Citation:

Wilson DR, Battiston M, Brzustowski J, Mennill DJ (2014) Sound Finder: a new software approach for localizing animals recorded with a microphone array. *Bioacoustics*, 23: 99–112. doi: 10.1080/09524622.2013.827588

Title: Sound Finder: A New Software Approach for Localizing Animals Recorded with a Microphone Array

Authors: [1] David R. Wilson (corresponding author), [1] Matthew Battiston, [2] John Brzustowski, [1] Daniel J. Mennill

Affiliation: [1] Department of Biological Sciences, University of Windsor, Windsor, ON N9B 3P4, Canada

[2] Department of Biology, Acadia University, Wolfville, NS B4P 2R6, Canada

Contact: D. Wilson (email: drwilson76@gmail.com; phone: 519-253-3000 ext. 2709), M. Battiston (email: mattbattist@gmail.com; phone: 519-253-3000 ext. 4725), J. Brzustowski (email: jbrzusto@fastmail.fm), D. Mennill (email: dmennill@uwindsor.ca; phone: 519-253-3000 ext. 4726)

Word count: 5861 words, excluding title page, references, tables, and figure legends

Acoustic localization is a powerful technique for monitoring the positions, movements, and behaviours of terrestrial animals. However, its prevalence in biological studies has been constrained by hardware and software that are custom-built, expensive, and difficult to use. We recently helped to relieve the hardware constraint by describing a microphone array that is affordable, portable, easy to use, and commercially available. Here, we help to relieve the software constraint by developing an acoustic localization program called “Sound Finder,” which is easy to use, freely available, and accurate for a variety of animals and recording conditions. It runs in the free software environment R, and in spreadsheet programs such as Microsoft Excel and the open source software LibreOffice. In this study, we describe how Sound Finder functions, and then test its accuracy by localizing natural sounds that were broadcast through loudspeakers and re-recorded with microphone arrays. We quantify Sound Finder’s accuracy by comparing its location estimates to known loudspeaker locations and to output from other localization approaches. We show that Sound Finder generates accurate location estimates for a variety of animal sounds, microphone array configurations, and environmental conditions. Furthermore, Sound Finder generates an error value that allows the user to assess its accuracy. In conclusion, Sound Finder provides accurate estimates of a vocalizing animal’s location. It is easy to use, requires only widespread and affordable software, and is freely available in a standard form as supplemental material to this paper.

Keywords: acoustic monitoring; multi-channel recording; radio tracking; Sound Finder; triangulation

Introduction

Behavioural biologists can gain critical insight by monitoring animal movements. For example, spatial data can shed light on social behaviour by revealing where and with whom an animal interacts (Rutz et al. 2012), on reproductive behaviour by showing where an animal defends its territory and seeks mating opportunities (Double and Cockburn 2000), and on foraging behaviour by elucidating an animal's food-searching strategies (Makino and Sakai 2004). Researchers can use a variety of methods to monitor animal movements in natural terrestrial environments, but each method has its own set of advantages and disadvantages. For example, observing animals directly can be a simple and reliable method, but the presence of human observers can inadvertantly affect an animal's behaviour ($\mathrm{M^c}$Donald et al. 2007). Furthermore, direct observation may not be possible over long periods of time, for large numbers of subjects, for cryptic species, for animals in visually occluded habitats, or for animals that are active at night. Radio tracking techniques can resolve these issues, but capturing the animal and fitting it with a radio transmitter can adversely influence the animal's behaviour (reviewed in Mech and Barber 2002). For smaller animals, the transmitter's weight and battery life can also be limiting factors.

Acoustic localization is a promising new technique for monitoring animal movements in terrestrial environments (reviewed in Blumstein et al. 2011). This technique uses an array of three or more microphones to localize animals based on the sounds they produce. Because sound travels at a predictable rate through air, the time required for a signal to reach each microphone will vary according to the signalling animal's position. The time-of-arrival-differences among simultaneously recording microphones can then be used to determine the location of the signalling animal. The benefits of acoustic localization are that it allows researchers to monitor the movements of multiple individuals over long periods of time, across large geographic areas (depending on the number of microphones, microphone density, and the active space of the signal of interest), in the absence of human observers, and in habitats where other monitoring techniques might be impossible. Furthermore, acoustic localization does not require animals to be captured or fitted with transmitters, so their behaviour will not be affected by this passive monitoring technique. The disadvantages of acoustic localization are that animals can only be localized when they emit sound, and that individuals can only be distinguished from each other if they produce individually distinctive signals or are associated with a particular location (e.g., territorial animals; Blumstein et al. 2011; Mennill 2011).

Acoustic localization involves three fundamental steps (Magyaret al. 1978; depicted in Fig. 1). First, a sound must be recorded with an array of at least three (two-dimensional localization) or four (three-dimensional localization) microphones (see example of a 4-channel array in Fig. 1a). The locations of the microphones must be measured precisely, and the recordings corresponding to the microphones must be synchronized with millisecond or sub-millisecond resolution. Traditionally, recordings have been synchronized by connecting the microphones via cables to a central multichannel recording device (e.g., Magyar et al. 1978; Mennill et al. 2006), though arrays composed of synchronized wireless recorders are now also possible (e.g., Collier et al. 2010; Mennill et al. 2012; Fig. 1a). Second, the sound’s time-of-arrival-differences among the microphones must be measured from the recordings with a high level of precision (Spiesberger and Fristrup 1990). This can be achieved using cross-correlation techniques (see description in Methods subsection 'Measuring Time-of-arrival-differences'; Fig. 1b-d), which are available in sound analysis software programs such as Raven Pro Interactive Sound Analysis Software (Cornell Lab of Ornithology Bioacoustics Research Program, Ithaca, NY, USA), Avisoft-SASLab Pro (Avisoft Bioacoustics, Berlin, Germany), and SIGNAL (Engineering Design, Belmont, MA, USA). Third, the time-of-arrival-differences must be used to estimate the location of the sound's source using one of several different mathematical approaches (details in Spiesberger and Fristrup 1990; Mennill et al. 2006; Collier et al. 2010).

