American Journal of Primatology 77:767–776 (2015)

RESEARCH ARTICLE

Validation of an Acoustic Location System to Monitor Bornean Orangutan (*Pongo pygmaeus wurmbii*) Long Calls

BRIGITTE SPILLMANN[1], MARIA A. VAN NOORDWIJK[1], ERIK WILLEMS[1], TATANG MITRA SETIA[2], URS WIPFLI[1], AND CAREL P. VAN SCHAIK[1]
[1]*Anthropological Institute and Museum, University of Zurich, Zurich, Switzerland*
[2]*Fakultas Biologi, Universitas Nasional, Jakarta, Indonesia*

The long call is an important vocal communication signal in the widely dispersed, semi-solitary orangutan. Long calls affect individuals' ranging behavior and mediate social relationships and regulate encounters between dispersed individuals in a dense rainforest. The aim of this study was to test the utility of an Acoustic Location System (ALS) for recording and triangulating the loud calls of free-living primates. We developed and validated a data extraction protocol for an ALS used to record wild orangutan males' long calls at the Tuanan field site (Central Kalimantan). We installed an ALS in a grid of 300 ha, containing 20 SM2+ recorders placed in a regular lattice at 500 m intervals, to monitor the distribution of calling males in the area. The validated system had the following main features: (i) a user-trained software algorithm (Song Scope) that reliably recognized orangutan long calls from sound files at distances up to 700 m from the nearest recorder, resulting in a total area of approximately 900 ha that could be monitored continuously; (ii) acoustic location of calling males up to 200 m outside the microphone grid, which meant that within an area of approximately 450 ha, call locations could be calculated through triangulation. The mean accuracy was 58 m, an error that is modest relative to orangutan mobility and average inter-individual distances. We conclude that an ALS is a highly effective method for detecting long-distance calls of wild primates and triangulating their position. In combination with conventional individual focal follow data, an ALS can greatly improve our knowledge of orangutans' social organization, and is readily adaptable for studying other highly vocal animals. Am. J. Primatol. 77:767–776, 2015.

Key words: microphone array; triangulation; passive recording and monitoring; loud call

INTRODUCTION

Remotely recorded animal vocalizations are useful in investigations of various research topics, including biodiversity assessment, species abundance, seasonality in communication, the effect of ambient noise, or anthropogenic sounds, adaptations in signal directionality, or social networks [Blumstein et al., 2011]. Additionally, some research questions require knowledge of the location of vocalizing animals. These requirements, namely identifying a call's occurrence both in time and in space, can be met by an Acoustic Location System (ALS) [McGregor et al., 1997]. Although an ALS has many benefits [Blumstein et al., 2011; Mennill et al., 2012], its main limitation is that a signal's propagation capacity greatly determines inter-microphone distances. Because primate loud calls are well adapted for long-distance propagation, they are promising potential candidates for an ALS.

Acoustic tracking of vocalizing animals was originally developed for marine animals, especially cetaceans [Watkins and Schevill, 1972]. More recently, studies of birds have increasingly come to use acoustic location to track individuals through their singing or calling to address different aspects of communication networks [Fitzsimmons et al., 2008; Foote et al., 2008; Lippold et al., 2008; Mennill and Vehrencamp, 2008; Peake et al., 2001]. So far, acoustic tracking studies of terrestrial mammals are still scarce [Collier et al., 2010; Thompson et al., 2010; Wrege et al., 2012], and in particular there is no study of primates that combines acoustic location and behavioral data collection via observation. There is growing interest, however, in applying this

Correspondence to:Anthropological Institute and Museum, University of Zurich, Zurich, Switzerland. E-mail: spillbri@gmx.ch

Received 13 October 2014; revised 19 February 2015; revision accepted 20 February 2015

DOI: 10.1002/ajp.22398
Published online 16 March 2015 in Wiley Online Library (wileyonlinelibrary.com).

technique to primates, especially to wild-living apes (chimpanzees and orangutans) where radio- or satellite-tracking are not possible [Piel, 2014].

