Research Paper

Density estimation of sound-producing terrestrial animals using single automatic acoustic recorders and distance sampling

Esther Sebastián-González [1,2], *Richard J. Camp* [3,4,5], *Ann M. Tanimoto* [1], *Priscilla M. de Oliveira* [1], *Bruna B. Lima* [1], *Tiago A. Marques* [5,6] and *Patrick J. Hart* [1]

[1]Department of Biology, University of Hawaiʻi at Hilo, Hilo, Hawaii, USA, [2]Applied Biology Department, Miguel Hernández University, Elche, Alicante, Spain, [3]Hawaiʻi Cooperative Studies Unit, University of Hawaiʻi at Hilo, Hawaiʻi National Park, Hawaiʻi, USA, [4]U.S. Geological Survey, Pacific Island Ecosystems Research Center, Hawaiʻi National Park, Hawaiʻi, USA, [5]Centre for Research into Ecological and Environmental Modelling, The Observatory, University of St. Andrews, St. Andrews, Scotland, [6]Centro de Estatística e Aplicações, Departamento de Estatística e Investigação Operacional, Faculdade de Ciências, Universidade de Lisboa, Lisboa, Portugal

ABSTRACT. Obtaining accurate information on the distribution, density, and abundance of animals is an important first step toward their conservation. Methodological approaches using automatic acoustic recorders for species that communicate acoustically are gaining increased interest because of their advantages over traditional sampling methods. In this study, we created and evaluated a protocol to estimate population density, which can be used to compute abundance of terrestrial sound-producing animals from single automatic acoustic recorders and using an automatic detection algorithm. The protocol uses cue rates from the target species, environmental conditions, and an estimate of the distance of the individual to the recorder based on the power of the received sound. We applied our protocol to estimate the density of a Hawaiian forest bird species (Hawaiʻi ʻAmakihi [*Chlorodrepanis virens*]) on the island of Hawaiʻi, USA. We validated our approach by comparing our density estimates with those calculated at the same stations using a traditional point-transect distance sampling method based on human observations. Overall density estimates based on recorded signals were lower than those based on human observations, but 95% confidence intervals of the two density estimates overlapped. This study presents a relatively simple but effective protocol for estimating animal density using single automatic acoustic recorders. Our protocol may easily be adapted to other sound-emitting terrestrial animals.

Densité d'animaux terrestres émettant des sons estimée au moyen d'enregistreurs audios automatiques et d'échantillonnage fondé sur la distance

RÉSUMÉ. L'obtention de données précises sur la répartition, la densité et l'abondance d'animaux est une première étape importante en vue de leur conservation. Les approches méthodologiques qui utilisent des enregistreurs audios automatiques pour les espèces communiquant par sons gagnent en intérêt en raison de leurs avantages par rapport aux méthodes d'échantillonnage traditionnelles. Dans la présente étude, nous avons créé et évalué un protocole visant à estimer la densité de population, laquelle peut ensuite être utilisée pour calculer, au seul moyen d'enregistreurs audios automatiques et d'un algorithme de détection automatique, l'abondance d'animaux terrestres émettant des sons. Le protocole repose sur le taux de signaux enregistrés d'espèces cibles, les conditions environnementales et une estimation de la distance de l'individu à l'enregistreur fondée sur la puissance du signal enregistré. Nous avons appliqué notre protocole pour estimer la densité d'une espèce d'oiseau forestière hawaïenne (Amakihi familier [*Chlorodrepanis virens*]), sur l'île d'Hawaii, aux États-Unis. Nous avons validé notre approche en comparant les estimations de densités avec celles calculées aux mêmes stations au moyen d'une méthode d'échantillonnage traditionnelle de points d'écoute et de transects fondée sur la distance (données récoltées par des personnes). Les estimations de densité basées sur les signaux enregistrés étaient globalement plus faibles que celles fondées sur les observations obtenues par des personnes, mais les intervalles de confiance à 95 % des deux types d'estimations se chevauchaient. Notre étude présente un protocole relativement simple mais efficace pour estimer la densité d'animaux au seul moyen d'enregistreurs audios automatiques. Ce protocole peut facilement être adapté pour d'autres animaux terrestres qui émettent des sons.

Key Words: *cue rate; Hawaiʻi ʻAmakihi; point count; transect; vocalization*

INTRODUCTION

Wildlife management and conservation is increasingly important as animal populations continue to decline across the globe. To develop effective management strategies for a given species, it is necessary to have accurate information on its distribution and abundance. The estimation of absolute abundance of animals is a topic that has received much attention by both researchers and managers (Burnham et al. 1980, Seber 1986, Buckland et al. 2015). Having accurate information on the population size of a species is the first step to tracking population trends and assessing the

---

Address of Correspondent: Esther Sebastián-González, Avda. Universidad S/n, Elche , Spain, 03202, esebgo@gmail.com

need and direction of management actions. Several methods have been developed to produce accurate abundance estimates and account for methodological problems, especially imperfect detection (MacKenzie and Kendall 2002, Kellner and Swihart 2014, Dénes et al. 2015). This is critical because imperfectly detecting animals can result in biased density estimates (Verner 1985, Marques et al. 2017) and lead to inappropriate conservation and management strategies.

Distance sampling is a widely used methodology for the estimation of animal population size (Buckland 2006, Camp et al. 2009). Because the probability of detecting an animal decreases with distance from an observer, functions that describe the probability of detection for each species of interest given its distance are modeled. This allows density and abundance estimation even when not every individual is detected in the sampling area (Buckland et al. 2001, 2015). The point-transect method, whereby an observer estimates the radial distance to each individual at points along a transect for a predetermined period of time, is a commonly used form of distance sampling (Buckland 2006). However, this method is sensitive to the experience and knowledge of the observer (Alldredge et al. 2007, 2008), and getting sufficient experienced observers to perform surveys can be challenging, especially for long-term monitoring programs. In addition, many animals inhabit remote areas that are difficult and expensive to access. This may severely limit the frequency of surveys and underdetect species whose activity levels do not coincide with the limited survey periods, especially species that are uncommon or rare, thereby resulting in a reduced sample size of detections to levels below what are required for reliable density estimation. An alternative to point-transects based on human detections that has become recently available is passive acoustic monitoring (e.g., Celis-Murillo et al. 2009, Dawson and Efford 2009, Marques et al. 2013, Shonfield and Bayne 2017). Many animals, including birds, cetaceans, amphibians, and insects, communicate primarily with sound and are more often detected by their sound rather than visually, especially forest birds (Camp et al. 2016); thus, acoustic cues may be used to determine animal density in sound-producing animal species.

Passive acoustic monitoring systems present several advantages over human observers (Marques et al. 2013). First, they are highly amenable to automated data collection over multiple points in time at a location, and therefore may be better able to detect uncommon species and species whose activity levels vary within and among days. Second, some factors that affect detection probability by automated systems may be less variable than those related to human observers, given the large variability in detection associated with different observers (Faanes and Bystrak 1981). Thus, automatic detectors may provide more comparable estimates among sampling units that are replicated either temporally or spatially. Also, unlike most visual surveys, passive acoustic surveys are able to operate under any light condition (e.g., both day and night, or when visibility is restricted, such as in fog) and in difficult working environments for human observers (e.g., remote locations or hazardous areas). Third, the acoustic information collected by the recorders can be quantified more objectively and data can be revisited. Lastly, automatic acoustic monitoring may be less invasive and disruptive to animals than other methods because humans are not present while data are being collected. All these characteristics make passive acoustic systems excellent candidates for long-term and large-scale monitoring of animal populations. However, there are still difficulties that have prevented their widespread use beyond simple estimates of presence–absence. For example, acoustic recording generates massive amounts of data that are very time-consuming to manually process. Even though estimating the number of temporal revisits needed for detection probability can reduce manual processing, this approach and the design of algorithms that automatically identify target species are often beyond the capability of many managers (e.g., Celis-Murillo et al. 2009, Dawson and Efford 2009).

