Journal of Field Ornithology

J. Field Ornithol. 80(4):387–398, 2009 DOI: 10.1111/j.1557-9263.2009.00245.x

Humans versus autonomous recording units: a comparison of point-count results

Richard L. Hutto[1] and Ryan J. Stutzman

Avian Science Center, Division of Biological Sciences, University of Montana, Missoula, Montana 59812, USA

Received 29 April 2009; accepted 17 August 2009

ABSTRACT. There is growing interest in the potential use of autonomous recording units (ARUs) to obtain point-count bird survey data in the absence of human observers. To determine possible reasons for differences in the point-count performances of humans and ARUs and to better understand the possible limitations of ARUs, we compared the results of point-counts conducted by human observers with those conducted using ARUs. Human observers and ARUs recorded birds at 56 point-count stations in three different habitats (green mixed-conifer forest, burned mixed-conifer forest, and mixed riparian cottonwood bottomland). Combined, human observers and ARUs generated a total of 858 detections of 86 different species. We found that 9.7% of detections were recorded by ARUs only, 40.9% by human observers only, and 49.4% by both ARUs and humans. The mean number of species detections per point was significantly greater for human observers (13.8) than for ARUs (9.1), and did not differ significantly among habitat types. Birds not detected using ARU recordings included those with songs or calls too distant to be recorded (52.7%), those detected by human observers using only visual cues (14.8%), and those too difficult to identify from recordings (10.3%) or simply overlooked (8.8%). About two-thirds (68.7%) of birds detected using ARUs, but missed by human observers in the field, were simply overlooked in the field; most of the rest were the result of misidentification in the lab. The failure of ARUs to record a large proportion of the detections recorded by human observers, combined with problems with detecting or identifying sounds in the lab and the extra time and cost associated with use of ARUs, suggest that they would not provide a cost-effective means of gathering data for traditional point-count surveys.

RESUMEN. **Unidades de grabación humanas vs. autónomas: una comparación de los resultados de conteos por punto**

Existe un crecimiento en el interés del uso de Unidades de Grabación Autónomas (UGA's) para obtener datos de conteo por punto en la ausencia de observadores humanos. Para determinar las posibles razones de las diferencias en el rendimiento de conteos por punto de humanos y de UGA's y para proveer información sobre las posibles limitaciones de UGA's, comparamos los resultados de conteos por punto realizados por observadores humanos con los resultados de conteos por punto usando UGA's. Los observadores humanos y las UGA's registraron aves en 56 puntos de conteo ubicados en tres hábitats diferentes (bosque verde de coníferos mixto, bosque quemado de coníferos mixto, y vegetación ribereño mixto). Los observadores humanos y las UGA's en combinación generaron un total de 858 detecciones de 86 especies diferentes. Encontramos que 9.7% de las detecciones fueron grabados solo por UGA's, 40.9% solo por los observadores y 49.4% por las UGA's y los humanos. El número promedio de especies detectados por punto fue significativamente mayor para los observadores humanos (13.8) que para UGA's (9.1), sin una diferencia significativa entre los tipos de hábitat. Las aves que no fueron detectadas usando las grabaciones de UGA's incluyeron las con cantos demasiados distantes para ser grabados (52.7%), las que fueron detectadas por observadores humanos mediante la observación visual (14.8%), las que fueron muy difíciles para identificar usando las grabaciones (10.3%), o las cuales simplemente no fueron detectadas (8.8%). Aproximadamente dos tercios (68.7%) de las aves detectadas usando UGA's, pero que no fueron detectadas por los observadores humanos en el campo, simplemente no fueron registrados por los humanos en el campo; la mayoría del resto de las aves que no fueron registradas por los humanos fueron resultado de una mala identificación en el laboratorio. La inhabilidad de las UGA's de grabar una gran proporción de las detecciones registradas por los observadores humanos, en combinación con los problemas de detectar o identificar los sonidos en el laboratorio y el tiempo extra y costo asociados con el uso de UGA's sugiere que no proveerían una manera económicamente efectiva de colectar datos para los muestreos tradicionales de conteo por punto.

Key words: acoustic survey, autonomous recording unit, monitoring, point count, sound recording

Birds can be useful for monitoring the effects of land-use practices (Hutto and Young 2002) and, for such monitoring, survey data are generally collected by field technicians who record all birds detected by sight and sound during 10-min point counts conducted according to a standard protocol (Ralph et al. 1995). Unfortunately, the use of human observers poses a number of

---

[1] Corresponding author. Email: hutto@mso.umt.edu

concerns. One is the cost of hiring and training field technicians every year. Another is the trade-off between the extensiveness and intensiveness of sampling; survey points are often visited multiple times during the season to provide an improved estimate of the probability of detecting species actually present, but this comes at a cost in terms of reduced sample sizes (Hutto et al. 1986, Ralph et al. 1995). Additional concerns are related to sampling biases associated with the use of human observers. For example, observers may vary in their abilities to detect and identify birds (Sauer et al. 1994, Cunningham et al. 1999, Diefenbach et al. 2003, Alldredge et al. 2007, Simons et al. 2007), some species may be less inclined to sing or call in the presence of a human (Hutto and Mosconi 1981, McShea and Rappole 1997, Bye et al. 2001), birds may be attracted or repulsed during the approach of a technician, producing biased detection distance profiles (Hutto and Mosconi 1981, Hutto and Young 2003), and similar sounding species may be misidentified (Sauer et al. 1994).