Despite the benefits of acoustic localization, two constraints have limited its widespread use as a tool for studying the spatial ecology of animals. First, the hardware comprising microphone arrays has traditionally been expensive, custom-manufactured, and difficult to deploy. Fortunately, these hardware constraints have recently been ameliorated by a new cable-free microphone array technology that is affordable, commercially available, and easy to use (Mennill et al. 2012). Second, the acoustic localization software needed to convert time-of-arrival-differences to estimates of an animal's location is not available commercially or from the peer-reviewed scientific literature. In previous studies (see, for example, all studies listed in the review by Blumstein et al. 2011), the software for conducting acoustic localization has been custom-written by individual authors and is now either unavailable or available only upon request from the authors. Consequently, the accuracy of such software may not be known and may change as the authors modify their software. Furthermore, existing software programs may be inaccessible to many biologists because they require expensive and advanced computer software environments such as MatLab (e.g., Mennill et al. 2006). Another disadvantage of previous custom-written software solutions is that they are often highly tailored to one specific animal, and their general applicability has never been assessed. The scientific community therefore has a pressing need for acoustic localization software that is affordable, accurate, easy to use, applicable to a variety of animals and environmental conditions, and available in a standard form.

In this methodological study, we developed an acoustic localization program called “Sound Finder,” which relies only on affordable software packages that are already owned by most research laboratories. We provide Sound Finder as supplemental material to this paper to ensure that all researchers will have perpetual access to the same version of the program, and that the published version will be the same as the one we describe in this paper. We also test Sound Finder’s accuracy by localizing natural sounds that had been broadcast through a loudspeaker and re-recorded with various microphone arrays. We localize a variety of animal sounds that had been recorded with a variety of microphone array configurations, including arrays that did or did not rely on microphone cables; arrays that had four, eight, or 16 microphones; arrays that were located in tropical or temperate environments; and arrays that were located in forested or open habitats. Our specific objectives are: (1) to describe how Sound Finder works, (2) to assess the accuracy of the location estimates provided by Sound Finder, (3) to determine if Sound Finder's error value can predict the localization accuracy of those location estimates, and (4) to compare the accuracy of Sound Finder to the accuracy of one of the most commonly used acoustic localization software approaches from previous studies.

Methods

Part 1: Sound Finder

Sound Finder is a computer software program that is available in two versions. The first version runs in the freely available software environment, R, which runs on a variety of UNIX, Windows, and Macintosh operating systems (R Core Team 2013; www.r-project.org). The second version runs in a variety of spreadsheet programs, including Microsoft Excel (Microsoft Corporation, Mountain View, CA, USA) and the freely available LibreOffice (www.libreoffice.org). We have run the spreadsheet version of Sound Finder successfully on multiple operating systems (Windows 7, Windows XP, Mac OS X) and in multiple spreadsheet programs (including Excel 2003, Excel 2007, Excel 2010, Excel X for Mac, Excel 2004 for Mac, Excel 2011 for Mac, and LibreOffice 4). We note that Sound Finder's batch processing feature does not function in Microsoft Excel 2008 for Mac, since Visual Basic is not contained in this version of Excel; however, sounds can still be localized individually in this version of Excel. The R version of Sound Finder, as well as example data and instructions for its use, are included in the supplemental material in a file entitled "S1 Sound Finder for R.zip". The spreadsheet version of Sound Finder, as well as example data and specific instructions for its use, are included in the supplemental material in a single Microsoft Excel workbook entitled "S2 Sound Finder for Spreadsheets.xls". Any future updates to Sound Finder will be hosted at http://discovery.acadiau.ca/R/SoundFinder/.

Sound Finder localizes sounds in two-dimensional or three-dimensional space using data that the user enters into a text file (R version) or an Excel worksheet (spreadsheet version). For each sound to be localized, the user enters the temperature at the time of recording, and the latitude, longitude, and altitude (altitude is necessary only for three-dimensional microphone arrays) of each microphone in the array (maximum = 64 microphones). The user also enters the time-of-arrival-differences of the sound at each microphone, having calculated these differences from other software. Time-of-arrival-differences can be generated using cross-correlation techniques that are available in several sound analysis software programs (see Introduction), including Raven Pro Interactive Sound Analysis Software (version 1.4), which we used here.

Sound Finder uses an automated batch process to estimate the origin of each sound specified by the user. First, Sound Finder uses the temperature at the time of recording to calculate the speed of sound, following the formula presented in Wölfel and McDonough (2009):

equation 1: $$\text{speed of sound (m/s)} = 331.5 * [(\text{temperature (ºC)} + 273.15) / 273.15]^{0.5}$$

Sound Finder does not consider humidity at the time of recording because humidity has negligible effects on the speed of sound (Wölfel and McDonough 2009). Second, Sound Finder estimates the location of the sound source by applying the least-squares solution that was developed for global positioning systems (Bancroft 1985; see also Muanke and Niezrecki 2007). Sound Finder automatically localizes sounds in three dimensions when the user provides altitude coordinates for the microphones in the array; if altitude is not provided, Sound Finder localizes sounds in two dimensions. Third, Sound Finder generates numerical output, including the latitude, longitude, and altitude of the sound's origin, the time at which the sound was produced relative to when it was detected at the first microphone, and an estimate of the error associated with the localization. Higher error values indicate lower confidence in the accuracy of the localization. The output is stored in a text file in the R version of Sound Finder and in a separate worksheet in the spreadsheet version of Sound Finder. The numerical output from Sound Finder can then be visualized in any mapping software, such as ArcGIS (Esri, Redlands, CA, USA).