For wild orangutans, direct focal observations provide only a very limited overview of their social system, especially on Borneo, where individuals—except for mother–offspring pairs—are usually solitary, and widely dispersed. Direct encounters, especially between flanged males, are rare [Galdikas, 1985a; MacKinnon, 1979; Mitra et al., 2009; Rijksen, 1978; van Schaik, 1999]. However, a conspicuous loud vocalization, the so-called "long call", given only by flanged male orangutans, often affects other individuals' (both males and females) ranging behavior, and mediates social relationships [Delgado, 2003; Galdikas, 1983; Mitani, 1985b; Spillmann et al., 2010]. Long call production and duration are highly variable: flanged males tend to give long calls several times each day and long calls last from 15 sec to 4 min [Delgado et al., 2009]. Additionally, van Schaik et al. [2013] showed that spontaneously given long calls announce future travel direction in Sumatran orangutans. The most effective observational approach—undertaking simultaneous individual focal follows— at best captures only a subset of the individuals and their interactions in a study area, and is logistically difficult. An ALS for long calls would provide a valuable complement to direct observations.

We studied Bornean orangutans in Tuanan, Central Kalimantan, a dense peat swamp forest with a dense orangutan population (4.25–4.5 individuals/km^2) [van Schaik et al., 2005]. We established an ALS to continuously record long call events in the study area, using 20 time-synchronized recorders placed in a lattice at 500 m intervals. The recorders were continuously active for 14 hr, from dawn to dusk. These time-synchronized recordings permit localizing long call occurrences through triangulation, via differences in the time of arrival of a signal at different recorders [McGregor et al., 1997; Mennill et al., 2006; Wilson et al., 2013]. Moreover, because long calls are individually distinctive in their acoustic structure [Delgado, 2003; Lameira and Wich, 2008; Spillmann et al., 2010], we could identify the individuals that had emitted the recorded long calls, even if they were not directly observed. The advantage of an ALS is the standardized monitoring and detection of multiple individuals' calling behavior simultaneously. In combination with conventional focal follows, an ALS should lead to a better understanding of the flanged males' communication in time, and space, as well as of the ranging responses of intended long call receivers and eavesdroppers.

The accuracy of triangulation depends on the density of microphones in an array, and is also affected by attenuation, scattering, and reverberation, more so in forest habitats than in open fields [McGregor et al., 1997; Mennill et al., 2012]. We strove to reach a biologically meaningful compromise between covering a relative large area to get a broad picture of flanged male abundance, ranging behavior, and identity on the one hand, and maximizing the accuracy in localizing long calling males' positions and their corresponding ranging behavior on the other.

The aim of this study was to validate this orangutan ALS, in two main steps (Table I). In the first step we asked whether all long calls were picked up in the area of the ALS. We examined 25 recording days (around 350 recording hr) to verify whether a trained, automated recognition algorithm (Song Scope) reliably recognized independently verified long call occurrences. To estimate the actual area covered by the microphone-grid, we examined the distances at which long calls are still reliably detected by a trained recognition algorithm.

In a second step, we checked whether the ALS correctly determined the locality of long calls. We established a validation procedure for the acoustic location measurements of 89 long calls given by seven flanged males who were followed by observers and whose GPS locations during long call events were therefore available for comparison. Spectrographic cross-correlation of signals from a pair of recorders aligns two spectrograms where the peak correlation value corresponds to the time of arrival difference of a signal at two recorders. With at least three recorders recording a long call, we could then use these pair-wise time-of-arrival differences of a signal to triangulate the location of a long calling

TABLE I. Summary of the Steps of the Validation Procedure

Step	Analysis	Validation
1	Locate long calls in sound files with a trained recognition algorithm. (Song Scope)	Does recognition algorithm locate all known long calls?
2a	Extraction of differences in time of arrival through cross-correlations. (Raven Pro 1.4)	
2b	Triangulation of long call position with different time of arrivals (Sound Finder)	Does Sound Finder provide correct locations for known LCs? Does the output error (ms) of Sound Finder correspond with distance to true location? Comparison of localization accuracy for LCs from inside versus outside the grid.

male. Additionally, we examined the effect of distance from the microphone array, because several studies have reported that acoustic location accuracy degrades with increasing distance from the microphone array [McGregor et al., 1997; Mennill et al., 2012]. The aim of this validation was to identify the area where an animal's position could be triangulated, and the area where the animals could be recorded but not triangulated. Additionally we wanted to explore the accuracy of this acoustic location method.

