Purple Finch	*Haemorhous purpureus*		X
Savannah Sparrow	*Passerculus sandwichensis*		X
Song Sparrow	*Melospiza melodia*	X	
Steller's Jay	*Cyanocitta stelleri*	X	
Townsend's Warbler	*Setophaga townsendi*		X
Western Wood-Pewee	*Contopus sordidulus*		X
White-tailed Kite	*Elanus leucurus*	X	

Point-level species richness was significantly positively correlated between in-person point counts and human-annotated ARU counts (Spearman's rank correlation: rho = 0.80, $p < 0.001$, S = 2819.3; Fig. 6A). A rho value close to 1 indicates a near perfect positive correlation, a value closer to -1 a near perfect negative correlation, and a value close to 0 denotes no significant pattern of correlation. Point-level species richness derived from in-person point counts and human-annotated ARU recordings at sites with low and high wind noise were significantly different, while points with medium noise were not significantly different (Appendix Table A2). Fewer species were tallied during in-person point counts than on ARU recordings in low and medium wind conditions, and more during point counts under high wind conditions (Figure 6B).

For 17 bird species that were detected at >14 survey points, we compared point-level abundance derived from in-person versus human-annotated ARU surveys using a Wilcoxon signed rank test. Counts derived from in-person point counts and ARU recordings differed significantly for 5 of 17 species (see Appendix Table A3). Counts of California Towhee detected during in-person surveys were significantly lower, and counts of California Quail, House Finch, and Mourning Dove during in-person surveys were significantly higher, in comparison to counts derived from human-annotated ARU recordings (Fig. 7).

Across all point surveys with low to medium wind noise (Appendix Table A4), the majority of bird species detected during human annotation of recordings were also detected by BirdNET at least once. The software detected 42 of 51 species detected during human annotation (82-% of species) when annotations were filtered to segments labeled with >0.65 confidence by BirdNET. The software detected 39 of 51 species detected during human annotation (64.5% of species) when annotations were filtered to segments

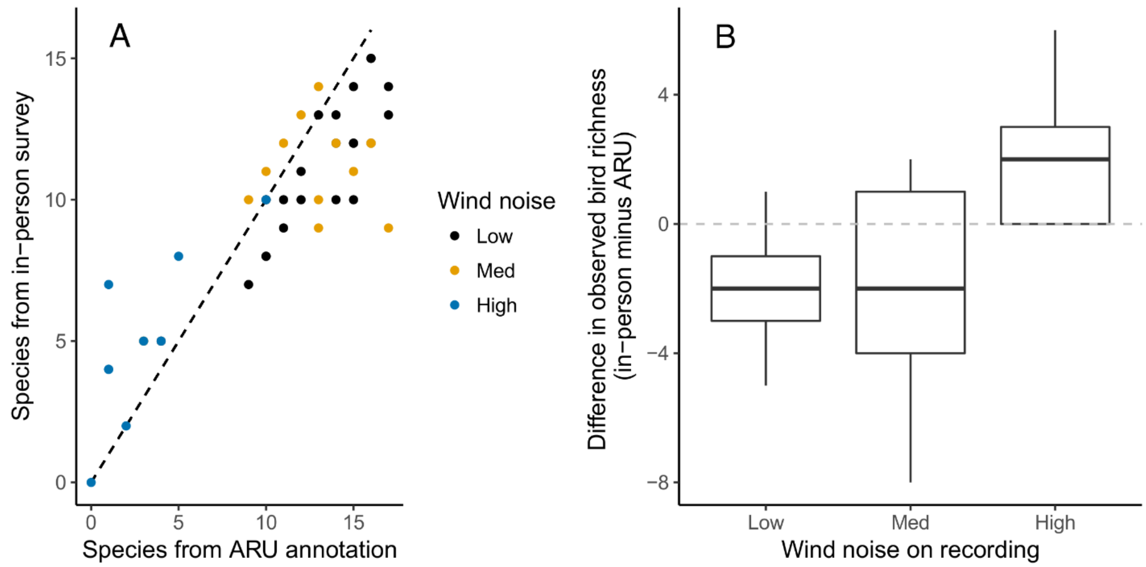

Figure 6. Point-level species richness derived from simultaneous in-person point count surveys and human-annotated ARU recordings of the same survey points (A). Diagonal dotted line represents at 1:1 relationship between richness derived from in-person point count richness and human-annotated ARU recordings. Boxplots of difference in point-level species richness (in-person point count species richness minus ARU derived species richness) under differing wind conditions (B). Horizontal dotted line represents a 1:1 relationship between in-person point count richness and ARU derived richness.