Part 2: Accuracy of Sound Finder

2.1 Sounds used to test Sound Finder

We tested the accuracy of Sound Finder by localizing animal sounds that had been broadcast through loudspeakers from known positions and re-recorded with a microphone array (see supplemental material, "S3 Sound Clips.zip"). In total, we used three different microphone array configurations, which we set up at 38 different locations during three previous studies (full details in Mennill et al. 2006; Mennill and Vehrencamp 2008; Lapierre et al. 2011; Mennill et al. 2012). In the first study, we set up an array of eight omnidirectional microphones at 20 different locations in a dense tropical forest habitat in Costa Rica. The average area bounded by each microphone array was 1.30 ha, and the average microphone density was 6.2 microphones/ha. The microphones were connected via cables to a centrally located computer that recorded the signals into an 8-channel audio file (Mennill et al. 2006). In the second study, we set up an array of 16 omnidirectional microphones at six different locations in an open field habitat in eastern Ontario, and, again, the microphones were connected via cables to a centrally located computer that recorded the signals into a single 16-channel audio file (Lapierre et al. 2011). The average area bounded by each microphone array was 6.65 ha, and the average microphone density was 2.4 microphones/ha. In the third study, we set up an array of four microphones in six open field locations and six forest locations in a temperate environment in southern Ontario (see Fig. 1a). The average area bounded by each microphone array in this study was 0.14 ha, and the average microphone density was 28.6 microphones/ha. The microphones in this study were mounted directly on independent digital recorders that were synchronized with a GPS signal. After recording was complete, we combined the four time-synchronized single-channel audio recordings into a single 4-channel audio file (Mennill et al. 2012).

We broadcast a different pre-recorded animal signal at two different locations in each of the 38 microphone arrays, resulting in 76 unique playback locations that we could attempt to localize with Sound Finder (see supplemental material, "S3 Sound Clips.zip"). In the first study (Mennill et al. 2006), one loudspeaker played the song of a male rufous-and-white wren (*Thryophilus rufalbus*) and the other played the song of a female rufous-and-white wren. In the second study (Lapierre et al. 2011), each loudspeaker played the song of a different male song sparrow (*Melospiza melodia*). In the third study (Mennill et al. 2012), one loudspeaker played the advertisement call of a male grey treefrog (*Hyla versicolor*) and the other played the advertisement call of a male spring peeper (*Pseudacris crucifer*). Birdsong stimuli were broadcast at a natural amplitude of 80 dB SPL, and frog call stimuli were broadcast at a natural amplitude of 90 dB SPL (measured 1 m from the loudspeaker with a RadioShack sound level meter; RadioShack Corporation, Fort Worth, TX, USA). Mennill et al. (2012) provide spectrograms and detailed descriptions of all five types of playback stimuli. To create a diversity of sound source locations, we positioned the loudspeakers inside the array for four types of stimuli (male and female rufous-and-white wren solo songs, male song sparrow songs, grey treefrog advertisement calls) and in the 50-m boundary surrounding the array for the fifth type of stimulus (spring peeper advertisement calls). By broadcasting natural animal sounds at natural amplitudes in natural habitats containing other vocalizing animals, we were able to conduct a realistic test of Sound Finder's performance under a variety of natural recording conditions.

We used a survey-grade Global Positioning System [Ashtech ProMark II in Mennill et al. (2006); Ashtech ProMark III in Lapierre et al. (2011) and Mennill et al. (2012); Santa Clara, CA, USA] to measure the actual locations of the microphones and loudspeakers used in our study. Resulting measurements had a horizontal accuracy of 1.26 ± 1.08 m (mean ± SD) for microphone positions, and 1.80 ± 0.71 m for loudspeaker positions. We do not report vertical accuracy because all microphones and speakers within a given array were placed on a horizontal plane.

We used Syrinx-PC sound analysis software (version 2.6h; J. Burt, Seattle, WA, USA) to browse through the long multi-channel recordings and identify the playback stimuli of interest. We then extracted each stimulus across all of the recording channels (see example in Fig. 1b). Because the playback stimulus reached each microphone at a slightly different time, we selected the beginning and end of each clip such that the clip included the entire playback stimulus in all of the channels in which the stimulus was audible. Clips were saved as 76 separate multichannel WAVE files (16-bit amplitude encoding, 22050 Hz sampling rate) for use in subsequent analyses. We saved our WAVE files with a sampling rate of 22050 Hz because the maximum frequency of our playback stimuli never exceeded the Nyquist frequency of 11025 Hz

(Mennill et al. 2012). Although we were not interested in the absolute times at which stimuli were recorded, we note that such information can easily be preserved during the clipping process and throughout the entire localization process. Specifically, users can name each clip with the name of its parent soundfile and the exact time at which the clip occurs within that file (e.g., "arrayrecording1.1hour.32min.WAV"). Programs such as Raven Pro Interactive Sound Analysis Software can even apply such file naming conventions automatically when extracting multiple clips from long recordings. If the original recordings are calibrated according to the Universal Time Code, then the clip's true time can also be preserved by including it in the clip's filename.