Density estimation using passive acoustic recorders has received much attention for aquatic species (e.g., Van Paijs et al. 2009, Marques et al. 2013). Several approaches have been taken in the marine environment to estimate detection probabilities using single sensors, including using static data loggers (Kyhn et al. 2012), fixed sensors (Küsel et al. 2011), or echolocation clicks (Hildebrand et al. 2015). However, studies focusing on terrestrial species are scarce. Previous studies that have aimed to estimate terrestrial animal density using passive acoustic monitoring have used different approaches such as deploying an array of recorders (see the review by Blumstein et al. 2011), using a capture-recapture approach (Dawson and Efford 2009, Stevenson et al. 2015), pairing point count data from humans and acoustic recorders (van Wilgenburg et al. 2017), or using calibration exercises, where a model of the number of detected cues as a function of density (using areas where the density is known) is applied to estimate density in new areas (Oppel et al. 2014). However, the first two methods require several recording devices per site, and the last method requires areas with known animal density. Our main objective is to create and evaluate a protocol to estimate the density of sound-producing animals from single automatic acoustic recorders (see also Harris et al. 2013, van Wilgenburg et al. 2017) using an automatic detection algorithm. As a case study, we applied our protocol to the density estimation of a Hawaiian forest bird on the island of Hawai'i, USA. We examined the effectiveness of our approach by comparing our density estimates with those calculated using standard human-based point-transect counts conducted at the same location and time.
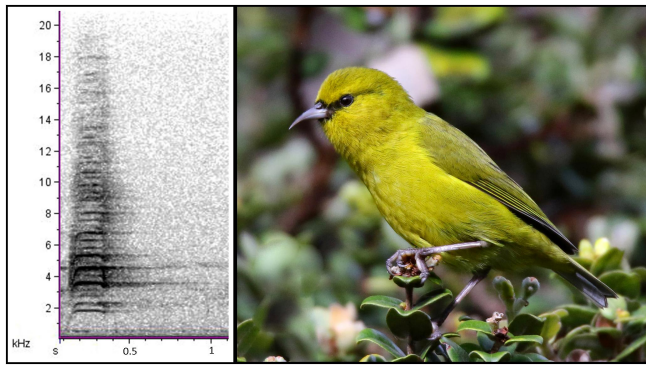
METHODS

Study area and species

Our study was performed in two native forests on the island of Hawai'i, USA. The first was located on Mauna Loa Volcano (19° 39' N, 155°21' W) in an area comprised of dozens of similarly aged forest fragments (i.e., kīpuka) that were isolated by lava flows approximately 150 years ago. The second study area was located within the Hakalau Forest National Wildlife Refuge (hereafter Hakalau forest; 19°47' N, 155°19' W). Both areas are native evergreen forests dominated by tall-statured (15–25 m) 'ōhi'a (*Metrosideros polymorpha*) and koa (*Acacia koa*) trees. Density estimates were performed only on the data from Hakalau forest, while data from Mauna Loa forest was used to calibrate the model (see *Step I*); however, the habitat is the same in both areas, and the forest structure is very similar (Sebastián-González et al. 2018).

We studied a native Hawaiian forest bird species that is widely distributed across the study areas. The Hawai'i 'Amakihi (*Chlorodrepanis virens*) is a generalist honeycreeper that consumes

nectar, fruits, and invertebrates (Lindsey et al. 1998). Like other oscine passerines, it learns song primarily through cultural transmission (Lynch 1996). Its most common call is very similar between different populations within the island (authors, *personal observations*). This call is short (duration: 0.48 ± 0.11 s), has a peak frequency ranging from 4087 to 5402 Hz (Sebastián-González et al. 2018) (Fig. 1), and is produced by both sexes year-round. ´Amakihi occur in high densities in some areas of the island, which facilitates data collection for this species and makes it a good study model for our objective.

Fig. 1. Picture and spectrogram of an ´amakihi at the study area.

Density estimation using human surveys

Field surveys
Bird density estimation by field observers was accomplished using point-transect distance sampling methods along a linear transect in the Hakalau forest (see Camp et al. 2016 for details). This transect is part of a sampling network in an area that has been surveyed annually since 1987. Surveys were conducted at eight independent stations separated by 150 m. On 27 and 31 July 2015, five experienced bird surveyors performed 8-min counts at each station between 0700 and 1100 hours. Surveyors identified the distance to each individual and the method of identification (visually or aurally). To increase the number of surveys, each station was surveyed several times the same day by different observers (between four and six times), with a minimum time difference between surveys of 30 min to ensure independence. Each station was surveyed 9–10 times over the two survey days (Table A1.1).

Statistical analysis
Density estimates using all data (auditory and visual) and the auditory-only detections from the standard point-transect counts were obtained using methods described in Camp et al. (2016). Per-station sampling effort equaled the number of times the station was surveyed. A species-specific detection function was modeled with program DISTANCE, version 7.1, release 1 (Thomas et al. 2010). The probability of detection was used to estimate bird density (birds ha^{-1}). Candidate models for the detection function were limited to half- normal and hazard-rate detection functions with expansion series of order two (Buckland et al. 2001) (half-normal was paired with cosine and Hermite polynomial adjustments, and hazard-rate was paired with cosine

and simple polynomial adjustments). Each detectability model in the candidate set was evaluated using information theoretic methods, where we selected the model with the lowest Akaike information criterion corrected for small sample sizes (AICc). Variances and confidence intervals were derived using bootstrap methods with 999 replicates. Buckland et al. (2001, 2004, 2015) and Thomas et al. (2010) describe distance-sampling procedures and analyses in detail.

Density estimation using bioacoustics
Our procedure for estimating bird density using bioacoustics had six steps (Fig. 2). To summarize, we recorded the acoustic signals and eliminated files with unfavorable weather conditions. Next, we automatically detected the calls or songs of our target species from the recordings using a published algorithm with measured performance (Sebastián-González et al. 2015). Then, we estimated the distance from the vocalizing individual to the recorder using field data on the relationship between the power of the sound (dB) and its distance to the bird (e.g., Efford et al. 2009) while taking into account weather variables. Finally, we used this information along with the cue rate (i.e., number of vocalizations per time unit) to calculate bird density using a similar approach to the point-transect distance sampling method outlined in the section *Density estimation using human surveys* (Gates and Smith 1972, Buckland 2006).

Fig. 2. Methodological steps suggested to estimate density from sound-emitting species.
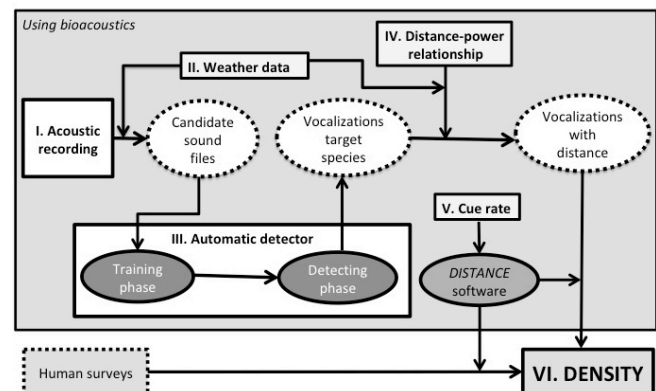
Step I: Acoustic recording
We collected acoustic data using automatic acoustic recorders (Songmeter SM2, Wildlife Acoustics Inc.) between 27 July and 17 August 2015 (see table A1.1 for acoustic sample sizes). Although the recording period was longer than the period when the human-driven surveys were done, ´amakihi are not known to change the frequency or timing of their vocalizations during this time. The recorders were stationed between 1.5 and 2 m from the ground and were located at the same points where the human-based point-count surveys were conducted. The Songmeters recorded daily from 0700 to 1100 hours in 5-min on–off duty cycles, matching the sampling period of the human-based surveys. The focal species starts vocalizing earlier in the day, but recorders

were set to intentionally avoid the dawn chorus, where there is an increased overlap in acoustic cues from our target species and other species, which increases the error in the automatic detection algorithm (see *Step III*). Recordings were made in .wav file format at a sampling rate of 44.1 kHz using a single omnidirectional microphone (SMX-II: Wildlife Acoustics) with a sensitivity of -35 dBV/pa and frequency response of 20–20,000 Hz.