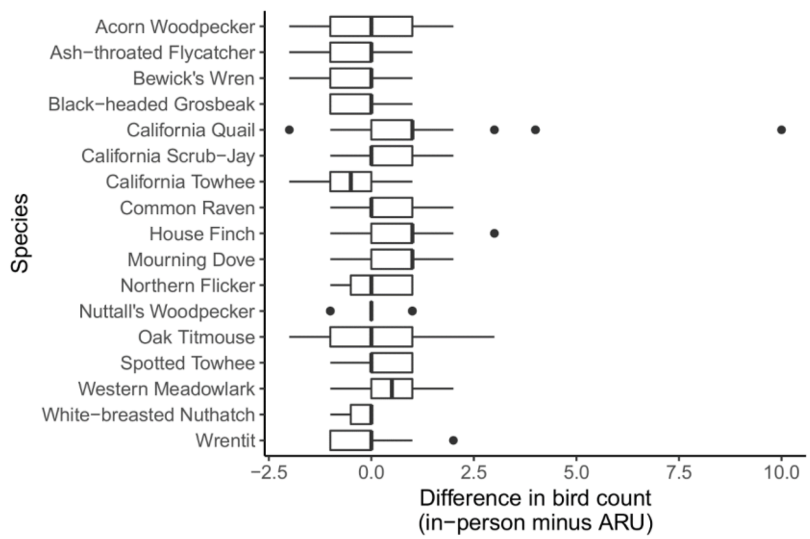

Figure 7. Difference in point-level counts of selected bird species derived from in-person point counts and human-annotated ARU recordings. Bold vertical bar for boxplot represents the median, the left and right thinner bars represent the first and third quartile respectively, the whiskers minimum and maximum, and dots are outliers for each species. Boxes to the left of zero indicate less birds were counted during in-person point counts than during human-annotation of ARU recordings.

labeled with >0.85 confidence by BirdNET. Additionally, the software identified 24 and 12 species not identified during human annotation or in-person point counts, for segments with >0.65 and >0.85 confidence respectively. The majority of species that were identified during human annotation of ARU recordings and during BirdNET automatic annotation had few instances of false positives (Fig. 8). However, because the high filtering threshold set for BirdNET species that were present at a location were frequently missed, rather than detected and assigned the incorrect bird species (Fig. 8 right panel).

Precision appeared to be modestly improved when we used a more stringent threshold for considering BirdNET detections to be valid (confidence of 0.65 versus 0.85) and recall was modestly reduced (proportion of false negatives; Fig. 8).

DISCUSSION

In our study human-annotated ARU recordings provided comparable, and often improved, estimates of bird species richness and abundance in low to medium wind conditions to in-person point counts. Indeed, more species were detected on ARU recordings than during paired passive point count surveys. Perhaps our annotated species richness was higher because the observer had the ability to review the recording multiple times and compare recorded sounds to databases of labeled high-quality bird sounds. For the majority of bird species there was no statistical difference in counts of individual birds detected during passive and ARU annotated recordings, suggesting that the ARUs we used provided a similar acoustic sampling radius to passive point counts.