2.2 Measuring time-of-arrival-differences

Spectrographic cross-correlation is a method for comparing the similarity of two spectrograms (see Fig. 1b-d). This technique involves overlaying two spectrograms and incrementally sliding one past the other in time while calculating a correlation coefficient at each time offset. The correlation coefficients are plotted as a function of the time offset, and the time offset corresponding to the peak correlation coefficient is used to predict when the signals contained in the two spectrograms are aligned. We used the spectrographic cross-correlation feature in Raven Pro Interactive Sound Analysis Software (version 1.4) to measure time-of-arrival-differences from our multichannel recordings. Specifically, we measured the time required for the playback stimulus to reach each microphone in the array, relative to when it reached the closest microphone in the array (see Fig. 1b-d). Spectrograms were generated using a 512-point FFT, 87.5% overlap, and a Hamming window, which resulted in a temporal resolution of 2.9 ms and a frequency resolution of 43 Hz. Audio files were filtered with a bandpass filter to remove background noise outside of the range of our target sounds (songs of the male rufous-and-white wren: 500–2800 Hz; songs of the female rufous-and-white wren: 600–3800 Hz; songs of the male song sparrow: 1500–8000 Hz; calls of the grey treefrog: 500–4000 Hz; calls of the spring peeper: 2200–3600 Hz) and were normalized to a peak amplitude of 0 dB within each audio channel. Correlation functions were then computed automatically by Raven. Importantly, we manually inspected each correlation function to ensure that the peak correlation value and the associated latency value were based on the signal of interest and not on an artefact contained in the audio recording. Such a situation was obvious because the target signal was misaligned between the two corresponding spectrograms. If the peak correlation value and the associated latency value were based on an artefact, then they were re-calculated following manual alignment of the two spectrograms. Similarly, if the playback stimulus was not visible on the spectrogram of a particular audio channel, then latencies associated with that channel were excluded from further analysis. The remaining latencies (see Fig. 1d) were input as the time-of-arrival-differences into Sound Finder, along with the temperature at the time of recording, and GPS coordinates of each microphone.

Although we used spectrographic cross-correlation in our analysis, we note that it is also possible to conduct cross-correlation on a signal's waveform (Zollinger et al. 2012). Since spectrograms have imperfect temporal resolution, waveform cross-correlation can potentially calculate time-of-arrival-differences with better accuracy. We used spectrographic cross-correlation in our analysis because the signal-to-noise-ratios of our target sounds within the array recordings were too low to detect the signals from the waveforms, even after filtering and normalizing the recordings.

2.3 Localizing sounds

For each sound, we defined Sound Finder's localization accuracy as the distance between the location estimate provided by Sound Finder and the GPS location determined for the loudspeaker.

We compare the location estimates from Sound Finder not only to the GPS measurements of the positions of the loudspeakers broadcasting the stimuli, but also to the location estimates generated from a previous software approach for microphone array analysis: ArrayGUI. This software is written in MatLab (Mathworks Inc., Natick, MA, USA) and is one of the acoustic localization programs most commonly described in the literature (see Mennill et al. 2006, 2012). ArrayGUI automatically computes spectrographic cross-correlation functions for predefined sections of a sound, and then uses an optimization procedure to estimate the sound’s origin. Importantly, we used the same spectrogram parameters and filter settings in ArrayGUI as we did in Sound Finder. Following the methods outlined in previous studies involving ArrayGUI software (see Mennill et al. 2006, 2012 for details), we attempted to localize each playback stimulus three times by applying the cross-correlation procedure to three short (i.e., < 1.0 s) non-overlapping sections of each playback stimulus. We defined ArrayGUI's localization accuracy as the distance between the location estimate with the lowest error (a value generated by ArrayGUI that reflects the probability that the location estimate is correct) and the GPS location of the loudspeaker.

2.4 Statistical analysis

In our first analysis, we described the accuracy of Sound Finder by comparing its location estimates to the known locations of the loudspeakers. We then used a linear mixed-effects model to test whether the ‘error value’ produced by Sound Finder could be used to assess the accuracy of a location estimate when the actual location of the sound source is unknown. We entered ‘error value’ in milliseconds as a covariate with fixed effects, ‘localization accuracy’ as the dependent variable, and ‘array’ as a subject variable with random effects to account for non-independence between the two speaker locations in each array. To facilitate comparisons between Sound Finder and other localization techniques, we repeated this analysis using the error values and localization accuracies derived from ArrayGUI (error values were derived from the column labelled 'error' in the ArrayGUI output).

In our second analysis, we used a linear mixed-effects model to compare the localization accuracy of Sound Finder to that of ArrayGUI. We accounted for non-independence between the two speaker locations in each array, and between the two localizations conducted on each acoustic signal, by including ‘array’ and ‘speaker location’ nested within ‘array’ as subject variables with random effects. We included ‘analysis software’ as a factor with fixed effects (i.e., Sound Finder vs. ArrayGUI) and ‘localization accuracy’ as the dependent variable. We also included descriptive statistics to describe the probability of achieving different degrees of localization accuracy with each software approach.

For all linear mixed-effects models, we used the restricted maximum likelihood method to estimate the fixed effects, and we modelled the subject effect(s) by assuming a variance components covariance structure. Residuals were not normally distributed in preliminary models, but were corrected by applying a log(10)-transformation to 'localization accuracy' and 'error value'. All other assumptions were satisfied in the final models. Statistical models were conducted in PASW (version 18 for Mac; IBM, Armonk, NY, USA) and results were considered statistically significant when $P \leq 0.05$.

Results

Sound Finder required only a fraction of a second to localize 76 loudspeakers broadcasting five types of animal sounds. The sounds were broadcast in a variety of environments, including field, forest, temperate, and tropical environments, and were recorded with three different kinds of microphone array, including a wireless array and two different cable-based arrays. The average distance between the loudspeaker position and the location estimated by Sound Finder was 4.3 m (N = 76 sounds; 95% CI: 2.9 – 6.2 m). We consider this distance to be highly accurate given that the sounds were recorded with large, dispersed microphone arrays (average area bounded by each microphone array was 1.78 ha) that had relatively low microphone densities (average microphone density was 12.7 microphones/ha). We note that this level of accuracy is based on the two-dimensional microphone arrays used in our study, and that future studies will need to establish Sound Finder's accuracy for sounds derived from three-dimensional arrays.