An alternative that may circumvent or alleviate some of these concerns is the use of autonomous recording units (ARUs) to detect birds through the use of recordings that can then be analyzed in the laboratory (Brandes 2008). ARUs have received considerable attention recently because they have been deployed in an effort to verify reports of Ivory-billed Woodpeckers (*Campephilus principalis*) in Arkansas (Fitzpatrick et al. 2005). ARUs or hand-held recorders have also been used by other investigators, and species lists generated using ARUs compare favorably to those generated by humans (Haselmayer and Quinn 2000, Hobson et al. 2002, Rempel et al. 2005).

The consistency and accuracy of recorded information makes the prospect of using ARUs for monitoring an attractive alternative to the use of human observers. Although others have evaluated ARU performance based on comparisons of simple species lists or diversity indices generated from sound recordings and human observers, only two published studies have involved a direct comparison of formal point count and ARU results (Hobson et al. 2002, Celis-Murillo et al. 2009) and both found that ARUs performed favorably. Nevertheless, a detailed analysis of the reasons for differences in the point-count performances of humans and ARUs is needed to better understand the possible limitations of ARUs.

Our objective was to test the efficacy of ARUs in detecting birds at points where observers simultaneously conducted traditional point counts in the field. Specifically, we: (1) compared species lists simultaneously generated by human observers and ARUs in the field, (2) determined the probable reason for any disagreement between the numbers of detections for each species on the two lists, and (3) provide an assessment of the performance of ARUs relative to humans.

METHODS

Description of ARUs. The ARUs used in our study were programmable, battery-operated, digital audio recorders developed by the Bioacoustics Research Program at the Cornell Lab of Ornithology (Ithaca, NY). Each ARU consisted of a water-resistant, cylindrical, polyvinyl chloride housing containing a microprocessor, 12-bit analog-to-digital converter, a clock, and a hard disk for storing audio data. The ARU was connected by a cable to an external battery pack. Each stereo ARU was equipped with two custom-made, 16-element compound microphones. Each element was an electret condenser microphone cartridge (WM61; Panasonic, Secaucus, NJ) with a sensitivity of −35 db (± 4 db) re 1V/Pa, or 17.8 mV/Pa. The elements were arranged in a linear configuration 10.3-cm long. The outputs of the individual microphone elements were summed. Although, in theory, this design could yield a combined sensitivity of 16 times that of a single element, laboratory estimates suggest, in practice, an actual sensitivity of approximately four times that of a single element. Each microphone was mounted in a vertical orientation inside an acoustically transparent, cylindrical windscreen made of artificial "fur." One microphone was mounted to the bottom of the cylindrical main ARU housing, and the second to a separate, smaller base connected to the main housing by a cable. A custom-made, configurable filter/amplifier provided high-pass filtering at 200 Hz (12 dB/octave rolloff), antialias (low-pass) filtering at 8000 Hz (30 dB/octave rolloff), and adjustable gain. The high-pass filter reduced low-frequency noise from wind and other sources. The amplifier gain was set to 45.6 dB. This setting has provided satisfactory sensitivity in a variety of habitats and environmental conditions while avoiding signal clipping from loud or nearby sounds. We made no specific effort, however, to optimize the gain setting for our study. The vertically mounted compound microphone was omnidirectional in azimuth. Audio data were stored as a series of AIF sound files with a sampling rate of 44.1 kHz and sample size of 16 bits.

Collecting ARU data. We deployed 10 ARUs at 63 points in forests near Missoula, Montana, from 30 May–4 July 2007. All points were near roads and were in one of three habitat types: burned mixed-conifer forest, green mixed-conifer forest, or mixed cottonwood bottomland and marshland (Fig. 1). These habitat types represented a continuum in vegetation structure from open with little vegetation obstruction, through moderately dense mid- and upper canopy, to dense understory and canopy. Deployments occurred in 3-d rotations, with initial deployment during the afternoon of the first day. Setting up a single unit generally took about 30 min, so we scheduled five units for deployment (2.5 h) on each set-up day. The main ARU housing and second microphone were mounted about a 1 m above ground on separate metal fence posts spaced about 2 m apart (Fig. 1). ARUs were powered up and the following information was then read aloud and recorded by the ARU: date, time of day, location, habitat type, GPS location, ARU number, and name of observer. The unit switched into a standby (nonrecording) mode after 5 min and then began recording at 05:00 for a continuous 5-h period on each of the following two mornings. All ARUs were retrieved after 10:00 on the second recording day. We rotated deployments between two groups of five ARUs; only once were all 10 ARUs recording on the same day in the same habitat.