METHODS

Study Area and Recording

Fieldwork took place at the Tuanan field station, Central Kalimantan, Indonesia (2.151° South; 114.374° East) from March 2012 until February 2013. The field site is part of the Mawas area, managed by a local non-governmental organization (Mawas) that aims to protect the rainforest from illegal logging, fires, and poaching. The field site encompasses a 1,000 ha trail-system in a peat swamp forest that was previously selectively logged. The area is flat and forest structure is homogeneous (unpubl. data). Researchers and well trained long-term field assistants conducted individual focal follows according to standardized field methods (Available online at http://www.aim.uzh.ch/Research/orangutannetwork/FieldGuidelines. html). Whenever possible, we followed subjects from morning to night nest over a span of several days (6–10 days). Subjects included all age–sex classes. Using hand-held GPS devices (Garmin GPS 78), observers recorded GPS points every 30 min to indicate the position of the focal individual. Observers recorded additional GPS points whenever they heard a long call and, if the focal subject was a flanged male, whenever he emitted a long call. This research project adhered to the American Society of Primatologists (ASP) principles for the ethical treatment of non-human primates.

We placed twenty off-line, time-synchronized (via GPS), SM2+ (SM2 firmware 3.1.9) autonomous recording units in a lattice at 500 m intervals. The ALS thus encompassed a grid of 300 ha. Each recorder was equipped with a single omni-directional, weatherproof SMX-II microphone. A 12 V, 18 Ah dry gel battery, charged by a 20 Watt solar-panel on site, powered each SM2+ recorder, allowing it to run independently for about three weeks (H. Kühl, & O. Wagner, pers. comm.). In camp preparation of the power-systems (solar-panel, solar charger, DCDC-converter, battery) and microphone wires for each recorder required approximately fourteen days, and installation of the recording grid in the forest approximately seven days.

We settled on an average microphone density of 0.066 microphones/ha, which is extremely low compared with previous studies in which average microphone densities were about 10 microphones/ha [marmots: Collier et al., 2010; passerines: Wilson et al., 2013]. Our microphone grid of 300 ha was correspondingly large. We chose this low microphone density and large grid size for three reasons. First, long calls are audible up to ca. 1 km for human observers on the ground and are well adapted to long-distance propagation because of their acoustic structure, relatively low fundamental frequency, high sound pressure level (100 dB at 1 m), and high repetition of call elements [Mitani, 1985b; Waser and Brown, 1986]. Second, acoustic analysis can identify individual long-calling flanged males up to distances of 300 m from the microphone [Lameira and Wich, 2008]. Third, a pilot study of re-recorded long call playbacks with distances up to 800–1000 m from the microphone showed that long calls are still identifiable as such in a spectrogram (unpublished results).

We placed microphones in trees at a height of ca. 10 m, above the dense understory. A correspondence between the height of the microphone and the signaler should considerably improve receiving conditions [Dabelsteen, 2005; Maciej et al., 2011; Mathevon et al., 2005]. For each recorder, the microphone hung down from a branch close to the tree's trunk and above the recorder's position. We used the mean of the GPS coordinates provided every 10 min by each SM2+ recorder (SD = 7 m) to define microphone positions. We set the sample rate at 22,050 Hz and sample size to 16-bit signed PCM. Continuous recording occurred from 04:30 until 18:30 (starting 1 hr before sunrise and ending 1 hr after sunset). Each recording file covered exactly 1 h 57 mins, with a new file starting after a delay of 3 min. We staggered starting times of different recorders to ensure that at all times at least 15 of the 20 recorders were active. We saved the recorded files in a compressed .wac format without data loss and uploaded them every month from SD cards to an external hard drive.

Long Call Extraction

To search for long calls in the numerous .wac files, we established a trained recognition algorithm using Song Scope software with batch processing. Song Scope classification algorithms are based on Hidden Markov Models (HMMs) [Agranat, 2009]. We established a long call recognition algorithm in three steps. First, we selected high-quality long call recordings of all flanged males that observers had recorded with a shotgun microphone (Sennheiser ME 67) during individual focal follows. We used this selection as a first training set to produce a "high quality recognition model". In a second step we incorporated long calls recorded by SM2+ of moderate quality (greater distance between long call position and microphone) of different individuals into this model (see Buxton and Jones [2012]).

In the final step, we established optimal model parameters, including minimum frequency (200 Hz), frequency range (700 Hz), and sample rate (4000 Hz), based on the properties of long call acoustics. Additionally, we set Fast-Fourier Transform window size (512), FFT overlap (1/2), dynamic range ([20 dB] i.e., how much signal energy is used to compare call components; dB relative to peak call signal dB), maximum syllable duration (1,500 ms), maximum syllable interval (800 ms), and maximum song duration (60,000 ms).