Sound Finder provided information for assessing the accuracy of location estimates (Table 1), which would be critical in applications where the actual location of the sound source is unknown. In our investigation of sounds that were produced at known locations, the error value generated by Sound Finder significantly predicted localization accuracy, with higher error values corresponding to less accurate location estimates (linear mixed-effects model: $F_{1,72}$ = 83.5, N = 76 sounds, $P < 0.001$; Table 1). In contrast, ArrayGUI's error value did not significantly predict its localization accuracy ($F_{1,73}$ = 0.6, P = 0.460) for our 76 playback stimuli.

The 76 location estimates produced by Sound Finder were, on average, significantly more accurate than those produced by ArrayGUI (linear mixed-effects model: $F_{1,113}$ = 49.6, N = 76 sounds, $P < 0.001$; Fig. 2a). Furthermore, Sound Finder localized 24% of the sounds to within 1 m of their actual location (compared to 3% by ArrayGUI), 42% to within 3 m (compared to 8% by ArrayGUI), 57% to within 5 m (compared to 14% by ArrayGUI), and 74% to within 10 m (compared to 22% by ArrayGUI). Only 26% of the sounds localized by Sound Finder had a localization accuracy of 10 m or more, whereas 78% of the sounds localized by ArrayGUI had a localization accuracy of 10 m or more (Fig. 2b).

Discussion

We developed a new software approach, Sound Finder, for localizing sounds recorded with a microphone array. Sound Finder is unique among acoustic localization software programs because it operates in the free software environment R, and in spreadsheet programs such as Microsoft Excel and the open source software LibreOffice. We include the software here as a free online supplement to ensure that it will be disseminated universally in a standard form (instructions for use of the software are contained in the spreadsheet version of Sound Finder in the worksheet entitled "Instructions" and in the R version of Sound Finder through the built-in help functions).

We showed that Sound Finder produces accurate location estimates for a variety of animal sounds. On average, the distance between the loudspeaker and the location estimate produced by Sound Finder was 4.3 m, which we consider to be highly accurate for the large field sites and low microphone densities used in our study. This high level of accuracy was not the result of limited sampling, as it was based on all 76 playback stimuli, including those with relatively poor recording quality (as assessed visually from spectrograms during the cross-correlation procedure). It was also based on multiple types of animal sounds, including three avian and two anuran signal types, and on a variety of microphone array configurations, including arrays that did or did not rely on microphone cables, arrays that had four, eight, or 16 microphones, arrays that were located in tropical or temperate environments, and arrays that were located in forested or open habitats.

The location estimates produced by Sound Finder were accurate, but not perfect. There were at least three sources of measurement error that may have contributed to the localization error reported in this study. First, the Global Positioning System used to measure microphone and loudspeaker positions had an average horizontal accuracy of 1.26 m for microphones and

1.80 m for loudspeakers, and thus probably contributed significantly to the 4.3 m of error associated with Sound Finder's location estimates. Although a Global Positioning System was deemed the best method for measuring microphone and speaker locations in the large and densely vegetated sites used in our study, it may not be the most accurate method in other situations. For example, instruments such as tape measures and compasses, or total station surveying equipment, may provide better accuracy in open sites, and may thus improve the quality of data that Sound Finder uses to localize animal sounds (see also Collier et al. 2010). Second, our analyses assumed that the microphones and loudspeakers were located on a horizontal plane. Although our study sites were generally flat, subtle, unmeasured variation in the altitude of the microphones and loudspeakers within an array could have contributed to overall localization error. Third, the spectrogram cross-correlation procedure used in our study had a temporal resolution of 2.9 ms. Since sound travels approximately 1 m in 2.9 ms, this error probably also contributed significantly to the 4.3 m of localization error. In future studies, it may be possible to reduce this error by replacing spectrogram cross-correlation with waveform cross-correlation, which has a superior temporal resolution that is limited only by the sampling rate of the recording. Waveform cross-correlation will be most feasible when the signal-to-noise ratio of the sounds being localized is high, which will tend to occur when the sounds being studied are loud, when background noise at the study site is low, and when microphone density is high. Alternatively, it may be possible to reduce error for some signals by improving the temporal resolution of spectrogram cross-correlation procedures.

Acoustic localization programs should provide researchers with an estimate of their localization accuracy. This is important because the true locations of the animals they are localizing are usually unknown. For Sound Finder, we showed that the error value generated by the program provides a reliable measurement of localization accuracy for a variety of animal sounds and microphone array configurations. Therefore, a researcher can use the error value from Sound Finder to estimate the accuracy of future localizations. Based on the sounds recorded in our study, for example, 75% of localizations with an error value between 1 and 2 ms had an accuracy of 3.2 m or less, and 75% of localizations with an error value less than 1 ms had an accuracy of 0.9 m or less (Table 1). For the best results, however, we recommend that researchers recalibrate the relationship between localization accuracy and the error value whenever they employ a new microphone array configuration, move to a new environment or habitat, or conduct research on a new type of animal signal. This is important because the factors that might affect the relationship between the error value and localization accuracy (e.g., recording conditions, accuracy of microphone positions, measurement error during cross-correlation, signal structure) are poorly understood. Recalibration is easily done by using a loudspeaker to broadcast a typical sound of the study animal, at a typical amplitude and from a typical position within the recording area, and then calculating the accuracy with which Sound Finder localizes the sound source.

Sound Finder generated location estimates that were, on average, seven times more accurate than the location estimates generated by one of the most commonly-used localization approaches, ArrayGUI (Fig. 2a). Compared to ArrayGUI, Sound Finder also generated accurate location estimates for a greater proportion of sounds (Fig. 2b). We suggest that these differences are not based on the mathematical algorithms used by each program to convert time-of-arrival-differences to location estimates, but, rather, that they are based exclusively on the accuracy of the time-of-arrival-differences themselves. For Sound Finder, time-of-arrival-differences were generated in separate software using spectrographic cross-correlation; critically, the correlation functions were inspected manually to ensure that their peak correlation was based on the signal of interest and not a non-target sound contained in the audio recording. In contrast, ArrayGUI does not permit the user to manually inspect correlation functions, so many of its peak correlations may have been based on non-target sounds, such as other animal vocalizations, background noise such as wind or traffic, or recording artefacts caused by reverberation or microphone interference. The difference in the accuracy of the two software programs shows that it is worthwhile to manually inspect cross-correlation functions, rather than rely on automated correlation procedures. We suggest that this is particularly important when the recordings have a low signal-to-noise ratio or when they contain frequent non-target sounds.