Point count protocol. We used two human observers, one (RLH) experienced (having conducted such counts for decades) and the other (RJS) relatively inexperienced (trained to conduct point counts within the year). Our goal was to obtain data likely to bracket the extremes in results that might be obtained due to differences in observer experience. The two observers recorded bird detections simultaneously, but independently, at 10 of the 63 points during a trial period used to refine the field protocol; otherwise, a single observer recorded the data used for comparison with the data from the ARU recording. Between 06:00 and 10:00 on one of the 2 d when the ARU was recording, an observer conducted a 10-min count using a basic point-count protocol (Ralph et al. 1995). After arrival at the point-count location, the observer checked that the ARU was operating correctly and then announced the date, time of day, and time the point count started. The observer recorded all individuals of all bird species heard and seen during the 10-min recording period, noting the 2.5-min time interval when the bird was detected. Distance from the point and the directional bearing toward each bird detected were also recorded, as were the auditory (none, song, call, or drum) and visual cues (present or absent) used to detect the bird.

Lab analysis protocol. On the day that each ARU was retrieved from the field, we downloaded data from the previous two days. Each ARU could store recordings for the entire season on the built-in hard drive (about 14 5-h recording sessions), but more frequent downloads minimized chances that data would be lost if a unit malfunctioned or was stolen, and allowed us to check and, if necessary, reset the internal clock. Each data download session took approximately 30 min for a single unit.

After ARU data were downloaded, we found the 10-min recording windows that corresponded with the point counts conducted simultaneously in the field. The same observer who conducted the point count in the field then extracted data from the 10-min recording by listening to the recording through three-piece Harman/Kardon speakers while looking at the sound spectrogram (1 min duration and five lines per page) generated by RAVEN Pro (Cornell Laboratory of Ornithology, Ithaca, NY). The listener also noted the time of first detection of any bird species that was identifiable. At the end of the 10-min recording, observers then went back to reevaluate the more difficult sounds to get the most accurate species list possible.

Data analysis. Because it is difficult to distinguish and count the number of individuals of a single species at a point using sound recordings alone, we used the detection of a species at a given point as the basis of comparison. Bird-detection data collected by the human observer were aggregated to include only the single closest detection of each species at a point, and the ARU data from the same point were then merged into

Fig. 1. Photographs of the three habitats where point-count surveys were conducted—burned mixed-conifer forest (top), green mixed-conifer forest (middle), and riparian cottonwood bottomland and marshland (bottom).

the same file so that each species detected by either the ARU or the human (or both) was listed only once for a survey point. Each bird species then fell into one of three categories: (1) detected by the human only, (2) detected by the ARU only, or (3) detected by both the human and the ARU. Because observers differ in their ability to identify birds by sound, the same observer who recorded birds in the field also recorded detections in the lab for the same point. This eliminated the possibility that a difference between the ARU-based data and the human-based data was due to the observer's ability to identify a sound.

We assumed that a species detected by both the human observer and ARU was indeed present during the point count. Species detected by either the human or the ARU alone were "disagreements" that were investigated further by both of us (RLH and RJS) listening to the recording at the time the bird was detected (in the case of an ARU-only detection) or by examining the method of detection and estimated distance to the bird (in the case of a human-only detection). If a species was recorded by the ARU only, we assigned one of four possible reasons for failure of the human observer to record the species in the field: (1) the bird was detected, but misidentified (this was assigned when a second observer was present in the field and that observer detected the same species recorded by the ARU), (2) the bird was misidentified in the lab, (3) the sound signature in the lab was unambiguous so the bird was present, but undetected, or (4) reason unknown. If a species was detected and recorded by the human observer only, we assigned one of seven possible reasons for why the species was not also recorded in the lab: (1) the bird was misidentified in the field (this was assigned when a second observer was present in the field and that observer did not detect the species, or when there was a similar-sounding species recorded by the ARU alone during the same 2.5-min time period), (2) the song was recorded by the ARU, but the sound was misidentified in the lab, (3) the bird was recorded by the ARU, but overlooked in the lab, (4) the bird was recorded, but was too difficult to identify by sound alone in the lab, (5) the bird was too distant for the ARU to pick up (this was assigned when the distance to the bird exceeded the maximum distance at which the species was detected by both the human and ARU), (6) the bird was detected by visual cues only (this was assigned when the field detection cue was listed as visual only), or (7) reason unknown.

We conducted χ^2 analyses to test for significant differences in results between observers, and among bird species and the three habitat types. We used ANOVA to determine if the mean number of species detections differed with mode of detection and observer.

RESULTS

Two observers conducted point counts simultaneously (while recording bird detections independently) at 10 points in burned forest, and single observers conducted counts at the remaining 46 points (28 by RLH and 18 by RJS). We were unable to retrieve data from seven of the 63 deployments either because wires on the recording units were chewed by elk (*Cervus canadensis*; four units) or the recorder malfunctioned (three units). Thus, we compared ARU- and human-based results from a total of 56 points (31 in burned forest, 15 in green forest, and 10 in cottonwood bottomland).

Comparison of human-based and ARU-based results. The two human observers and the ARUs generated a total of 858 bird detections representing 86 species. Overall, 9.7% of the detections were recorded by ARUs only, 40.9% by human observers only, and 49.4% by both ARUs and human observers. These percentages differed for the two human observers ($\chi^2_2 = 19.5$, $P < 0.001$), with the more experienced observer (RLH) detecting a greater percentage of species recorded by both ARU and human, relatively more species than by ARUs only, and relatively fewer by human observers only (Table 1). Restricting species detections to those judged to be within 100 m of the observer (a commonly used cutoff distance for fixed-radius data analyses), the total number of detections dropped from 858 to 571, but the percentage of detections recorded by both human observers and ARUs was still only 54.5%. The main difference was twice (rather than four times) as many detections by human observers alone (31%) than by ARUs alone (14.5%).