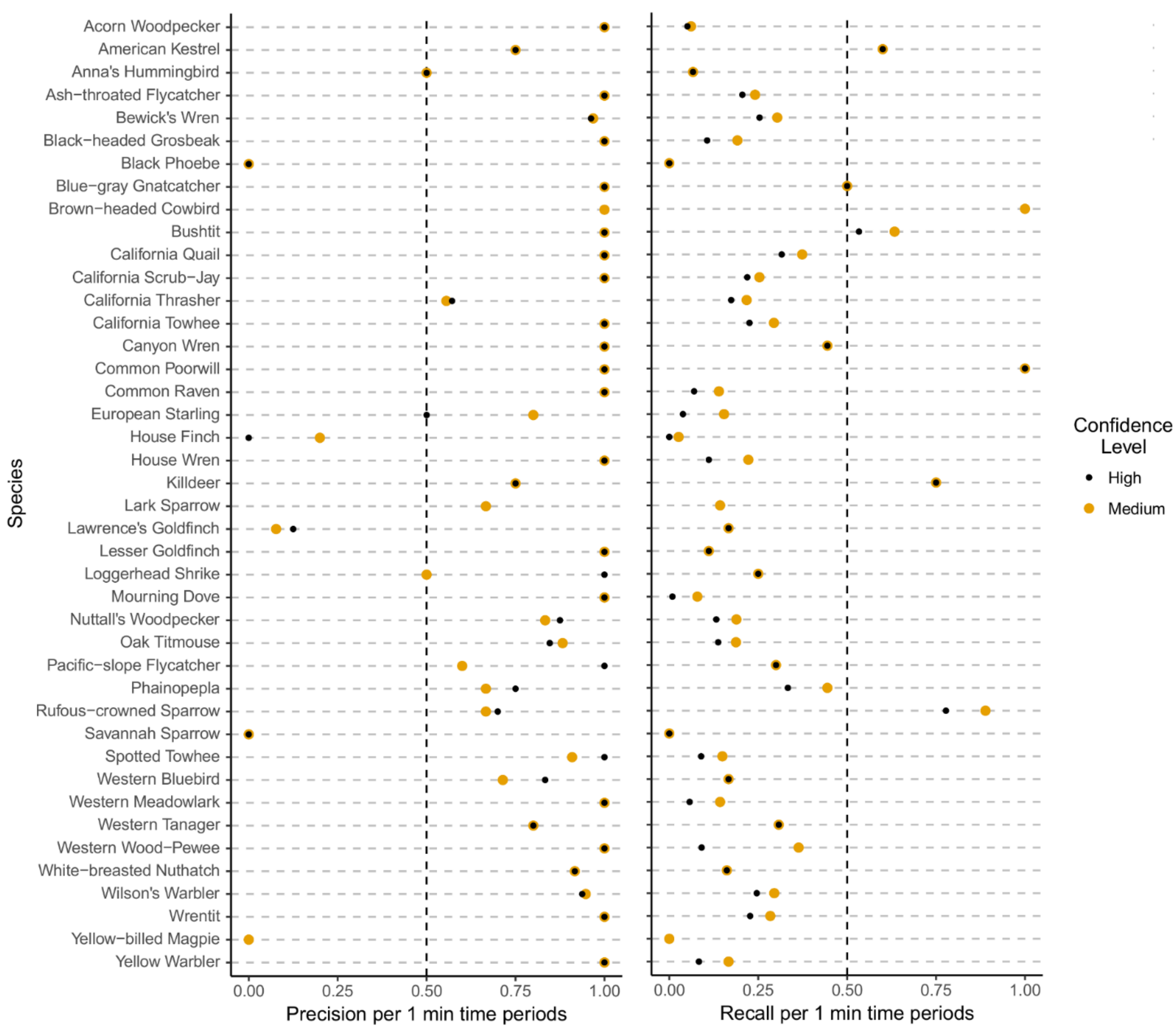

Figure 8. Precision (how often a species was identified corrected when detected) and recall (how often a species was detected when actually present) statistics for bird species detected during automatic annotation using BirdNET software. Only species detected during human annotation and automatic annotation of recordings are presented. Statistics are reported for detections screened by BirdNET's confidence in the species' identities and are denoted as medium (detections with confidence >0.65) and high (detections with confidence >0.85). Formula for precision and recall are provided in Methods.

The BirdNET automatic bird sound annotation software was effective at correctly labeling bird species when they were detected on the recording. However the majority of species were not detected by BirdNET as frequently as they were during human annotation of recordings. Given that false negatives were a frequent occurrence with BirdNET it might be beneficial to use automatic annotation on a longer recording segment (e.g., 30 minutes rather than 10 minutes) to maximize the probability of detecting a species at least once. We suspect that on "acoustically busy" recordings (i.e., many different bird species vocalizing simultaneously) one bird species may be masked by others and the probability of detection by the BirdNET algorithm is lowered. The BirdNET software was very recently developed and it will likely increase its effectiveness with time.