We note that the location estimates generated by ArrayGUI were less accurate in our study than in Mennill et al. (2006, 2012), even though our analyses relied on array recordings derived from those previous studies. This discrepancy does not affect the comparison of Sound Finder and ArrayGUI, but it does warrant explanation. In the Mennill et al. (2012) study, localization accuracy was based on a subset of localizations that were deemed ‘reliable’ (i.e., 60% of the playback stimuli that were initially localized; see Mennill et al. 2012 for details). Since reliability correlates with localization accuracy, the exclusion of ‘unreliable’ localizations would have improved localization accuracy in that study. In contrast, localization accuracy in our study was based on all of the playback stimuli. The greater inclusivity allowed us to test Sound Finder’s ability to localize sounds with low signal-to-noise ratios, but it also worsened the localization accuracy of ArrayGUI and Sound Finder because the faint signals were more challenging to cross-correlate.

Sound Finder provides a simple, accurate, available, and affordable software solution for localizing animal sounds recorded with a microphone array. As with previous software solutions, however, Sound Finder has certain limitations. First, Sound Finder does not generate time-of-arrival-differences, but, rather, relies on cross-correlation procedures contained in other software. This affords the user the flexibility to use preferred and dedicated bioacoustics software for cross-correlation analysis, but may also require the user to purchase that software if is not already available. Second, as with any acoustic localization software, localizing sounds can be time-consuming. The user must extract the target sounds from the array recordings, cross-correlate the signal in separate software, and then copy the time-of-arrival-differences into Sound Finder. A benefit of Sound Finder, however, is that it then localizes all of the sounds automatically as a batch process in only a fraction of a second.

In conclusion, Sound Finder is a new approach for acoustic localization that provides accurate estimates of a vocalizing animal’s location. It is easy to use, it is available in a standard form as supplemental material to this paper (see supplemental material, “S1 Sound Finder for R.zip” and "S2 Sound Finder for Spreadsheets.xls"), and it requires only readily available software. Sound Finder therefore provides an additional software solution for localizing animal sounds recorded with a microphone array, and should provide additional opportunities for researchers to use acoustic localization in future studies of animal ecology and behaviour.

Acknowledgments

We thank Pierre-Paul Bitton for assistance with programming Sound Finder. We thank Beth MacDougall-Shackleton and Janet Lapierre for their collaboration in collecting sparrow array recordings, and the many assistants who helped in collecting array recordings of sparrows, wrens, and frogs. We thank several anonymous reviewers for providing helpful feedback on earlier versions of our paper and Sound Finder. We thank the Natural Sciences and Engineering Research Council of Canada (NSERC) for providing equipment grants for microphone arrays and the survey-grade GPS, and for providing support through the Undergraduate Summer Research Award program to MB, the Post-Doctoral Fellowship program to DRW, and the Discovery Grants program to DJM. The Canada Foundation for Innovation and the Government of Ontario also supported this research through grants to DJM.

References

Bancroft S. 1985. An algebraic solution of the GPS equations. IEEE Transactions on Aerospace and Electronic Systems 21:56–59.

Blumstein DT, Mennill DJ, Clemins P, Girod L, Yao K, Patricelli G, Deppe JL, Krakauer AH, Clark C, Cortopassi KA, Hanser SF, McCowan B, Ali AM, Kirschel ANG. 2011. Acoustic monitoring in terrestrial environments using microphone arrays: applications, technological considerations and prospectus. Journal of Applied Ecology 48:758–767.

Collier TC, Kirschel AN, Taylor CE. 2010. Acoustic localization of antbirds in a Mexican rainforest using a wireless sensor network. Journal of the Acoustical Society of America 128:182–189.

Double M, Cockburn A. 2000. Pre-dawn infidelity: females control extra-pair mating in superb fairy-wrens. Proceedings of the Royal Society of London Series B 267:465–470.

Lapierre JM, Mennill DJ, MacDougall-Shackleton EA. 2011. Spatial and age-related variation in use of locally common song elements in dawn singing of song sparrows *Melospiza melodia*; old males sing the hits. Behavioral Ecology and Sociobiology 65:2149–2160.

Magyar I, Schleidt WM, Miller B. 1978. Localization of sound producing animals using the arrival time differences of their signals at an array of microphones. Experientia 34:676–677.

Makino TT, Sakai S. 2004. Findings on spatial foraging patterns of bumblebees (*Bombus ignitus*) from a bee-tracking experiment in a net cage. Behavioral Ecology and Sociobiology 56:155–163.

M^{c}Donald PG, Kazem AJN, Wright J. 2007. A critical analysis of 'false feeding' behavior in a cooperatively breeding bird: disturbance effects, satiated nestlings or deception? Behavioral Ecology and Sociobiology 61:1623–1635.

Mech LD, Barber SM. 2002. A critique of wildlife radio-tracking and its use in national parks: a report to the U.S. National Park Service. Biological Resource Management Division, U.S. National Park Service, Fort Collins, USA.

Mennill DJ. 2011. Individual distinctiveness in avian vocalizations and the spatial monitoring of behaviour. Ibis 153:235–238.

Mennill DJ, Burt JM, Fristrup KM, Vehrencamp SL. 2006. Accuracy of an acoustic location system for monitoring the position of duetting tropical songbirds. Journal of the Acoustical Society of America 119:2832–2839.