On a per-point basis, the mean number of species detected was greater for human observers than for ARUs (13.8 ± 0.10 [SE] vs. 9.1 ± 0.07, $F_{1,110} = 27.7$, $P < 0.001$). The distribution of

Table 1. The numbers of species detections recorded by human observers only, both human observers and ARUs, and ARUs only for each observer.

Method	Observer RLH	Observer RJS	Totals
Human only	185 (35.6%)	166 (49.1%)	351 (40.9%)
Human and ARU	272 (52.3%)	152 (45.0%)	424 (49.4%)
ARU only	63 (12.1%)	20 (5.9%)	83 (9.7%)
Totals	520 (100.0%)	338 (100.0%)	858 (100.0%)

detections among the three categories (human-only, ARU-only, and both) did not differ among the three vegetation types ($\chi^2_4 = 7.0$, $P = 0.14$). For any of the three vegetation types, roughly half of the species detections were recorded by both human observers and ARUs, and about three times more were detected by human observers alone than by ARUs alone (36% vs. 12%; Table 2).

Detections by human observers only. Reasons why bird species were detected by human observers alone varied, and differed between observers ($\chi^2_6 = 49.9$, $P < 0.001$). More than half (52.7%) of the human-only detections were birds too distant for ARUs to detect (Table 3), 14.8% were birds detected by visual cues, and most other human-only detections were sounds either too difficult to identify from recordings because they were too faint or obscured by other sounds (10.3%) or were simply overlooked (8.8%). Few human-only detections (6%) were due to field misidentification, and most were by the less experienced observer. The percentages associated with each explanation did not differ (using data from RLH only) among the three vegetation types ($\chi^2_{10} = 9.1$, $P = 0.52$). Restricting species detections to those judged to be within 100 m of the observer, the numbers of human-only detections became more evenly distributed among the possible reasons, but the two most frequent reasons were still that cues were either visual only or too distant (Table 3).

Detections by ARUs only. On average, 68.7% of species detections recorded by ARUs alone were simply overlooked and not recorded by human observers in the field, and 26.5% were due to song misidentification in the lab (Table 4). Again, the percentages associated with each explanation did not differ (using data from RLH only) among the three vegetation types ($\chi^2_4 = 2.1$, $P = 0.71$).

Combining data from the two observers, we found marked differences in the relative proportions of ARU-only, human-only, and joint detections among bird species. Our data are too sparse (nearly 80% of the contingency table cells have expected values less than 5) to calculate a reliable χ^2 statistic, but these differences are undoubtedly nonrandom. For example, compared to most other species, the probability of human-only detection was greater for species like Calliope Hummingbirds (*Stellula calliope*) that were detected exclusively using visual cues, and like Red-breasted Nuthatches (*Sitta canadensis*), American Robins (*Turdus migratorius*), and Chipping Sparrows (*Spizella passerina*) that were often detected by human observers in the field at distances too great for ARUs to pick up clearly, if at all.

DISCUSSION

The two human observers in our study recorded 83 (95.4%) of the 87 species detected

Table 2. The numbers of species detections that were recorded by the human observer only, both human observers and ARU, and ARU only, for each of three habitat types.

Method	Habitat: Burned forest	Habitat: Green forest	Habitat: Riparian	Totals
Human only	63 (44.4%)	67 (30.9%)	55 (34.2%)	185 (35.6%)
Human and ARUs	64 (45.1%)	122 (56.2%)	86 (53.4%)	272 (52.3%)
ARUs only	15 (10.6%)	28 (12.9%)	20 (12.4%)	63 (12.1%)
Totals	142 (100.0%)	217 (100.0%)	161 (100.0%)	520 (100.0%)

Table 3. Reasons why birds were detected during point counts by human observers and not by the ARUs. Most human-only detections were birds too distant for ARUs to record or were identified based only on visual cues, regardless of whether we used all bird detections or only those determined to be within 100 m of observers.

	All data			Detections within 100 m		
	Observer			Observer		
Reason	RLH	RJS	Totals	RLH	RJS	Totals
Unknown	19 (10.3%)	6 (3.6%)	25 (7.1%)	17 (16.0%)	6 (8.5%)	23 (13.0%)
Misidentified in field	3 (1.6%)	18 (10.8%)	21 (6.0%)	1 (0.9%)	18 (25.4%)	19 (10.7%)
Misidentified in lab	0 (0.0%)	1 (0.6%)	1 (0.3%)	0 (0.0%)	1 (1.4%)	1 (0.6%)
Missed in lab (overlooked)	14 (7.6%)	17 (10.2%)	31 (8.8%)	13 (12.3%)	13 (18.3%)	26 (14.7%)
Missed in lab (too hard)	27 (14.6%)	9 (5.4%)	36 (10.3%)	23 (21.7%)	8 (11.3%)	31 (17.5%)
Missed in lab (too distant)	80 (43.2%)	105 (63.3%)	185 (52.7%)	19 (17.9%)	15 (21.1%)	34 (19.2%)
Missed in lab (visual only)	42 (22.7%)	10 (6.0%)	52 (14.8%)	33 (31.1%)	10 (14.1%)	43 (24.3%)
Totals	185 (100.0%)	166 (100.0%)	351 (100.0%)	106 (100.0%)	71 (100.0%)	177 (100.0%)