The recordings at the Mauna Loa Volcano field sites were taken on different days and locations during spring 2015, and were used to train the automatic detection algorithm (see *Step III*). These recordings were made using the same automatic acoustic recorders, which were also stationed in trees between 1.5 and 2 m from the ground. From those recordings, we selected files that contained the cue type and species of interest. Further details about these training data can be found in Table A1.2.

Step II: Weather data
Both rain and wind may affect sound propagation and, in turn, density estimation using automatic acoustic recorders. Thus, we placed a weather station (ACU RITE Professional Weather Station, model 01036) proximate to the survey stations to collect automated measurements of the climatic conditions. The weather station recorded measurements of rain (mm) and wind (km/h) every 12 min. We paired the 5-min recordings to the closest measurement from the weather station. We classified total rain and average wind per 12-min interval independently into three classes as follows: for rain, Class 0: < 0.5 mm, Class 1: 0.5–1 mm, Class 2: > 1 mm; for wind, Class 0: < 15 km/h, Class 1: 15–30 km/h, Class 2: > 30 km/h. Because we needed to be able to assess the rain and wind class in the field while taking the cue rate and the power–distance measurements (see *Step IV* and *Step V*), we selected classes that were different enough that could be easily identified in the field. Because strong rain and wind completely saturate the recordings (i.e., spectrograms are totally black and all other sounds are obscured, prohibiting identification), we excluded files with rain and wind Class 2 from our analyses.

Step III: Automatic detection algorithm
We used the algorithm described in Sebastián-González et al. (2015) to identify cues (i.e., vocalizations) of the target species from the automatic recordings. This algorithm has two phases: training and detecting. The (1) training phase uses known cues (in this study, data from the Mauna Loa Volcano; see *Step I*) to train a data classification tool (e.g., a Support Vector Machine) (Cortes and Vapnik 2009) that is used in the (2) detecting phase to identify cues from the target species. The algorithm first selects candidate cues (called "selections"; N = 13,256) that are within the time length and frequency of the target species using the Band Limited Energy Detector (Mills 2000) from the Raven 1.5 software (Bioacoustics Research Program 2014). Next, candidate cues are manually sorted to identify which cues correspond to the target species. Selections from target species are then used to train the classification tool and to calculate true/false positives/negatives (see *Step III*). The accuracy of the detector was calculated using the Balanced Accuracy metric (BAC) (Féret and Asner 2012) (Eq. 1):

$$BAC = \frac{t_p + t_n}{2} \times 100 \qquad (1)$$

where t_p are true positives, and t_n are true negatives. Following the recommendations in Knight et al. (2017), we also calculated the

precision, which is the proportion of true detections of the target species (Eq. 2):

$$Precision = \frac{t_p}{t_p + f_p} \qquad (2)$$

where f_p are the false positives; recall, or the proportion of target vocalizations detected (Eq. 3):

$$Recall = \frac{t_p}{t_p + f_n} \qquad (3)$$

where f_n are the false negatives; and finally, we calculated a metric that combines both recall and precision: the F-score (Eq. 4):

$$F - score = \frac{(\beta^2 + 1) * precision * recall}{\beta^2 * precision + recall} \qquad (4)$$

where β is a parameter that is changed depending on the importance given to the precision or the recall (here, $\beta = 1.5$); t_p, t_n, f_p, and f_n were calculated using a cross-validation approach. We randomly separated 70% of the data as a training set and 30% as validation set. Training files were not used for the validation, and validation files were also never used in the training analysis. The training set was used to predict binary classes (presence or absence of the target species), and the validation set was used to evaluate the model performance. We repeated this procedure 1000 times by random selection of the training and validation of the data set, and subsequently, we calculated the mean values and SE for each parameter. These analyses were performed using R 3.4.1 (R Development Core Team 2017). See Sebastián-González et al. (2015) for further details on the detector.

Similar to the training phase, in the detecting phase we identified candidate cues in the recordings from the area where we wanted to estimate density using the Band Limited Energy Detector from Raven. Then, we used the classification tool to select only the cues from the target species. After running the detectors, we had a list of cues per 5-min recording.

Step IV: Distance–power relationship
The probability of detecting a cue (i.e., vocalization) is related to the distance of the species vocalizing to the recorder, among other factors. We therefore estimated the distance from the individual vocalizations to the recorder and used this information in the calculation of the detection probability for our density estimation. Because the power (i.e., sound energy) of an onmidirectional sound depends on the distance from the source, we estimated the distance from the recorder to the bird using the power of the cue. To measure the relationship between the power of the cues (dB) and the distance to vocalizing individuals, we collected data from the two study areas using a songmeter SM2 to record the vocalizations, and a range finder (Nikon Forestry 550) to measure the distance (m) to each vocalizing individual. We haphazardly walked through the forest and stopped when we detected an individual of the focal species. Then, we recorded the vocalizations and measured the distance to the individual bird, given that it was visually detected and its position could be determined with precision. We tried to maximize the range of distances measured by approaching or walking away from an individual when possible. We collected data until we had a fair

coverage of the possible distance and power values. Also, to reduce the chances that the same individual was recorded twice, we surveyed different areas every day, and after monitoring one individual, we walked about 50 m before monitoring a different one. The microphone of the recorder was always oriented in the same direction (up) as those placed at the sampling stations. We also noted the sampling conditions (i.e., weather and noise). These measurements were taken in 2015 (8, 9, 23, 24 June and 1, 7, 13 July). We included information on 62 vocalizations from at least 15 individuals.

After data collection, we used generalized linear models to examine the relationship among the power of the calls measured using Raven software (predictor) and the distance to the vocalizing bird measured in the field (response variable). Since sound may be affected by the climatic conditions, we also tested the effect of wind strength and rain intensity on the relationship by including them as covariates in the model. We used a Gaussian distribution with a log-link, and we selected the model with the best fit using AICc (see model diagnostics in Fig A1.2). We considered that a model was better than another if the difference in their AICc was > 2. We computed all possible combinations of univariate and multivariate models including one or both covariates. We calculated the proportion of explained deviance as a measure of fit of the model. To investigate the accuracy of this estimate, we also used the model to estimate the predicted distances for the set of observations for which we had measured the real distance to the bird. Then, we related the measured distance with the predicted one using a linear model. As a final step, we used the selected model and the power of the detected cues to calculate the distances from the recorder to the vocalizing birds.

Step V: Cue rate
We calculated the cue rate of the target species in July–August 2015 as the number of cues (i.e., calls) per minute. To do so, we searched for individuals at the two field sites by walking quietly through the forest. When individuals were visually located, we waited 5 s, then counted the number of cues until we lost visual contact with the individual. For each observation, we also noted the climatic conditions (rain and wind, similar to *Step IV*), the time of the day, and the total time of observation (in s). Because the total observation time needs to be very large to calculate an accurate cue rate, we monitored all visually located individuals, even if the same individual was potentially monitored more than one time. Then, we used nonparametric Kruskall-Wallis and Mann-Whitney tests in R to determine if the cue rates differed among climatic conditions and among times during the day. Preliminary analyses revealed that cue rates were not affected by wind and rain Classes 0 to 1, study area, or time of the day (see *Results*); therefore, we calculated the cue rate using all data. Following Marques et al. (2009), we estimated the average cue rate as a weighted average of the individual cue rates, with weights corresponding to the amount of time an individual was followed. Variance of the weighted average cue rate was calculated using Cochran's approximation (Marques et al. 2009). We analyzed only calls because 'amakihi did not produce their characteristic song trill during the mid-summer study period.