Validation of Extraction Performance

We checked long calls identified by the trained recognition algorithm of Song Scope to remove false positive results (background noise or non-target calls). The sonogram allowed easy detection of false positive results which we subsequently discard from the spreadsheet [Buxton and Jones, 2012]. Nevertheless, we calculated the false positive rate of the trained recognition algorithm. However, false negatives give much more insight into the accuracy of the trained recognition algorithm. To evaluate the success rate of Song Scope, we compared the success of Song Scope with known occurrences of "long calls" in the area, based on individual focal follows. We analyzed 25 recording days (350 recording hr of the ALS), during which we recorded 145 long calls with known time and location that were given by flanged male focal subjects. All follows took place in the recording area of our ALS. In total, the ALS recorded 981 long calls during these 25 days. Therefore, during individual focal follows of flanged males, human observers recorded 15% of long calls given in the entire study area within this time period.

Area Covered by Microphone Array

To assess the maximum distances at which recorders recorded long calls and the trained algorithm exclusively recognized them, we examined the probability of detection as a function of distance to the long call position from each of all 20 recorders. We excluded distances to non-active recorders. For this analysis we used 89 long calls given by seven focal flanged males. This procedure gave us an estimate of the area covered by our ALS outside the recorder grid (Fig. 1 of supporting information).

Time of Arrival Differences

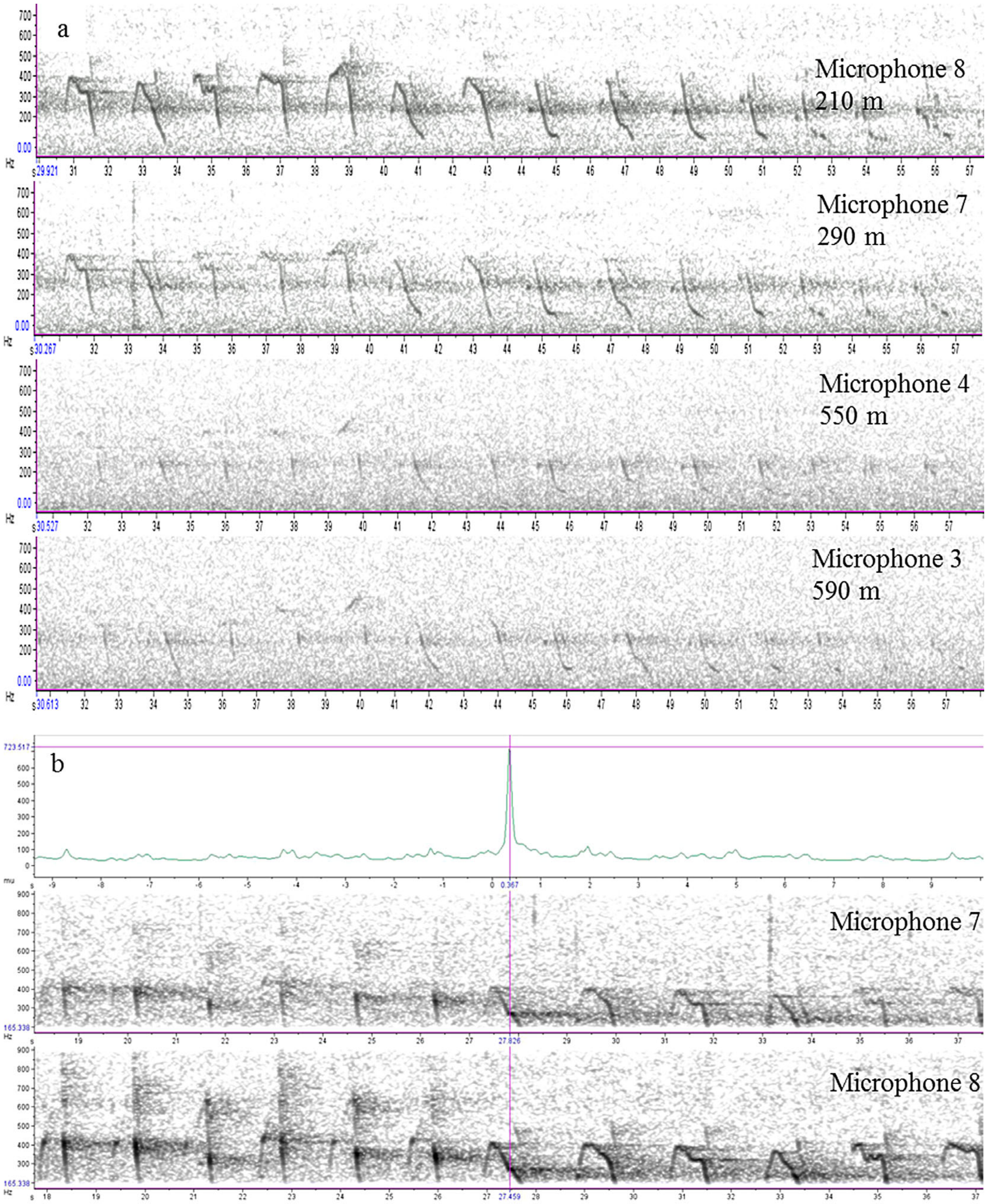
To initiate the second step (triangulation of long call locality), we first converted compressed .wac files into .wav files with WAC2WAV Converter Utility 3.3.0 (Wildlife Acoustics), because Raven Pro 1.4 does not recognize .wac files. We subsequently downsampled the .wav files from 22,048 to 6,000 Hz using WaveLab6 to reduce file size. This should not affect the results because a long call's fundamental frequency does not exceed 3000 Hz and the Niquist-Shannon sampling theorem states that a sampling rate is sufficient when the bandwidth of interest is half the sampling rate [Bradbury and Vehrencamp, 1998].