by either humans or ARUs, whereas the ARUs recorded 69 (79.3%) of total number of species detected. Thus, using equipment comparable to that used in our study, human observers generated a more complete species list than ARUs in the same amount of sampling time. Across all counts, this translates into 19 species recorded by human observers that were undetected by ARUs, and four species recorded by ARUs that were undetected by human observers. All ARU-only and many human-only detections were single-instance detections that represented little more than chance differences between methods, but relatively silent waterfowl, hummingbird, and flyover species would likely be missed routinely by ARUs. However, because landbird surveys rarely include waterfowl and flyover data, none of the species routinely reported in landbird surveys by human observers went entirely unrecorded by the ARUs (e.g., we heard ARU recordings of hummingbird chip-threat calls and diving flights outside the 10-min point count period). Our results are, therefore, similar to those reported in Peru (Haselmayer and Quinn 2000), where human observers were deemed slightly better because of their ability to detect rare species, but where recordings generally worked better in locations where species richness was high.

For most management purposes, however, cumulative species lists are not as valuable as presence/absence information. The presence/absence information is not only the most consistent and reliable point-based metric, but is the basis for calculating probabilities of occurrence across larger areas. In our study, two human observers recorded 775 (90.3%) of the 858 bird detections recorded by either humans or ARUs, whereas the ARUs recorded 507 (59.1%) of the bird detections recorded by either source. Even if we restrict the analysis to detections judged to be within 100 m of the observer ($N = 571$), human observers still recorded 488 (85.5%) of the birds detected by either source, whereas the ARUs recorded 337 (59.0%) of the birds detected by either source.

The 83 instances where the ARUs detected species missed by human observers included 44 (51%) of the 87 species detected. Thus, misses or misidentifications by human observers appeared to be primarily context dependent (i.e., related to ambient noise, distractions, or multiple species singing simultaneously) and

Table 4. Possible reasons why birds were detected using ARUs, but not by human observers. Most ARU-only detections were missed by human observers in the field.

Most likely reason	Observer		Totals
	RLH	RJS	
Unknown	0 (0.0%)	2 (10.0%)	2 (2.4%)
Misidentified in field	1 (1.6%)	1 (5.0%)	2 (2.4%)
Misidentified in lab	18 (28.6%)	4 (20.0%)	22 (26.5%)
Missed in field	44 (69.8%)	13 (65%)	57 (68.7%)
Totals	63 (100.0%)	20 (100.0%)	83 (100.0%)

were not dependent on the species involved. Indeed, other authors have either suggested (Bystrak 1981) or shown (Simons et al. 2007, Breeden et al. 2008) that ambient noise and background singing by multiple species can cause difficulties for observers. Although the 83 ARU-only detections amounted to only 9.7% of all detections in our study, we were still surprised by the number of times (57) we missed a species that was clearly present in the field and that the percentage did not differ among habitat types. We suspect that, regardless of habitat type, observers are generally not very good at multitasking when several species are singing and calling at once, especially at the start of a count when an observer may be more likely to forget to note a species that was actually detected. Indeed, most (47%) ARU-only detections missed in the field were recorded during the first 2.5 min of the 10-min count, significantly more than the 25% expected if misses were independent of the hectic first few minutes of a count. Celis-Murillo et al. (2009) also suggested that the first several minutes of a bird survey can be challenging for observers because of the large number of new detections. Celis-Murillo et al. (2009) also noted that an advantage of recordings is that sounds can be replayed to allow detection of individuals that might otherwise have been overlooked and this, in turn, can help meet the assumptions associated with the use of a temporal-removal modeling approach to estimate detection probabilities.

Concerning human-only detections, most (273 of 351, or 78%) birds missed by ARUs in our study were due to limitations of ARU technology, including the inability of ARUs to use visual cues, to use a broader field context for identification, and to pick up more distant sounds or sounds embedded in other noise. Even restricting the analysis to detections within 100 m of observers, 34 of 177 (19.2%) detections were still deemed too distant for ARUs to record. Over half of those detections involved Red-breasted Nuthatches, Swainson's Thrushes (*Catharus ustulatus*), American Robins, and Townsend's Warblers—species that sometimes sang or called too softly for ARUs to pick up, even well within 100 m. Improvement in recorder sensitivity could eliminate many of the more distant detections recorded by humans only, but there is no technical solution for the absence of visual-only or context-dependent detections that humans routinely record (these amounted to 14.8% and 10.3% of human-only detections, respectively), and no sure way to avoid missing or misidentifying sounds in the lab (9.1% of human-only detections).