Step VI: Density estimation
We estimated bioacoustics-based bird density following the same approach as for the human-based point- transect counts, but by

using the detections and power-based distance estimates from the acoustic recordings. We estimated the density across all recording dates, where per-station sampling effort equaled the total number of 5-min intervals the station was surveyed (Table A1.1). Marques et al. (2013) provide the description and parameterization of cue-count methods and analyses, including best practices for estimating the proportion of false positive detections, and a multiplier (in this case, cue rate) to convert cue density to an estimated bird density using Eq. 5:

$$\widehat{D} = \frac{n(1 - \hat{f}_p)}{\hat{p}_v K a \hat{T} \hat{r}} \tag{5}$$

where n is the number of vocalizations, $\hat{f}_p$ is the estimated proportion of false positive detections, $\hat{p}_v$ is the estimated probability of detecting a vocalization within the area a, and $\hat{r}$ is the estimated cue rate. Note that K, T, and a are constants: a is the area covered by each point-transect ($a = \pi w^2$, where w is the truncation distance); K is the number of points the recordings were made over, and T is the time spent recording in each point (measured in the same units as the cue rate). In *Step III*, we describe how we calculated the proportion of detections that were false positives, and the number of bird cues per 5-min interval. All parameters were quantified from data that were gathered in the same location, environment, and time as acoustic sampling. The bioacoustics-based detection probability was modeled with program DISTANCE, version 7.1, release 1 (Thomas et al. 2010), following guidance by Marques et al. (2013). We used the delta method to estimate the density coefficient of variation (CV) as the square root of the combined squared CVs of the false positives, cue rate, and cue density, assuming independence of the components. The confidence interval was computed by the two-sided α-level *t*-distribution percentile where the degrees of freedom were computed using the Satterthwaite method (see Buckland et al. 2001:89). All the data used in this study are available on ScienceBase: https://doi.org/10.5066/F7PZ571Q.

RESULTS

Detection algorithm, distance-power relationship, and cue rate
We used 13,256 selections to train the automatic detection algorithm (1015 of them included an 'amakihi call). The error of the Raven Band Limited Energy Detector (i.e., the proportion of 'amakihi calls that were not selected by the detector) was 6.1% (see Table A1.2), the BAC of the automatic detection algorithm was 92.3 ± 0.9 (mean ± SD), the precision was 0.86, the recall was 0.81, and the F- score was 0.57.
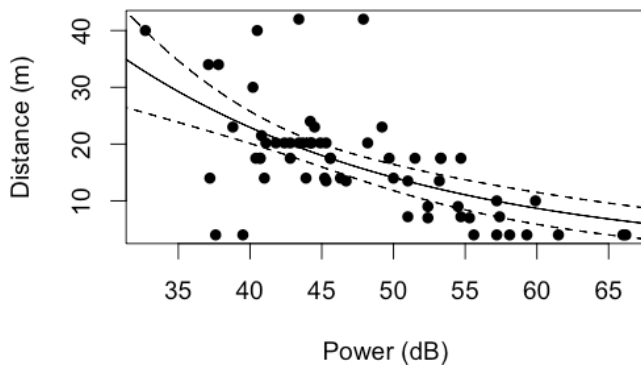
The power of the cue was significantly related to the distance of the vocalizing individual (GLM, coefficient ± SE = -0.048 ± 0.009, Intercept ± SE = 5.077 ± 0.406, t-value = -5.221, DF = 60, P < 0.001, explained deviance = 37.45%) (Fig. 3, Fig. A1.1). This relationship was not affected by the presence of rain or wind (variables not included in the model with the lowest AIC). Also, the measured distance was significantly related to the predicted one (LM, coefficient ± SE = 1.043 ± 1.04, Intercept ± SE = -0.807 ± 3.085, P < 0.001, R^2 = 0.36) (Fig. A1.2).

The cue rate was 0.63 calls/min (Table 1), calculated using all the observation time (4.74 h) and all the recorded calls (number of

Table 1. Parameters estimated for this study (mean and SE) and samples sizes of the data used to calculate them.

Description	Estimated value	SE	*N*
Proportion of false-positive detections	0.014	0.001	3976
Probability of detecting a vocalization within area "a": acoustic recorder	0.848	0.005	6830
Human visual+acoustic	0.451	0.040	257
Human acoustic	0.494	0.044	227
Cue rate (cues/minute)	0.630	0.043	206

Fig. 3. Relationship between power and distance to the vocalizing bird. We also represent the regression line from the model used to estimate the distances.

Table 2. Comparison of the Hawaiʻi ʻAmakihi mean density (individuals/ha) at all the stations calculated with human and acoustic-based surveys. We also include the lower and upper 95% confidence intervals (CI) and the standard error (SE) computed from bootstrap methods.

Method	Type	Density	SE	L 95% CI	U 95% CI
Human	Visual+ acoustic	10.50	1.52	8.25	13.89
Human	Acoustic	8.48	1.09	6.85	11.02
Acoustic	Call	6.02	1.58	3.73	8.31

calls = 179; number of recorded individuals = 206), including data from nonvocalizing individuals (*N* = 141). The cue rate was not affected by wind and rain classes, and did not change among study areas or times of the day (Kruskall-Wallis and Mann-Whitney tests; all *p* > 0.27).

Acoustic versus human density estimation

For density estimation, we used 6903 ʻamakihi detections out of 266.25 h of acoustic recordings, 261 detections from the eight human-based audio-only counts, and 289 detections from the eight human-based audio and visual counts. Truncation distances were selected specific to the individual data sets: acoustic-based truncation was 22.9 m, human-based audio-only at 47.8 m, and human-based audio and visual detections at 48.0 m. For each data set, a hazard rate detection function model without adjustment terms or covariates was selected (Table A1.3, Fig. A1.3). Detection probability for the acoustic-based detections was 0.848 (95% CI 0.837–0.859), while for the human-based audio-only detections, it was 0.494 (95% CI 0.415–0.587), and for the human-based audio and visual detections, it was 0.451 (95% CI 0.379–0.537).

In general, the density estimates using bioacoustics were lower than those based on point-counts by field observers; however, they were closer to the human-based survey that included only acoustic detections than to those that included both acoustic and visual detections (Table 2). The density for ʻamakihi was 29% lower using automatic acoustic recorders than estimates from human-based acoustic survey (6.02 versus 8.48 individuals/ha). Both confidence intervals substantially overlapped each other and

nearly bracketed the point estimates, where the coefficient of variation for the human-based audio-only density was about half that of the acoustic recorder (%CV of 12.85 and 27.40, respectively).

DISCUSSION

In this study, we describe and test a protocol for using single (i.e., not arrays) automatic acoustic recorders to estimate the density of sound-producing animals in terrestrial ecosystems. Our protocol uses information such as sampling conditions, estimates of the cue rate (i.e., number of cues per minute) of the focal species, and the relationship between the power of the cues and the distance to the recorder. Climatic conditions are easy to gather, but both cue rates and the distance–power relationship require that individuals of the species be easily detected in the field. Thus, this protocol is generalizable to other species, including birds, arthropods, or amphibians, given that this information can be collected.

By following our protocol, we could estimate similar densities with overlapping CIs using automatic sound recorders to those estimated from human observations. Our estimates were higher when the surveys were performed by humans than when density was estimated with acoustic recorders (as in van Wilgenburg et al. 2017). There are a couple of possible explanations for this pattern. The first is that human surveys are based on both aural and visual cues (Buckland 2006, Camp et al. 2009) where nonvocalizing birds can be detected, which is not possible in the acoustic-based estimates. With the correct cue rate, this should not cause a bias, but if the individuals selected for cue rate estimation are more vocal than individuals on average, a bias could occur. This may happen, for example, if the cue rate varies with pairing success (Gibbs and Wenny 1993). Given that we tend to locate animals aurally at first, and if these individuals are more

vocal than others, then this will bias cue rates up, and consequently bias density down. The second reason is that the human-based surveys are much more prone to violating the assumption that animals are detected by observers prior to animals detecting and responding to the observer (Turnock and Quinn 1991, Buckland 2006). Even if the movement were at random (and worse, if directional toward the observer), density estimates would be biased up. It would be interesting to determine if human-based estimates are biased up or acoustic recorder estimates are biased down. However, that would be possible only by evaluating the methods under a known population scenario (i.e., all the individuals of the population are known and marked).

Further, the two density estimation methods we employed have different associated estimation errors. For example, the estimation of the distance to a singing bird by field observers depends largely on the ability of the person performing the surveys (Alldredge et al. 2007, 2008, Kühl and Burghardt 2013), while the automatic estimation is based on the physical properties of the sound transmission; thus, it does not include the subjectivity coming from different observers. Moreover, the proportion of false positives and false negatives can be more precisely estimated in automated surveys than in surveys performed by humans (Guschanski et al. 2009, Miller et al. 2012). We note in passing that false negatives are intrinsically dealt with in distance sampling by the detection function. Provided there are no false positives at/near the point, the detection probability corrects for the calls missed. Automatic surveys also allow much larger effective sample sizes (i.e., sampling during longer time periods), thereby facilitating modeling the detection function and thus improving the signal-to-noise ratio. The overall density estimate is computed as an average of the densities in each sampling station, weighted by the sampling effort (i.e., the number of times each sampler was surveyed). Thus, even though density estimates using automatic and human surveys differed, we cannot consider either method to be more accurate. In our study system, both estimates were reasonably similar; therefore, long-term monitoring studies based on one or the other method may presumably be used to reliably compare density estimates and track trends over time.