We established spectrographic cross-correlations to the time of arrival differences whenever at least three recorders recorded a long call. Spectrographic cross-correlation makes it possible to compare two spectrograms and align them according to the peak correlation coefficient (Fig. 1). The peak correlation coefficient corresponds to the time offset between the two spectrograms. We used a band-pass filter from 200 to 1000 Hz to remove background noise outside of the signal of interest. We set the window size to 256 samples with an overlap of 50% and window-type Hanning to generate fast fourier transformation (FFT), and normalized each audio file to 0 dB. Even though Raven Pro 1.4 generates correlation functions automatically, we manually inspected all of them, because the quality of the signal of interest was sometimes so low that Raven incorrectly assigned the highest cross-correlation. These differences in time of arrival of a long call recorded at different recorders were the basis of triangulating its position.

Localization of Long Call Source

We used Sound Finder, a free software program developed by Wilson et al. [2013]; to calculate the origin of a sound. It estimates the location of the sound source by applying the least-squares solution that was developed for global positioning systems. It also takes into account the temperature at the time of recording to calculate the corresponding speed of sound. Although we used the mean hourly temperature in all subsequent analysis provided here, a comparison with results based on mean day temperature revealed a highly significant correlation between the two ($r = 0.99$, $t = 59.43$, $P < 0.001$). Moreover the difference in accuracy between these two approaches was not significant (Paired t-test: $t = -1.33$, $P = 0.188$, $df = 72$) indicating that future studies at Tuanan can also use mean day temperature, which is far easier to calculate. Note that this might be site-specific because of the small temperature fluctuation over the course of the day in a tropical forest habitat at low altitude.

We used Sound Finder to localize long call sources in two-dimensional space. The output of Sound Finder includes latitude and longitude of the long call's origin. Additionally, Sound Finder estimates the time it takes for a sound to travel to the closest recorder (measured in sec) and an error associated with the localization (measured in milliseconds) where higher errors denote lower confidence in the accuracy of the localization [Wilson et al., 2013].



c	Microphone 8	Microphone 7	Microphone 4	Microphone 3
Microphone 8	0	-0.363	-0.832	-1.109
Microphone 7	0.363	0	-0.491	-0.768
Microphone 4	0.832	0.491	0	-0.277
Microphone 3	1.109	0.768	0.277	0

Fig. 1. (a) Sonogram of a sequence of a long call recorded by four different microphones, (b) crosscorrelation of two long calls with the peak correlation value that corresponds to the time of arrival difference, (c) results of peak correlation value between all recording files (see text).