Mennill DJ, Vehrencamp SL. 2008. Context-dependent functions of avian duets revealed by microphone array recordings and multispeaker playback. Current Biology 18:1314–1319.

Mennill DJ, Battiston M, Wilson DR, Foote JR, Doucet SM. 2012. Field test of an affordable, portable, wireless microphone array for spatial monitoring of animal ecology and behaviour. Methods in Ecology and Evolution 3:704–712.

Muanke PB, Niezrecki C. 2007. Manatee position estimation by passive acoustic localization. Journal of the Acoustical Society of America 121:2049–2059.

Rutz C, Burns ZT, James R, Ismar SMH, Burt J, Otis B, Bowen J, St Clair JJH. 2012. Automated mapping of social networks in wild birds. Current Biology 22:R669–R671.

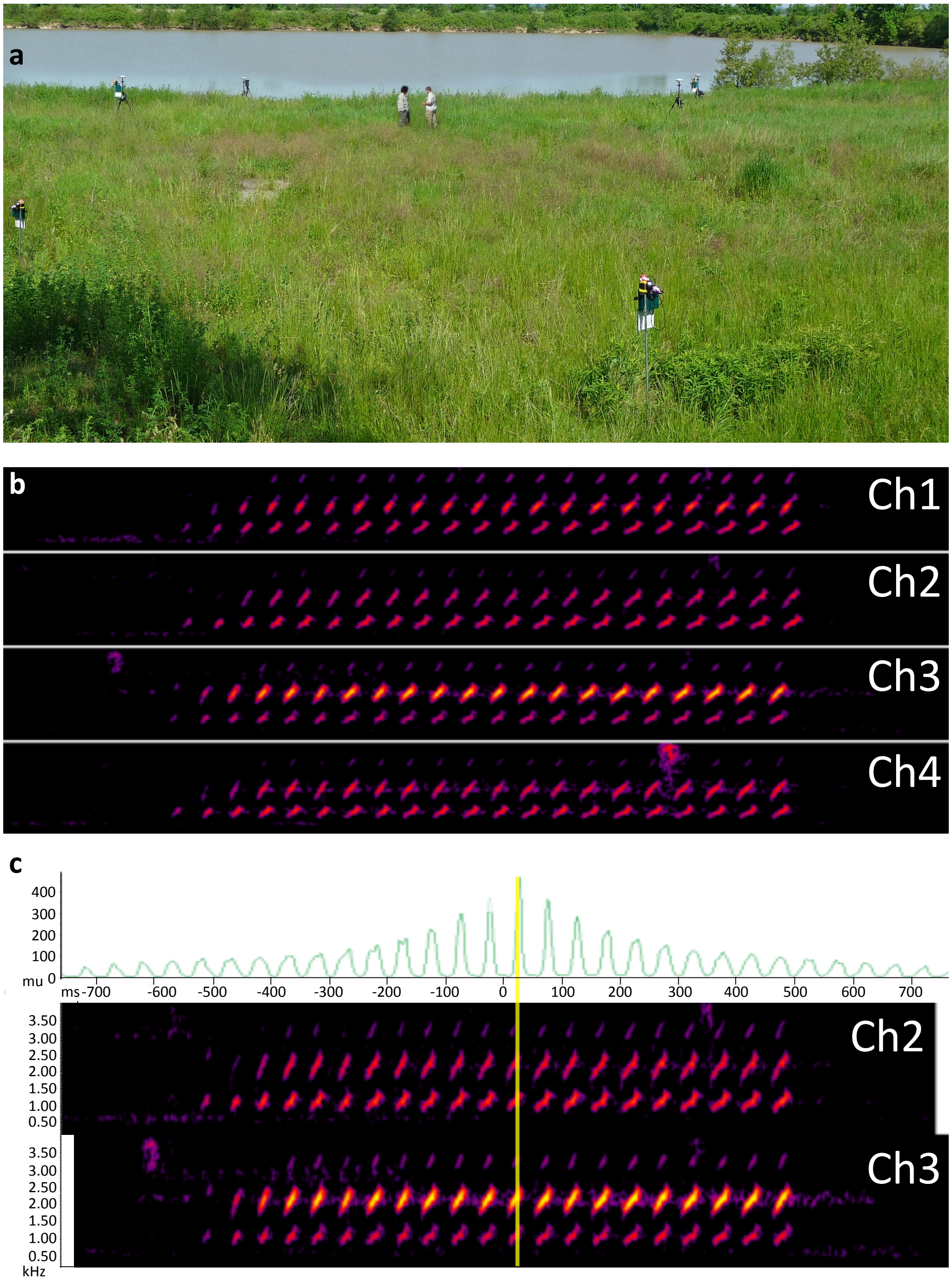
Spiesberger JL, Fristrup KM. 1990. Passive localization of calling animals and sensing of their acoustic environment using acoustic tomography. American Naturalist 135:107–153.

Wölfel M, McDonough J. 2009. Distant speech recognition. West Sussex: John Wiley & Sons Ltd.

Zollinger SA, Podos J, Nemeth E, Goller F, Brumm H. 2012. On the relationship between, and measurement of, amplitude and frequency in birdsong. Animal Behaviour 84:E1–E9.

Table 1. Sound Finder produces an error value in milliseconds that can be used to assess the accuracy of its location. Shown for each error value are five common percentiles of localization accuracy, including the 50th, 75th, 90th, 95th, and 100th percentiles. Localization accuracy is the distance between the location of the sound's origin, as estimated by Sound Finder, and the actual location of the loudspeaker, as determined by a Global Positioning System. *N* is the number of sounds localized. As an example of how to interpret this table, 50% of the localizations with an error value between 0 and 1 ms have a localization accuracy of 0.3 m or less, whereas 50% of the localizations with an error value between 1 and 2 ms have a localization accuracy of 1.3 m or less.