In this study, we estimated the cue rate and the relationship between the distance of the individuals and the power of the sound in the same area and time of the surveys, thus reducing errors associated with temporal (e.g., due to seasonality) and spatial (e.g., due to between-population differences) variability in cue rates (LaPerriere and Haugen 1972, McShea and Rappole 1997, Marques et al. 2013). It is important that future studies using this method or a similar one also use data taken in the same conditions and at the same time. Another important factor comes from the presence of nonvocalizing individuals because the automatic acoustic recorder does not detect them. We tried to minimize this effect when estimating the cue rates by looking for nonvocalizing individuals and including the time they were observed in our estimates. However, our cue rates are unavoidably biased because it is easier to find an individual that is vocalizing than one that is not. Also, several studies have already indicated that sampling conditions may affect the density estimation (e.g., Baumgartner and Fratantoni 2008, Marques et al. 2011). Birds may change their cue rates with strong wind or rain, and the noise produced by wind and rain may obscure some cues, particularly

quiet or distant vocalizations. Pairing the automatic acoustic recordings with a weather station was useful to account for the bias in the estimates produced by climatic conditions because it provided us with very simple corrections for the acoustic rates and for the distance estimates. This also allowed us to a priori drop the sampling periods when detections were not optimal, which minimizes variability.

Another source of variability in the density estimation may come from the calculation of the distance from the bird to the recorder. The true distances recorded for ´amakihi were made with relatively little measurement error (see Fig. 3). As the true distance increases, the variation in power decreases; calls from far away always arrive with lower power, while calls from close by will typically but not always have higher power. The consequence is that for an observed low power of 40 dB, we predict a distance of about 25 m, when in fact the observed distances were mixed, occurring at small (e.g., 4 m), moderate (e.g., 20 m), and large distances (e.g., 49 m). One of the causes of this variability may come from the orientation of the vocalizing individual because it may modify acoustic parameters of the cue (Patricelli et al. 2008), but our method assumes signal propagation occurs uniformly across all directions. However, if all cues are oriented randomly with respect to the recorder, the variability in the estimation of the distance will be larger, but the differences in the sound parameters due to the bird orientation will average over all the positions, and the mean prediction will be unbiased. Source level or transmission loss patterns (temporal or spatial) will also cause the same received level to represent several different ranges. In addition, confidence intervals are too narrow when measurement errors are ignored, and when measurement errors are substantial, it may be difficult to fit the distance data adequately, resulting in model misspecification (Borchers et al. 2010). It is important to note that the details on the power–distance relationship shown here are valid only for our species and study system. Other vocalizations may be affected differently by climatic conditions and sound degradations. Also, animal vocalizations degrade with distance much faster in areas with dense vegetation than in open areas (Forrest 1994). Thus, our data can be applied only to areas with similar vegetation structure.

With cue-rate surveys, as with the closely related point-transect methods, even relatively small amounts of measurement error may become problematic. When measurement error is small (coefficient of variation approximately 10%), density and variance estimates are nearly unbiased, and it may be safe to ignore measurement error stemming from the model to predict distance from received levels (Borchers et al. 2010). This condition is unlikely using single automatic acoustic recorders. Incorporating measurement error might be required to reduce bias. Density estimates could be multiplied by a bias correction factor for the effect of measurement error (following the approach proposed by Marques 2004) or in a likelihood framework by Borchers et al. (2010). The former method, however, does not perform well for point-transect methods and by extension, cue-rate surveys. The latter approach of modeling measurement error as a likelihood performs better for point- transect surveys, has practical advantages associated with maximum likelihood estimator theory, and employs standard regression methods and freely available software using generalized linear models and/or generalized additive models in R. Finally, another possibility to include bias is to use posterior estimates from Bayesian methods. While we obtained a model to estimate distance

from detected power, we have not propagated the variance associated with this model through to the density estimates. This perhaps allows a fairer comparison with the human observers (where the distances estimated with errors are also used as true distances); however, our confidence intervals might be too narrow. Therefore, if the goal was more than a proof-of-concept and comparison with the human observers, it would be desirable to propagate the variance in the distance estimation model. That should be straightforward using a nonparametric bootstrap to estimate the distances at each bootstrap iteration.

Our methods will work better when source sound level has lower variability. If the original source levels were highly variable, say across individuals, then there is a confounding between the original sound level and distance in the observed received level. Similarly, the more ominidirectional the source, the more "distance" can be explained by the received sound level. These two characteristics of the species of interest will be determinant in identifying when the methods might work or when there might be too little information on sound level to infer distance; hence, other bypass methods that require ranges might be preferable. Under such contexts, spatially explicit capture-recapture methods (e.g., as in Dawson and Efford 2009) might be useful, but they come at the additional cost of requiring multiple sensors across which detections would have to be matched. Finally, we intentionally avoided the dawn chorus in our sampling to increase the performance of the automatic detection algorithm; however, this may also be a source of bias in the results that should be considered by the users of the protocol (Streby et al. 2012). However, if the density estimates will be compared with others where the dawn chorus is also avoided, the comparisons should be correct.

It is also important to note that our spatial sample size was relatively small (eight stations, one transect). This study was designed to be a proof-of-concept for the protocol and has served to identify strengths and weaknesses of the method. Thus, our results should be taken with caution, and further studies are required for a more accurate and generalizable application of the protocol. In summary, we demonstrated that it is possible to use single automatic acoustic recorders to obtain fairly accurate abundance estimates for terrestrial species that communicate acoustically. This may be useful for collecting large-scale and long-term information on animal populations, particularly those that are rare or that live in remote areas that are difficult to access. Moreover, we identified potential limitations to the approach, and suggested methods to minimize their effect. Our approach can also be easily adapted for use on other sound-producing taxa such as mammals, amphibians, and arthropods.

Responses to this article can be read online at:
http://www.ace-eco.org/issues/responses.php/1224

LITERATURE CITED

Alldredge, M. W., K. Pacifici, T. R. Simons, and K. H. Pollock. 2008. A novel field evaluation of the effectiveness of distance and independent observer sampling to estimate aural avian detection probabilities. *Journal of Applied Ecology* 45:1349-1356. http://dx.doi.org/10.1111/j.1365-2664.2008.01517.x

Alldredge, M. W., T. R. Simons, and K. H. Pollock. 2007. A field evaluation of distance measurement error in auditory avian point count surveys. *Journal of Wildlife Management* 71:2759-2766. http://dx.doi.org/10.2193/2006-161

Baumgartner, M. F., and D. M. Fratantoni. 2008 Diel periodicity in both sei whale vocalization rates and the vertical migration of their copepod prey observed from ocean gliders. *Limnology and Oceanography* 53:2197-2209. http://dx.doi.org/10.4319/lo.2008.53.5_part_2.2197

Blumstein, D. T., D. J. Mennill, P. Clemins, L. Girod, K. Yao, G. Patricelli, J. L. Deppe, A. H. Krakauer, C. Clark, K. A. Cortopassi, S. F. Hansen, B. McCowan, A. M. Ali, and A. N. G. Kirschel. 2011. Acoustic monitoring in terrestrial environments using microphone arrays: applications, technological considerations and prospectus. *Journal of Applied Ecology* 48:758-767. http://dx.doi.org/10.1111/j.1365-2664.2011.01993.x

Borchers, D. L., T. A. Marques, T. Gunnlaugsson, and P. E. Jupp. 2010. Estimating distance sampling detection functions when distances are measured with errors. *Journal of Agricultural, Biological, and Environmental Statistics* 15:346-361. http://dx.doi.org/10.1007/s13253-010-0021-y

Buckland, S. T. 2006. Point-transect surveys for songbirds: robust methodologies. *Auk* 123:345- 357. http://dx.doi.org/10.1642/0004-8038(2006)123[345:PSFSRM]2.0.CO;2

Buckland, S. T., D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas. 2001. *Introduction to distance sampling*. Oxford University Press, Oxford, UK.