Acoustic ARUs provide a promising method for monitoring bird populations within SVRAs. Recording large amounts of acoustic data with relatively little effort compared to in-person point count surveys is one of the main benefits of using ARUs. These data can be archived, allowing not only analysis in the year they are collected, but also reanalysis when more refined statistical or automated annotation analysis techniques are developed. The BirdNET software could significantly reduce the time and expertise required for conducting bird surveys within each park, because recordings could be processed using cloud computing systems (e.g., Amazon Web Services) that perform automatic annotation remotely with relatively little human intervention. However, even though human annotation of recordings can be time-consuming, it still appears to be the best method for detecting birds and correctly identifying them to species at this time. Species richness derived from human-annotated ARUs has often matched in-person species richness across multiple studies (Darras et al. 2018). However, we found that counting the number of distinct individuals on a recording was difficult; it may therefore make more sense to use occupancy modeling to analyze annotated recording data because data are collapsed to species presence or absence. It is possible, with significant effort, to determine the characteristics of vocalizations for a given species and calculate the rate of sound decay with distance to determine the density of a particular bird species (e.g., Sebastián-González et al. 2018) but this requires a substantial time investment for density calculation across all species present.

Additionally, false positives and negatives are likely to be present in any annotation of recorded sound (either automatic or manual) and may bias annotations. Fortunately methods have recently been developed to account for false positive errors from automatic classification algorithms when a small segment of data (as little as 1% of detections) is validated by a human (Chambert et al. 2018). Human-annotated data can also have false positives and negatives accounted for when multiple observers annotate the same recording (Rempel et al. 2019). If ARU recordings are used for monitoring bird populations within and across SVRAs we recommend that any analysis of data incorporate a method of accounting for species misidentifications. We believe that ARUs may be beneficial to the future of scalable wildlife surveys, but are not a panacea for all studies (e.g., in instances where surveyors can only access a site once, or are surveying for species that are primarily detected visually). However, few bird species are detected only visually at the SVRAs, so the caveat listed above may not be pertinent. ARUs can be deployed for long-periods of time, with minimal training and expertise required, and enable researchers and managers to cover large regions with relatively low cost and effort. While we did not examine the ability of ARUs to detect OHV users close to sampling points because our sampling periods were during low OHV use periods, this technology could easily be used to quantify sound disturbance at each sampling point.

We emphasize that skilled observers will likely be necessary for annotation of recordings in the foreseeable, either manual annotation of the full recording set, or to evaluate the ability of an automated algorithm to correctly identify bird species on a subset of the data. Recordings should also have wind and background noise quantified, because we have demonstrated that excessive noise reduces an annotator's (and likely the BirdNET algorithm) ability to identify bird species. It may also be beneficial to select a segment of the recording for annotation that has low or medium wind noise, an option made possible because of the long sampling duration of ARUs. It was exciting to see that an automated algorithm was capable of effectively identifying bird species and it holds great potential for monitoring at SVRAs in the future.

REFERENCES

Buckland, S.T., E.A. Rexstad, T.A. Marques, and C.S. Oedekoven. 2015. *Distance Sampling: Methods and Applications*. Springer, New York, New York.

Chambert, T., J.H. Waddle, D.A.W. Miller, S.C. Walls, and J.D. Nichols. 2018. A new framework for analysing automated acoustic species detection data: Occupancy estimation and optimization of recordings post-processing. *Methods in Ecology and Evolution* 9:560–570.

Darras, K., P. Batáry, B. Furnas, A. Celis-Murillo, S.L. Van Wilgenburg, Y.A. Mulyani, and T. Tscharntke. 2018. Comparing the sampling performance of sound recorders versus point counts in bird surveys: A meta-analysis. *Journal of Applied Ecology* 55:2575–2586.

Lankau, H.E., A. MacPhail, M. Knaggs, and E. Bayne. 2015. *Acoustic Recording Analysis Protocol*. Edmonton, Alberta, Canada.

R Core Team. 2017. *R: A language and environment for statistical computing*. R Foundation for Statisical Computing, Vienna, Austria.

Rempel, R.S., J.M. Jackson, S.L. Van Wilgenburg, and J.A. Rodgers. 2019. A multiple detection state occupancy model using autonomous recordings facilitates correction of false positive and false negative observation errors. *Avian Conservation and Ecology* 14:1.