Some ARU-only detections (28.6%) and some human-only detections (16.2%) were due to human error in the detection or identification of sounds. Rempel et al. (2005) also reported considerable variation among sound interpreters, and that variation increased exponentially as bird abundance decreased. Thus, we can expect a certain amount of identification error even when recordings are available. In a few instances, the more naïve listener in our study incorrectly recorded the presence of a species that was mimicked by another. For examine, in our study area, the call of Red-tailed Hawks (*Buteo jamaicensis*) can be mimicked well by Gray Jays (*Perisoreus canadensis*), Olive-sided Flycatchers (*Contopus cooperi*) by Pine Grosbeaks (*Pinicola enucleator*), Western Tanagers (*Piranga leudoviciana*) by Cassin's Finches (*Carpodacus cassinii*), and Western Meadowlarks (*Sturnella neglecta*) by European Starlings (*Sturnus vulgaris*). Although infrequent, such mimicry presents a challenge to unsuspecting field technicians, and also to the idea that we can automate the task of bird identification through bird recordings—a task already made daunting by intraspecific variation in song, temporal song overlap among species, ambient noise levels, and other sounds that can result in false positive identifications (Kogan and Margoliash 1998, Swiston and Mennill 2009).

Comparison of time and cost. Using ARUs is more time consuming than having a human observers conduct bird surveys. This is because of the additional set-up and take-down time associated with ARUs, as well as the time needed to download data and listen to the recorded data for at least as long as the duration of the point-count in the field. Because of the weight of the batteries and the number of components, it took either two trips or two people in a single trip to deploy the ARUs in our study. We tried to minimize deployment time by setting up ARUs within 50 m of the edge of locked and gated dirt roads, but set-up and take-down time still required about 3 h for every five units—time that could have been used to visit roughly twice the number of survey points in a season. In the laboratory, we generally took about 20 min to review a 10-min recording and, in rare instances, may have taken as much as 30 min, just as Celis-Murillo et al. (2009) reported in their study. Using an automated species identification algorithm could reduce processing time and eliminate variation among sound interpreters in the lab, but, as noted above, intraspecific variation in song, interspecific similarity in sounds produced, temporal song overlap among species, and the presence of ambient noise makes the prospect of being able to do so unlikely.

The financial costs associated with using human observers and ARUs depend on the nature of the survey being conducted, but to conduct a 300-point survey in the manner that observers typically do in the Northern Region Landbird Program (Hutto and Kowalski 2006) would require a field technician and travel expenses over a typical 2-mo field season (in current US dollars, about $5000). Collecting survey data from the same number of points using ARU recordings alone would save the cost of employing a field technician to collect bird data, but someone would still have to be paid to drive in, set up, drive out, drive back in, take down, and drive back from the survey locations (180 h for deployment of units at 300 different points), and for the time to employ a skilled technician to download recordings and listen to the ARU recordings to extract data (about $10 000). In addition, there would be the cost of the 10 ARUs ($20 000 for the prototype experimental units that we used, but about $6000 for commercially available units; Wildlife Acoustics, Inc., Concord, MA). Thus, in current dollars, we estimate a cost of about $5000 to employ a person to conduct traditional point counts and about three times that amount ($16 000) to pay for equipment and to hire the technical help needed to deploy ARUs to do the same job. Continued improvements in commercially available units will certainly reduce deployment times and equipment costs in the future, but we believe that the time and cost associated with gathering data using ARUs will always be somewhat greater than that using human observers alone.

Reliability and amount of data collected. Although ARUs are not only adequate, but probably superior, for identifying species by song, many bird calls can probably be identified to species more easily in the field than in the lab because of subtle cues humans can use in combination with sound in the field (e.g., the species of plant where a calling bird is located, the height of the sound source, or visual cues associated with the sound). Indeed, some bird species typically provide only visual cues (e.g., hummingbirds) and others are more often detected using visual cues (e.g., swallows, bluebirds, American Robins, and Dark-eyed Juncos). In addition, there are flight noises, copulation calls, and other call notes that may be useful for identifying species in the field, but may be harder to identify in the lab when taken out of the field context. Finally, humans can detect birds at greater distances than the ARUs we used (Fig. 2). Thus, even if cost were not an issue, ARUs likely cannot replace trained human observers when it comes to conventional point-count monitoring.

During our study, we also lost data from ARUs at seven (11%) of 63 locations. Newer units may be more reliable, but the possibility of data loss due to equipment malfunction remains another potential disadvantage of using ARUs.

Traditional point-count data. If the objective of a study is to obtain point-count data from a series of locations, we believe using human observers will be more cost-effective than using ARUs. Nonetheless, because sound recordings in our study did reveal the presence of species sometimes overlooked by human