Buckland, S. T., D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas. 2004. *Advanced distance sampling*. Oxford University Press, Oxford, UK.

Buckland, S. T., E. A. Rexstad, T. A. Marques, and C. S. Oedekoven. 2015. *Distance sampling: methods and applications*. Springer, New York, USA. http://dx.doi.org/10.1007/978-3-319-19219-2

Burnham, K. P., D. R. Anderson, and J. L. Laake. 1980. Estimation of density from line transect sampling of biological populations. *Wildlife Monographs* 72:3-202.

Camp, R. J., K. W. Brink, P. M. Gorresen, and E. H. Paxton. 2016. Evaluating abundance and trends in a Hawaiian avian

community using state-space analysis. *Bird Conservation International* 26:225-242. http://dx.doi.org/10.1017/S0959270915000088

Camp, R. J., M. H. Reynolds, P. M. Gorresen, T. K. Pratt, and B. L. Woodworth. 2009. Pages 83-107 *in* T. K. Pratt, C. T. Atkinson, P. C. Banko, J. D. Jacobi, and B. L. Woodworth, editors. *Monitoring Hawaiian forest birds. Conservation biology of hawaiian forest birds: implications for island avifauna*. Yale University Press, New Haven, Connecticut, USA.

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. http://dx.doi.org/10.1111/j.1557-9263.2009.00206.x

Cortes, C., and V. Vapnik. 2009. Support-vector networks. *Machine Learning* 20:273-297. http://dx.doi.org/10.1007/BF00994018

Dawson, D. K., and M. G. Efford. 2009. Bird population density estimated from acoustic signals. *Journal of Applied Ecology* 46:1201-1209. http://dx.doi.org/10.1111/j.1365-2664.2009.01731.x

Dénes, F. V., L. F. Silveira, and S. R. Beissinger. 2015. Estimating abundance of unmarked animal populations: accounting for imperfect detection and other sources of zero inflation. *Methods in Ecology and Evolution* 6:543-556. http://dx.doi.org/10.1111/2041-210X.12333

Efford, M. G., D. K. Dawson, and D. L. Borchers. 2009. Population density estimated from locations of individuals on a passive detector array. *Ecology* 90:2676-2682. http://dx.doi.org/10.1890/08-1735.1

Faanes, C. A., and D. Bystrak. 1981. The role of observer bias in the North American Breeding Bird Survey. *Studies in Avian Biology* 6:353-359.

Féret, J. B., and G. P. Asner. 2012. Tree species discrimination in tropical forests using airborne imaging spectroscopy. *IEEE Transactions on Geoscience and Remote Sensing* 51:73-84. http://dx.doi.org/10.1109/TGRS.2012.2199323

Forrest, T. G. 1994. From sender to receiver: propagation and environmental effects on acoustic signals. *Integrative and Comparative Biology* 34:644-654. http://dx.doi.org/10.1093/icb/34.6.644

Gates, C. E., and W. B. Smith. 1972. Estimation of density of Mourning Doves from aural information. *Biometrics* 28:345-359. http://dx.doi.org/10.2307/2556152

Gibbs, J. P., and D. G. Wenny. 1993. Song output as a population estimator: effect of male pairing status. *Journal of Field Ornithology* 64:316-322.

Guschanski, K., L. Vigilant, A. McNeilage, M. Gray, E. Kagoda, and M. M. Robbins. 2009. Counting elusive animals: comparing field and genetic census of the entire mountain gorilla population of Bwindi Impenetrable National Park, Uganda. *Biological Conservation* 142:290-300. http://dx.doi.org/10.1016/j.biocon.2008.10.024

Harris, D., L. Matias, L. Thomas, J. Harwood, and W. Geissler. 2013. Applying distance sampling to fin whale calls recorded by single seismic instruments in the northeast Atlantic. *Journal of the Acoustical Society of America* 134:3522-3535. http://dx.doi.org/10.1121/1.4821207

Hildebrand, J. A., S. Baumann-Pickering, K. E. Frasier, J. S. Trickey, K. P. Merkens, S. M. Wiggins, M. A. McDonald, L. P. Garrison, D. Harris, T. A. Marques, and L. Thomas. 2015. Passive acoustic monitoring of beaked whale densities in the Gulf of Mexico. *Scientific Reports* 5:16343.

Kellner, K. F., and R. K. Swihart. 2014. Accounting for imperfect detection in ecology: a quantitative review. *PLoS ONE* 9:e111436. http://dx.doi.org/10.1371/journal.pone.0111436

Knight, E. C., K. C. Hannah, G. 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(2):14. http://dx.doi.org/10.5751/ACE-01114-120214

Kühl, H. S., and T. Burghardt. 2013. Animal biometrics: quantifying and detecting phenotypic appearance. *Trends in Ecology & Evolution* 28:432-441. http://dx.doi.org/10.1016/j.tree.2013.02.013

Küsel, E. T., D. K. Mellinger, L. Thomas, T. A. Marques, D. J. Moretti, and J. Ward. 2011. Cetacean population density from single fixed sensors using passive acoustics. *Journal of the Acoustical Society of America* 129:3610-3622.

Kyhn, L. A., J. Tougaard, L. Thomas, L. R. Duve, J. Stein-back, M. Amundin, G. Desportes, and J. Teilmann. 2012. From echolocation clicks to animal density – acoustic sampling of harbour porpoises with static dataloggers. *Journal of the Acoustical Society of America* 131:550-560.

Laperriere, A. J., and A. O. Haugen. 1972. Some factors influencing calling activity of wild Mourning Doves. *Journal of Wildlife Management* 36:1193-1199. http://dx.doi.org/10.2307/3799248

Lindsey, G. D., E. A. VanderWerf, H. Baker, and P. E. Baker. 1998. Hawaii Amakihi (*Hemignathus virens*), Kauai Amakihi (*Hemignathus kauaiensis*), Oahu Amakihi (*Hemignathus chloris*), and Greater Amakihi (*Hemignathus sigittirostris*). *In* A. Poole and F. Gill, editors. *The birds of North America*. No. 360. Cornell Lab of Ornithology, Cornell University, Ithaca, New York, USA.

Lynch, A. 1996. The population memetics of bird song. *Ecology* 15:181-197.

MacKenzie, D. L., and W. L. Kendall. 2002. How should detection probability be incorporated into estimates of relative abundance? *Ecology* 83:2387-2393. http://dx.doi.org/10.1890/0012-9658(2002)083[2387:HSDPBI]2.0.CO;2

Marques, T. A. 2004. Predicting and correcting bias caused by measurement error in line transect sampling using multiplicative error models. *Biometrics* 60:757-763. http://dx.doi.org/10.1111/j.0006-341X.2004.00226.x

Marques, T. A., L. Munger, L. Thomas, S. Wiggins, and J. A. Hildebrand. 2011. Estimating North Pacific right whale *Eubalaena japonica* density using passive acoustic cue counting. *Endangered Species Research* 13:163-172. http://dx.doi.org/10.3354/esr00325

Marques, T. A., L. Thomas, M. Kéry, S. T. Buckland, D. L. Borchers, E. Rexstad, R. M. Fewster, D. I. MacKenzie, J. A. Royle, G. Guillera-Arroita, C. M. Handel, D. C. Pavlacky, Jr. and R. J.

Camp. 2017. Model-based approaches to deal with detectability: a comment to Hutto (2016a). *Ecological Applications* 27:1694-1698. http://dx.doi.org/10.1002/eap.1553

Marques, T. A., L. Thomas, S. W. Martin, D. K. Mellinger, J. A. Ward, D. J. Moretti, D. Harris, and P. L. Tyack. 2013. Estimating animal population density using passive acoustics. *Biological Reviews* 88:287-309. http://dx.doi.org/10.1111/brv.12001

Marques, T. A., L. Thomas, J. Ward, N. DiMarzio, and P. L. Tyack. 2009. Estimating cetacean population density using fixed passive acoustic sensors: an example with Blainville's beaked whales. *Journal of the Acoustical Society of America* 125:1982-1994. http://dx.doi.org/10.1121/1.3089590

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.