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

APPENDIX

Table A1. Summary of which bird species were detected during in-person point count surveys or heard by a human annotator on automated recording unit (ARU) surveys of the same 44 survey points within Carnegie SVRA. *Continued on the next page.*

Common Name	Scientific Name	Detected during in - person point count	Detected by human annotator during ARU recording
Acorn Woodpecker	*Melanerpes formicivorus*	X	X
American Crow	*Corvus brachyrhynchos*	X	X
American Kestrel	*Falco sparverius*	X	X
Anna's Hummingbird	*Calypte anna*	X	X
Ash-throated Flycatcher	*Myiarchus cinerascens*	X	X
Bewick's Wren	*Thryomanes bewickii*	X	X
Black-headed Grosbeak	*Pheucticus melanocephalus*	X	X
Black-throated Gray Warbler	*Setophaga nigrescens*	X	
Black Phoebe	*Sayornis nigricans*	X	X
Blue-gray Gnatcatcher	*Polioptila caerulea*	X	X
Brewer's Blackbird	*Euphagus cyanocephalus*	X	
Brown-headed Cowbird	*Molothrus ater*		X
Bullock's Oriole	*Icterus bullockii*	X	X
Bushtit	*Psaltriparus minimus*	X	X
California Quail	*Callipepla californica*	X	X
California Scrub-Jay	*Aphelocoma californica*	X	X

California Thrasher	*Toxostoma redivivum*	X	X
California Towhee	*Melozone crissalis*	X	X
Canyon Wren	*Catherpes mexicanus*	X	X
Cassin's Vireo	*Vireo cassinii*	X	
Cliff Swallow	*Petrochelidon pyrrhonota*	X	
Common Poorwill	*Phalaenoptilus nuttallii*	X	X
Common Raven	*Corvus corax*	X	X
Dark-eyed Junco	*Junco hyemalis*	X	X
European Starling	*Sturnus vulgaris*	X	X
Hermit Warbler	*Setophaga occidentalis*	X	
Horned Lark	*Eremophila alpestris*	X	
House Finch	*Haemorhous mexicanus*	X	X
House Wren	*Troglodytes aedon*	X	X
Hutton's Vireo	*Vireo huttoni*	X	
Killdeer	*Charadrius vociferus*	X	X
Lark Sparrow	*Chondestes grammacus*	X	X
Lawrence's Goldfinch	*Spinus lawrencei*	X	X
Lesser Goldfinch	*Spinus psaltria*	X	X
Loggerhead Shrike	*Lanius ludovicianus*	X	X
Mourning Dove	*Zenaida macroura*	X	X
Northern Flicker	*Colaptes auratus*	X	X
Northern Mockingbird	*Mimus polyglottos*	X	
Nuttall's Woodpecker	*Picoides nuttallii*	X	X
Oak Titmouse	*Baeolophus inornatus*	X	X
Olive-sided Flycatcher	*Contopus cooperi*	X	X
Pacific-slope Flycatcher	*Empidonax difficilis*	X	X
Phainopepla	*Phainopepla nitens*	X	X
Purple Finch	*Haemorhous purpureus*		X
Red-tailed Hawk	*Buteo jamaicensis*	X	X
Ruby-crowned Kinglet	*Regulus calendula*	X	X
Rufous-crowned Sparrow	*Aimophila ruficeps*	X	X
Savannah Sparrow	*Passerculus sandwichensis*		X
Song Sparrow	*Melospiza melodia*	X	
Spotted Towhee	*Pipilo maculatus*	X	X
Steller's Jay	*Cyanocitta stelleri*	X	
Swainson's Thrush	*Catharus ustulatus*	X	X
Townsend's Warbler	*Setophaga townsendi*		X
Warbling Vireo	*Vireo gilvus*	X	X
Western Bluebird	*Sialia mexicana*	X	X
Western Kingbird	*Tyrannus verticalis*	X	X
Western Meadowlark	*Sturnella neglecta*	X	X
Western Tanager	*Piranga ludoviciana*	X	X
Western Wood-Pewee	*Contopus sordidulus*		X
White-breasted Nuthatch	*Sitta carolinensis*	X	X

White-tailed Kite	*Elanus leucurus*	X	
Wild Turkey	*Meleagris gallopavo*	X	X
Wilson's Warbler	*Cardellina pusilla*	X	X
Wrentit	*Chamaea fasciata*	X	X
Yellow-billed Magpie	*Pica nuttalli*	X	X
Yellow Warbler	*Setophaga petechia*	X	X

Table A2. Results of a Wilcoxon-signed ranked test comparing point-level species richness derived from passive point count surveys and ARU annotated recordings in low, medium, and high wind noise conditions. "V" is the sum of ranks assigned to differences with a positive sign.