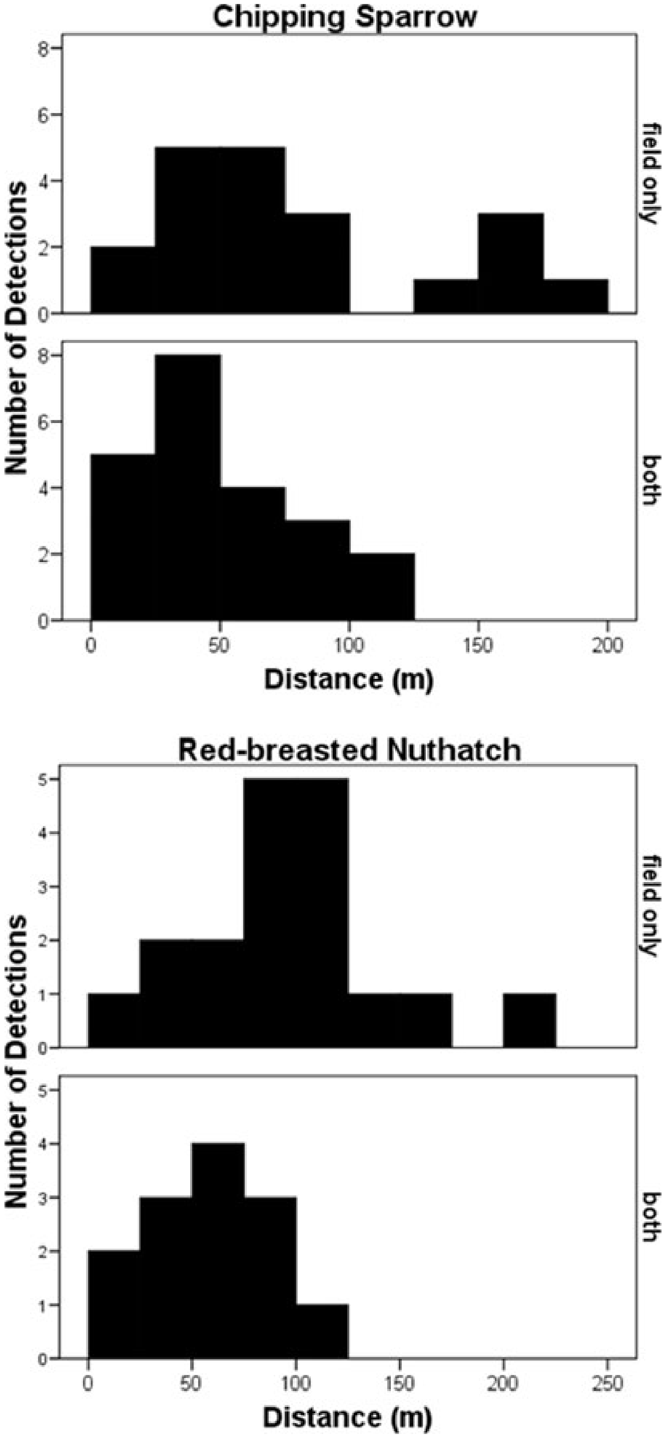

Fig. 2. Detection distances to Chipping Sparrows and Red-breasted Nuthatches when detected by human observers alone or by both human observers and ARUs. The absence of detections by the ARU when the distance to the bird was beyond about 100 m suggests that distance-based ARU recording limitations led to human-only detections when distance from the recording unit to birds was too great.

observers, and because recordings aided in the correct identification of species that technicians may have a hard time distinguishing in the field (e.g., Chipping Sparrow vs. Dark-eyed Junco, or Hammond's vs. Dusky flycatcher), we recommend the routine use of portable recorders as part of the formal point-count protocol. The use of portable recorders would not require additional time, but would provide recordings that could be reviewed later. The value of greater interobserver consistency and of having a permanent archive of the sounds associated with each point count cannot be overstated, as Acevedo and Villanueva-Rivera (2006) and Celis-Murillo et al. (2009) have noted previously.

Best use of ARUs for monitoring purposes. ARUs, no matter how technologically sophisticated, are not likely to serve as a cost-effective and satisfactory replacement for human observers for point counts. This conclusion is based largely on a comparison of cost. Indeed, ARUs would probably perform better than humans at detecting species from a point if recordings of longer duration were made at each point. Those results would not be comparable to results from 10-min point counts, however, and longer recordings would greatly increase the cost of analysis.

Although we are more skeptical than Hobson et al. (2002) that ARUs will ever serve as a cost-effective substitute for the use of human observers, there are clearly situations where the performance or cost-effectiveness of using ARUs is likely to exceed that of human observers. One is when the goal is to determine whether a particular species is present in a fairly restricted area (e.g., Ivory-billed Woodpecker). In addition, if a species rarely vocalizes, the benefit of using an ARU may outweigh the cost of deployment because it can be set up to record for an extended period. Finally, ARUs might also be ideal for nocturnal surveys because distracting ambient sounds are minimal at night and danger to human observers is probably at its maximum.

ACKNOWLEDGMENTS

We thank B. VanHorne for suggesting that we conduct this study, and R. Charif, R. Farve, and S. Kowalski for logistical support. The manuscript benefitted from comments by R. Charif, R. Farve, J. Young, and three anonymous reviewers. The USFS Northern Region and USFWS Lee Metcalf Wildlife Refuge provided access to protected areas to conduct the surveys. The project was funded by the USFS San Dimas Technology & Development Center's Inventory and Monitoring Program, and by a University of Montana Undergraduate Research Award. Equipment and software were provided by the Cornell Lab of Ornithology.

LITERATURE CITED

ACEVEDO, M. A., AND L. J. VILLANUEVA-RIVERA. 2006. Using automated digital recording systems as effective tools for the monitoring of birds and amphibians. Wildlife Society Bulletin 34: 211–214.

ALLDREDGE, M. W., T. R. SIMONS, AND K. H. POLLOCK. 2007. Factors affecting aural detections of songbirds. Ecological Applications 17: 948–955.