Miller, D. A. W., L. A. Weir, B. T. McClintock, E. H. C. Grant, L. L. Bailey, and T. R. Simons. 2012. Experimental investigation of false positive errors in auditory species occurrence surveys. *Ecological Applications* 22:1665-1674. http://dx.doi.org/10.1890/11-2129.1

Mills, H. G. 2000. Geographically distributed acoustical monitoring of migrating birds. *Journal of the Acoustic Society of America* 108:2582. http://dx.doi.org/10.1121/1.4743594

Oppel, S., S. Hervías, N. Oliveira, T. Pipa, C. Silva, P. Geraldes, M. Goh, E. Immler, and M. McKown. 2014. Estimating population size of a nocturnal burrow-nesting seabird using acoustic monitoring and habitat mapping. *Nature Conservation* 7:1-13. http://dx.doi.org/10.3897/natureconservation.7.6890

Patricelli, G. L., M. S. Dantzker, and J. W. Bradbury. 2008. Acoustic directionality of Red-winged Blackbird (*Agelaius phoeniceus*) song relates to amplitude and singing behaviours. *Animal Behavior* 76:1389-1401. http://dx.doi.org/10.1016/j.anbehav.2008.07.005

R Development Core Team. 2009. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Sebastián-González, E., J. Pang-Ching, J. M. Barbosa, and P. J. Hart. 2015. Bioacoustics for species management: two case studies with a Hawaiian forest bird. *Ecology and Evolution* 5:4696-4705. http://dx.doi.org/10.1002/ece3.1743

Sebastián-González, E., J. van Aardt, K. Sacca, J. M. Barbosa, K. Kelbe, and P. J. Hart. 2018. Testing the acoustic adaptation hypothesis with native and introduced birds in Hawaiian forests. *Journal of Ornithology*. http://dx.doi.org/10.1007/s10336-018-1542-3

Seber, G. A. F. 1986. A review of estimating animal abundance. *Biometrics* 42:267-292. http://dx.doi.org/10.2307/2531049

Shonfield, J., and E. M. Bayne. 2017. Autonomous recording units in avian ecological research: current use and future applications. *Avian Conservation and Ecology* 12(1):14. http://dx.doi.org/10.5751/ACE-00974 -120114

Stevenson, B. C., D. L. Borchers, R. Altwegg, R. J. Swift, D. M. Gillespie, and G. J. Measey. 2015. A general framework for animal density estimation from acoustic detections across a fixed microphone array. *Methods in Ecology and Evolution* 6:38-48. http://dx.doi.org/10.1111/2041-210X.12291

Streby, H. M., J. P. Loegering, and D. E. Andersen. 2012. Spot-mapping underestimates song-territory size and use of mature forest by breeding Golden-winged Warblers in Minnesota, USA. *Wildlife Society Bulletin* 36:40-46. http://dx.doi.org/10.1002/wsb.118

Thomas, L., S. T. Buckland, E. A. Rexstad, J. L. Laake, S. Strindberg, S. L. Hedley, J. R. B. Bishop, T. A. Marques, and K. P. Burnham. 2010. Distance software: design and analysis of distance sampling surveys for estimating population size. *Journal of Applied Ecology* 47:5-14. http://dx.doi.org/10.1111/j.1365-2664.2009.01737.x

Turnock, B. J., and T. J. Quinn II. 1991. The effect of responsive movement on abundance estimation using line transect sampling. *Biometrics* 47:701-715. http://dx.doi.org/10.2307/2532156

Van Parijs, S. M., C. W. Clark, R. S. Sousa-Lima, S. E. Parks, S. Rankin, D. Risch, and I. C. Van Opzeeland. 2009. Management and research application of real-time and archival passive acoustic sensors over varying temporal and spatial scales. *Marine Ecology Progress Series* 395:21-36. http://dx.doi.org/10.3354/meps08123

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(1):13. http://dx.doi.org/10.5751/ACE-00975 -120113

Verner, J. 1985. Assessment of counting methods. *Current Ornithology* 2:247-302. http://dx.doi.org/10.1007/978-1- 4613-2385-3_8

Appendix 1. Supplementary tables and figures

Table A1.1. Sample sizes for the different recording stations for both human and acoustic samples.

Station	Human samples	Acoustic samples
1	9	405
2	10	419
3	10	101
4	10	456
5	9	657
6	9	437
7	9	655
8	9	392

Table A1.2. Details on the configuration of the Band Limited Energy Detector in Raven. SNRt: Sound to Noise Ratio. Occupancy: the minimum proportion of pixels that are above the SNR level. Block size: Width of block size used to calculate background noise. Hop size: Space between noise calculations. Percentage: Percentage of the ranked noise accounted as background.

Minimum frequency (Hz)	1800
Maximum frequency (Hz)	9500
Minimum time (s)	0.108
Maximum time (s)	1
Minimum delta time (s)	0.102
SNRt	6
Occupancy	30
Block size	2.0
Hop size	0.5
Percentage	20

Table A1.3. Model parameters and model-selection results for Hawai'i 'amakihi for human-based surveys (audio-only and audio+visual detections) and acoustic recorder surveys. Models were sorted by differences in second-order Akaike's information criterion corrected for small sample size (ΔAICc) between each candidate model and the model with the lowest AICc value. Models examined included half-normal (HN) and hazard-rate (HR) key detection functions. None of the models incorporating series expansions converged. Preliminary analysis revealed that covariates did not improve AICc values. The number of estimated parameters (# Params), and negative log-likelihood (-LogL) are presented.

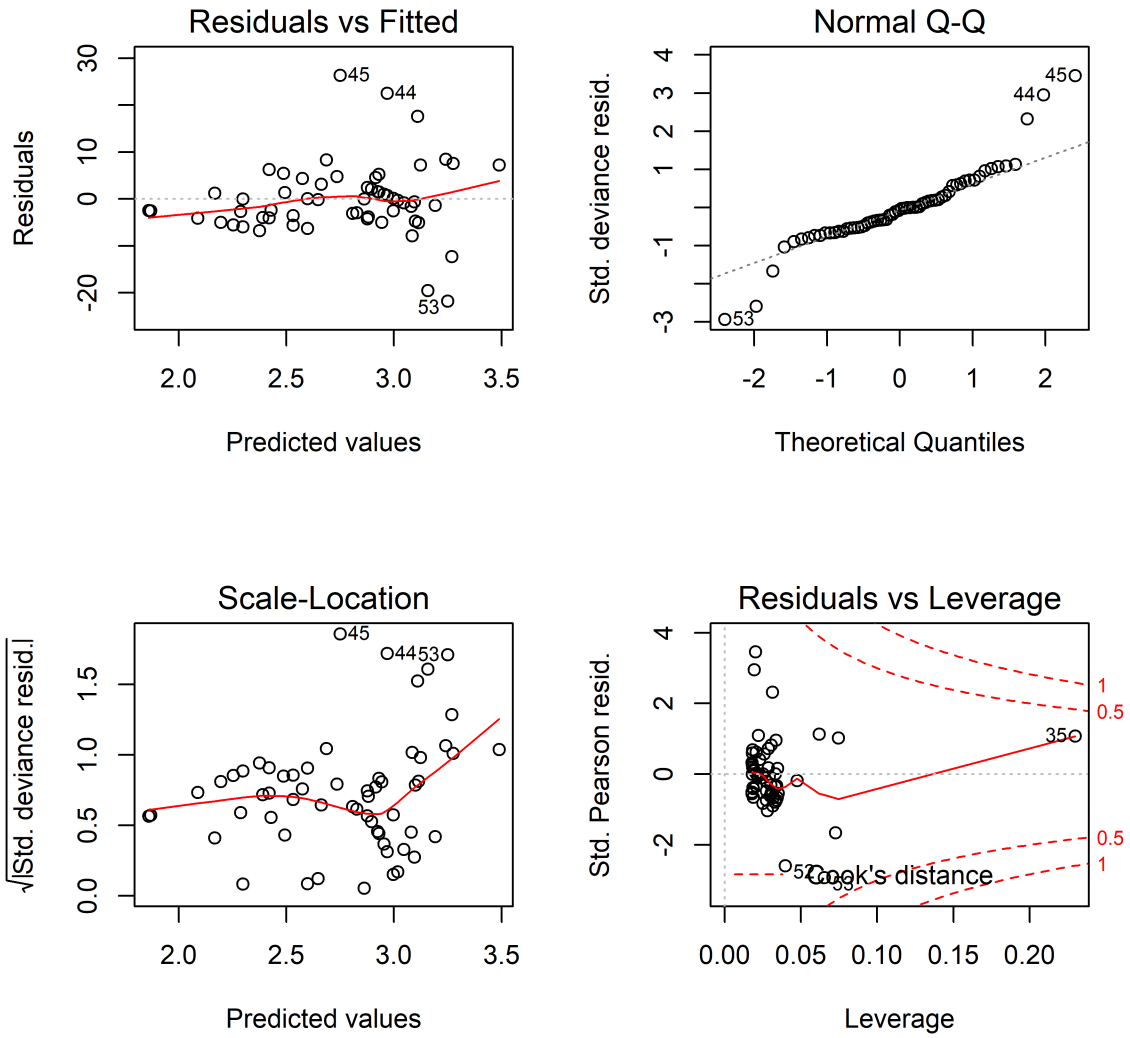| Detection function	# Params	-LogL	AICc	ΔAICc
Human-based Audio-only				
HR	2	830.76	1665.57	0
HN	1	836.66	1675.34	9.77
Human-based Audio+Visual				
HR	2	945.45	1894.96	0
HN	1	951.01	1904.04	9.08
Acoustic recorder-based				
HR	2	19810.26	39624.52	0
HN	1	20097.11	40196.22	571.7