Validation of Acoustic Location

One factor that affects localization accuracy is the position of the sound source in relation to the microphone grid. The result of a triangulation is more reliable when the sound originates from within the area bounded by the microphones rather than from outside it [McGregor et al., 1997; Mennill et al., 2012]. We therefore first concentrated on acoustic location results obtained for long calls originating within the microphone grid. We checked whether the error value produced by Sound Finder could be used to assess the accuracy of an acoustic location result. The accuracy of location was the distance in meters between the acoustic localization established by Sound Finder through triangulation and the corresponding GPS-position obtained by an observer who followed the calling male with a handheld GPS unit. It is important to note that coordinates measured with a handheld GPS unit have error margins of approximately 8–12 m at Tuanan.

Using a Wilcoxon rank-sum test we then compared results obtained from triangulations within the area bounded by microphones versus results obtained from outside the microphone array. Acceptable Sound Finder errors never produced a localization error of >100 m. We therefore set the cut-off distance from the microphone grid where triangulations became more inaccurate than those obtained inside the grid, at 200 m to stay within this acceptable localization error (Fig. 2).

This procedure allowed us to recognize three recording areas: in = inside the microphone grid, edge = up to 200 m outside the microphone grid, and out = >200 m outside the microphone grid. We conducted robustness tests (bootstrap with 9999 resamples) to assess whether these three areas differed according to the accuracy of triangulation results. All statistical tests were performed in R version 3.0.1 [R Core Team, 2013] and were two-tailed.

RESULTS

Success of Recognition Model

The success rate of the trained recognition algorithm using Song Scope was 99%: the recognition algorithm found 143 of 145 long calls observed during follows of flanged males. The two long calls that the trained recognition algorithm failed to recognize belonged to a flanged male with an exceptionally high-pitched voice (as had been obvious to human observers when naming him "Helium") and in addition background noise masked one of these two long calls. The system did, however, recognize another 21 long calls (> 90%) by this same male. Thus, the algorithm recognized all "normal" long calls known to have been emitted from within the study grid. Every fourth annotation of the trained recognition algorithm noted a long call. Therefore, on average, 75% of the annotations were false positives. Despite these errors, the advantage of using an automated method over a manual method to recognize long calls is apparent, given that every month included an average of 21 recording days, resulting in approximately 5880 recording hours (20 recorders × 14 hr × 21 days).

Coverage Area of ALS

The trained recognition algorithm recognized 77% of 89 long calls given by seven different flanged

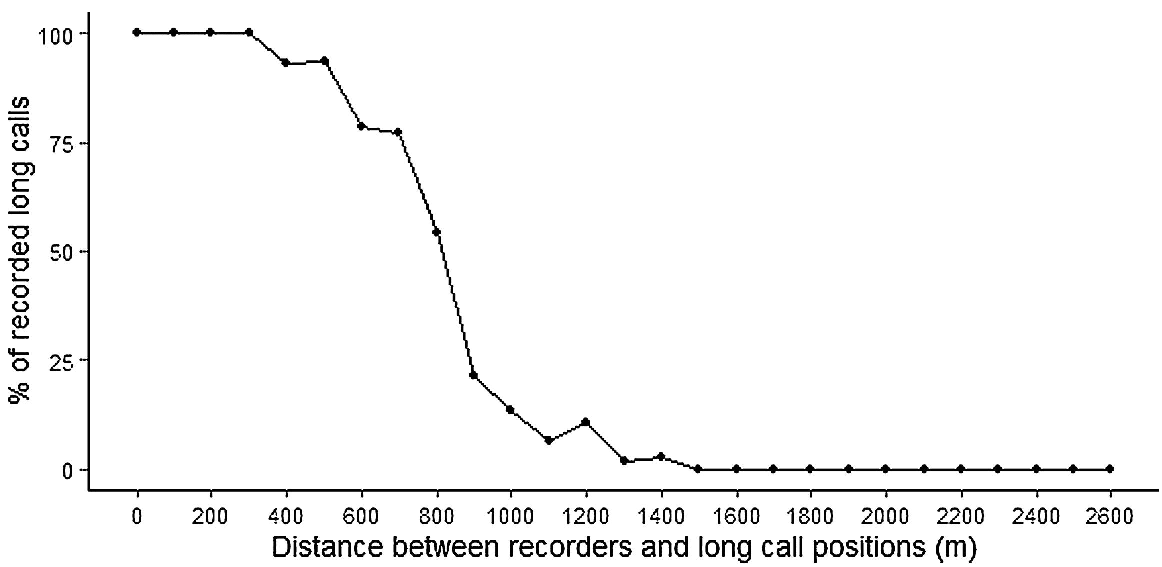


Fig. 2. The effect of distance on the recognition success by the trained recognition algorithm of long calls in the sound files.