BRANDES, T. S. 2008. Automated sound recording and analysis techniques for bird surveys and conservation. Bird Conservation International 18: S163–S173.

BREEDEN, J. B., F. HERNANDEZ, R. L. BINGHAM, N. J. SILVY, AND G. L. WAGGERMAN. 2008. Effects of traffic noise on auditory surveys of urban White-winged Doves. Wilson Journal of Ornithology 120: 384–389.

BYE, S. L., R. J. ROBEL, AND K. E. KEMP. 2001. Effects of human presence on vocalizations of grassland birds in Kansas. Prairie Naturalist 33: 249–256.

BYSTRAK, D. 1981. The North American Breeding Bird Survey. Studies in Avian Biology 6: 34–41.

CELIS-MURILLO, A., J. L. DEPPE, AND M. F. ALLEN. 2009. Using soundscape recordings to estimate bird species abundance, richness, and composition. Journal of Field Ornithology 80: 64–78.

CUNNINGHAM, R. B., D. B. LINDENMAYER, H. A. NIX, AND B. D. LINDENMAYER. 1999. Quantifying observer heterogeneity in bird counts. Australian Journal of Ecology 24: 270–277.

DIEFENBACH, D. R., D. W. BRAUNING, AND J. A. MATTICE. 2003. Variability in grassland bird counts related to observer differences and species detection rates. Auk 120: 1168–1179.

FITZPATRICK, J. W., M. LAMMERTINK, M. D. LUNEAU, T. W. GALLAGHER, B. R. HARRISON, G. M. SPARLING, K. V. ROSENBERG, R. W. ROHRBAUGH, E. C. H. SWARTHOUT, P. H. WREGE, S. B. SWARTHOUT, M. S. DANTZKER, R. A. CHARIF, T. R. BARKSDALE, J. V. REMSEN Jr, S. D. SIMON, AND D. ZOLLNER. 2005. Ivory-billed Woodpecker (*Campephilus principalis*) persists in continental North America. Science 308: 1460–1462.

HASELMAYER, J., AND J. S. QUINN. 2000. A comparison of point counts and sound recording as bird survey methods in Amazonian Southeast Peru. Condor 102: 887–893.

HOBSON, K. A., R. S. REMPEL, G. HAMILTON, B. TURNBULL, AND S. L. V. WILGENBURG. 2002. Acoustic surveys of birds using electronic recordings: new potential from an omnidirectional microphone system. Wildlife Society Bulletin 30: 709–720.

HUTTO, R. L., AND S. KOWALSKI. 2006. Northern Region Landbird Monitoring Program: a USFS-University of Montana partnership designed to provide both short-term and long-term feedback for land managers. In: Monitoring science and technology symposium: unifying knowledge for sustainability in the Western Hemisphere (C. Aguirre-Bravo, P. J. Pellicane, D. P. Burns, and S. Draggan, eds.), pp. 936–943. Proceedings RMRS-P-42CD, USDA Forest Service, Fort Collins, CO.

Hutto, R. L., and S. L. Mosconi. 1981. Lateral detectability profiles for line transect bird censuses: some problems and an alternative. Studies in Avian Biology 6: 382–387.

Hutto, R. L., S. M. Pletschet, and P. Hendricks. 1986. A fixed-radius point count method for nonbreeding and breeding season use. Auk 103: 593–602.

Hutto, R. L., and J. S. Young. 2002. Regional landbird monitoring: perspectives from the northern Rocky Mountains. Wildlife Society Bulletin 30: 738–750.

Hutto, R. L., and J. S. Young. 2003. On the design of monitoring programs and the use of population indices: a reply to Ellingson and Lukacs. Wildlife Society Bulletin 31: 903–910.

Kogan, J. A., and D. Margoliash. 1998. Automated recognition of bird song elements from continuous recordings using dynamic time warping and hidden Markov models: a comparative study. Journal of the Acoustical Society of America 103: 2185–2196.

McShea, W. J., and J. H. Rappole. 1997. Variable song rates in three species of passerines and implications for estimating bird populations. Journal of Field Ornithology 68: 367–375.

Ralph, C. J., J. R. Sauer, and S. Droege. 1995. Monitoring bird populations by point counts. General Technical Report PSW-GTR-149, USDA Forest Service, Pacific Southwest Research Station, Berkeley, CA.

Rempel, R. S., K. A. Hobson, G. Holborn, S. L. Wilgenburg, and J. Elliott. 2005. Bioacoustic monitoring of forest songbirds: interpreter variability and effects of configuration and digital processing methods in the laboratory. Journal of Field Ornithology 76: 1–11.

Sauer, J. R., B. G. Peterjohn, and W. A. Link. 1994. Observer differences in the North American Breeding Bird Survey. Auk 111: 50–63.

Simons, T. R., M. W. Alldredge, K. H. Pollock, and J. M. Wettroth. 2007. Experimental analysis of the auditory detection process on avian point counts. Auk 124: 986–999.

Swiston, K. A., and D. J. Mennill. 2009. Comparison of manual and automated methods for identifying target sounds in audio recordings of Pileated, Pale-billed, and putative Ivory-billed woodpeckers. Journal of Field Ornithology 80: 42–50.