Fig A1.1. Model diagnostics for the relationship between the power of the sound (m) and the distance to the vocalizing bird (m) represented in Fig. 3.
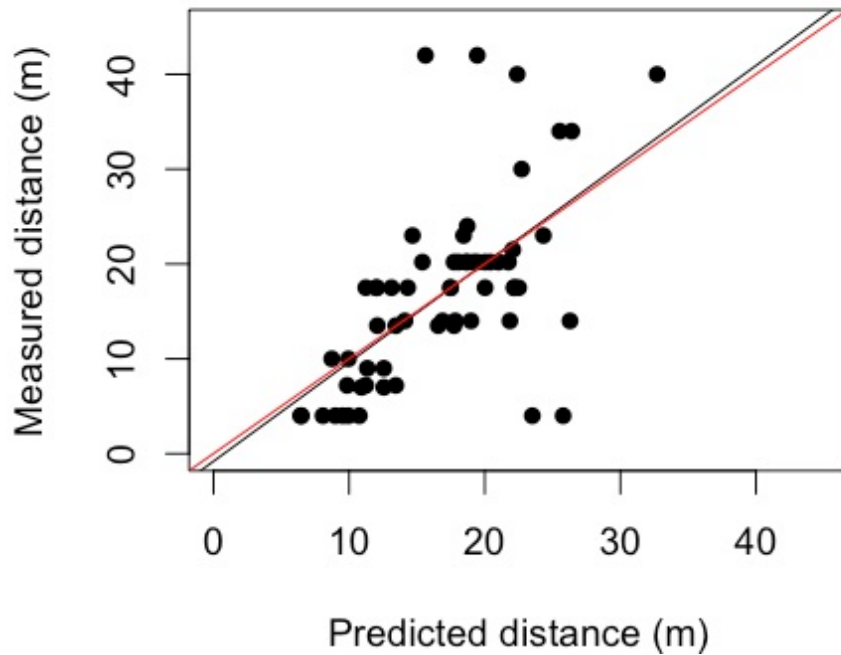
Figure A1.2. Graph showing the relationship between the measured distance from the recorder to the bird and the distance predicted by the model. We also show the regression between the two variables (black line) and the 1:1 relationship (red line).
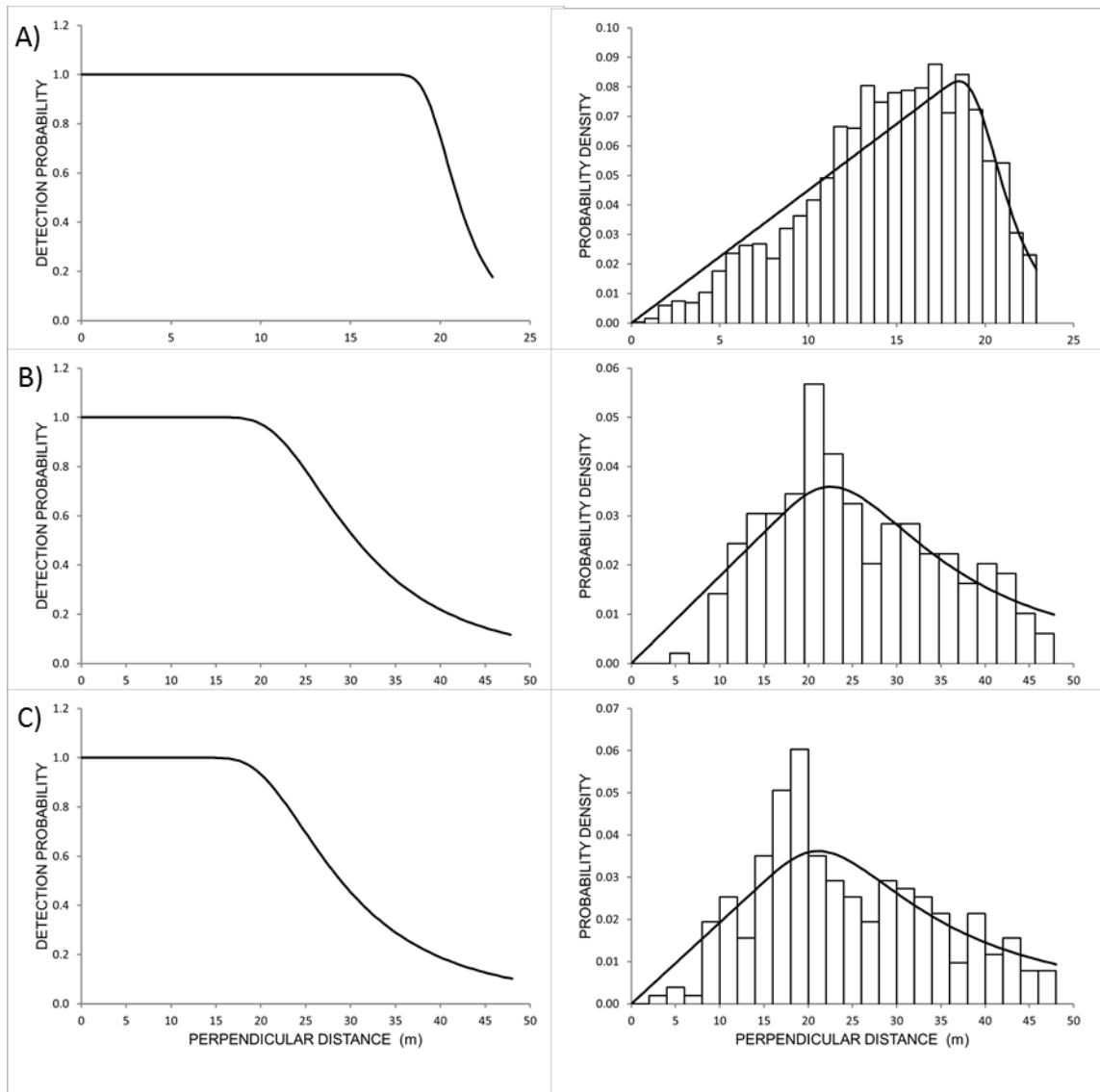
Figure A1.3. Detection probability and probability density plots overlaid on detection distance histogram for fitted detection function model to (A) acoustic-based with truncation at 22.9 m, (B) human-based audio-only with truncation at 47.8 m, and (C) human-based audio and visual data with truncation at 48.0 m. For each data set a hazard rate model without adjustment terms or covariates was selected (AICc values were 571.70, 9.77 and 9.06 units better than next best approximating model, respectively).