		Localization accuracy (m)				
Error value (ms)	*N*	50%	75%	90%	95%	100%
0-1	10	0.3	0.9	6.9	17.9	28.8
1-2	12	1.3	3.2	48.3	58.6	65.3
2-3	13	2.4	3.4	4.7	5.0	5.0
3-5	10	4.0	6.1	64.5	66.5	68.4
5-10	15	8.0	18.1	69.1	70.4	71.6
10+	16	19.9	47.5	64.7	89.4	119.9



d

	Channel 1	Channel 2	Channel 3	Channel 4
Channel 1	0.00000	-0.00290	0.01887	0.01596
Channel 2	0.00290	0.00000	**0.02177**	0.01887
Channel 3	-0.01887	-0.02177	0.00000	-0.00290
Channel 4	-0.01596	-0.01887	0.00290	0.00000

Figure 1. Summary of the acoustic localization technique used to locate loudspeakers broadcasting five types of animal signal. (a) Arrays of microphones were set up in multiple habitats, including this open field habitat in southern Ontario. (b) Sounds such as this single trilled grey treefrog advertisement call were broadcast through loudspeakers and re-recorded as mulitchannel audio files. (c) Spectrographic cross-correlation was used to measure time-of-arrival-differences between each pair of microphones. (d) Time-of-arrival-differences between each microphone in the array and the first microphone reached by the sound (highlighted in yellow) were input into Sound Finder to estimate the position of each loudspeaker. Axes and axis labels have been digitally redrawn to improve clarity.

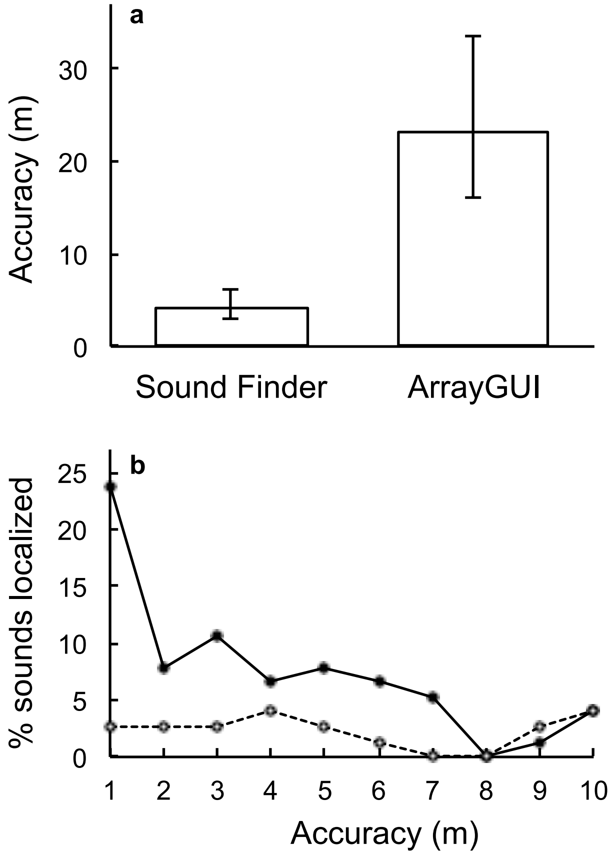


Figure 2. Accuracy of two software programs used for localizing 76 animal sounds. Accuracy is defined as the distance between the location of the sound's origin, as estimated by the software, and the actual location of the loudspeaker, as determined by a GPS. (a) Overall accuracy of the two software programs. Means and the 95% confidence interval are reverse log10-transformed from the estimated marginal means generated by our linear mixed-effects model. (b) Probability that each program accurately localizes sounds. Shown as a function of localization accuracy is the percentage of 76 sounds localized by Sound Finder (solid circles, solid line) and by ArrayGUI (open circles, hatched line). Values on the x-axis represent a range, where, for example, "1" represents 0–1, and "2" represents 1–2.

Supplemental information

S1 Sound Finder for R.zip Acoustic localization program. Contained in this single file are separate files that contain the R package for Windows operating systems (SoundFinder_1.0.zip), the R package for UNIX and Macintosh operating systems ("SoundFinder_1.0.tar.gz "), microphone coordinates for a hypothetical three-dimensional 18-microphone array ("example mics.csv"), and time-of-arrival-differences for six hypothetical sounds recorded with that array ("example sounds.csv").

S2 Sound Finder for Spreadsheets.xls Acoustic localization program. Contained in this single Microsoft Excel document are detailed instructions for using the spreadsheet version of Sound Finder (worksheet "Instructions"), example data corresponding to six sounds derived from a hypothetical three-dimensional 18-microphone array (worksheets "mic" and "sound"), spreadsheet calculations for localizing a particular sound (worksheets "prep", "find2D", and "find3D"), Visual Basic code for applying the localization process to multiple sounds (macro "Sound_Finder"), and a worksheet for storing the results of multiple localizations (worksheet "results (batch)").

S3 Sound Clips.zip Sound clips used to test Sound Finder. Contained in this single file are the 76 multichannel sound files that were used to test Sound Finder. Filenames indicate the corresponding array (1-38), whether the sound was the first or second sound broadcast within the array (pb 1 or pb 2), and the species identity ("RWWR" is rufous-and-white wren; "SPPE" is spring peeper; "GRTR" is grey treefrog; "SOSP" is song sparrow). Also included is an Excel file containing the corresponding metadata ("S3 sound clip metadata"); the worksheet entitled "sound clips" provides the filenames of each sound clip, the temperature at the time of recording (°C), and the GPS coordinates of the loudspeaker used to broadcast the sound (UTM coordinates); the worksheet entitled "mic coordinates" provides the GPS coordinates of each microphone in each array (UTM coordinates).