Wind Class	V	p-value
Low	4	<0.001
Medium	21.5	0.097
High	21	0.034

Table A3. Wilcoxon ranked-sign test results comparing point-level species richness derived from in-person point counts to species richness from human-annotated ARU recordings. "V" is the sum of ranks assigned to differences with a positive sign.

Common Name	V	p-value
Acorn Woodpecker	64	0.846759
Ash-throated Flycatcher	91	0.008457
Bewick's Wren	65	0.143868
Black-headed Grosbeak	30	0.350648
California Quail	39	0.003552
California Scrub-Jay	45	0.094761
California Towhee	67	0.0209
Common Raven	48	0.263395
House Finch	24	0.034596
Mourning Dove	27	0.000656
Northern Flicker	20	0.789726
Nuttall's Woodpecker	12	0.776814
Oak Titmouse	63	0.877831
Spotted Towhee	12	0.776814
Western Meadowlark	16.5	0.119952
White-breasted Nuthatch	10	0.071861
Wrentit	20	0.821098

Table A4. List of birds ever detected on ARU recordings at sites with low and medium wind noise when recordings were annotated either by a skilled listener (human annotation column), or species that were automatically annotated by BirdNET software, filtered to species ever detected during previous point count surveys at Carnegie, and filtered again to those species that BirdNET was 65% confident (auto-annotation at 0.65 column) or 85% confident (auto-annotation at 0.85 column) of the identity. Rows with an "X" denotes species that were detected using a given method. Continued on next page.

Common name	Human annotation	Auto-annotation at 0.65 confidence	Auto-annotation at 0.85 confidence
Acorn Woodpecker	X	X	X
American Crow	X		
American Kestrel	X	X	X
American Pipit		X	X
Anna's Hummingbird	X	X	X
Ash-throated Flycatcher	X	X	X
Band-tailed Pigeon		X	
Bell's Sparrow		X	X
Belted Kingfisher		X	
Bewick's Wren	X	X	X
Black Phoebe	X	X	X
Black-chinned Hummingbird		X	
Black-headed Grosbeak	X	X	X
Blue-gray Gnatcatcher	X	X	X
Brewer's Blackbird		X	X
Brown-headed Cowbird	X	X	
Bullock's Oriole	X		
Bushtit	X	X	X
California Quail	X	X	X
California Scrub-Jay	X	X	X
California Thrasher	X	X	X
California Towhee	X	X	X
Calliope Hummingbird		X	X
Canyon Wren	X	X	X
Cedar Waxwing		X	X
Cliff Swallow		X	X
Common Poorwill	X	X	X
Common Raven	X	X	X
Dark-eyed Junco	X		
Downy Woodpecker		X	
European Starling	X	X	X
Golden-crowned Sparrow		X	X
Grasshopper Sparrow		X	
Hairy Woodpecker		X	X

Hermit Thrush		X	
Hooded Oriole		X	X
Horned Lark		X	
House Finch	X	X	X
House Wren	X	X	X
Killdeer	X	X	X
Lark Sparrow	X	X	
Lawrence's Goldfinch	X	X	X
Lazuli Bunting		X	X
Lesser Goldfinch	X	X	X
Lewis's Woodpecker		X	X
Loggerhead Shrike	X	X	X
Mourning Dove	X	X	X
Northern Flicker	X		
Nuttall's Woodpecker	X	X	X
Oak Titmouse	X	X	X
Olive-sided Flycatcher		X	
Orange-crowned Warbler		X	
Pacific-slope Flycatcher	X	X	X
Phainopepla	X	X	X
Purple Finch	X		
Rock Pigeon		X	
Rock Wren		X	
Rufous-crowned Sparrow	X	X	X
Savannah Sparrow	X	X	X
Spotted Towhee	X	X	X
Swainson's Thrush	X		
Townsend's Warbler	X		
Tree Swallow		X	
Vaux's Swift		X	X
Warbling Vireo	X		
Western Bluebird	X	X	X
Western Kingbird	X		
Western Meadowlark	X	X	X
Western Tanager	X	X	X
Western Wood-Pewee	X	X	X
White-breasted Nuthatch	X	X	X
Wilson's Warbler	X	X	X
Wrentit	X	X	X
Yellow Warbler	X	X	X
Yellow-billed Magpie	X	X	