males at distances up to 600–700 m from a recorder (Fig. 3), including one long call that occurred at an exceptional distance of 1400 m. When we excluded "Helium" from this analysis, 87% of all long calls were still recognized at a distance of 600–700 m. If two or more microphones were within this distance, the probability of the long call being detected by at least one was 98%. Therefore we can conclude that our ALS encompassed an area of approximately 700 m, outside of its own perimeter. This corresponds to a total area of around 900 ha covered by the microphone array (300 ha within the grid, and 600 ha outside).

Triangulation: Localization of Long Calls With Sound Finder

Triangulations of 66 long calls that occurred in the area bounded by the recorders showed errors of 8–94 m (mean: 58 ± SEM 7.2 m; median: 50 m), with the exception of a single call, which yielded a deviation of 496 m from its actual location. These values were reasonable, if one takes into account that the directly observed location was also subject to error because of variable animal–human distance and the error of the handheld GPS unit. Sound Finder also reports the error in arrival times in milliseconds (ms). A perfect triangulation would produce an error of 0 ms. Therefore, because all acceptable long call triangulations showed an error value by Sound Finder <200 ms (the error of the failed triangulation was 2,000 ms), we included triangulations of long calls in the area bounded by the microphones and for which Sound Finder reported error values of <200 ms in our subsequent analysis.

To delineate how far the triangulation area extended outside the microphone grid, we also compared the triangulation accuracy of long calls given inside the grid (median error = 50.50 m) with long calls emanating from outside the grid (median error = 102.0 m), and found far greater triangulation errors for long calls occurring outside the area bounded by microphones (Wilcoxon signed-rank test, $Z = -4.60$, $W = 267.5$, $P = <0.001$, $n = 89$).

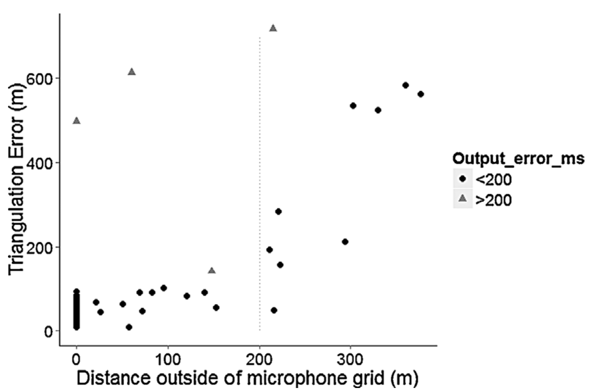


Fig. 3. Relationship between triangulation error (m) and the distance outside of the microphone grid where a long call occurs: Spearman rank correlation, $r_s = 0.596$, $P < 0.05$. Triangulations with low output errors (<200 ms) established by the software (dots) and triangulations with high output errors (triangles) are shown in the graph. All long calls from within the grid were given a distance of 0 m.

To compare error margins from within vs. outside the grid, we first established that triangulation error increased with distance from the grid (Spearman rank correlation, $r_s = 0.596$, $P < 0.05$). The error (in ms) established by Sound Finder predicted the accuracy of the triangulation (Fig. 2). Therefore triangulations up to 200 m outside the microphone grid seemed to reflect the accuracy of triangulations inside the grid, and we called this area the edge. However, triangulations from farther out were less accurate, even if they were accompanied by low errors (in ms) established by Sound Finder. To compare these three triangulation areas (in, edge, and out) we conducted a bootstrapping test to assess whether triangulation results of long calls originating from the edge of the microphone grid could be used for further analysis. For this analysis, we excluded triangulations that failed because of error >200 ms, as explained above. The results show that until approximately 200 m outside the microphone grid, the triangulation errors were the same as those from inside the grid, whereas errors rapidly increased when calls emanated from farther outside the grid (Table II). We therefore used the 200 m extension of the microphone grid to delineate the triangulation area.

DISCUSSION

We aimed to validate data extraction performance from an Acoustic Location System (ALS) set up to record the vocalizations of wild male Bornean orangutans. This validation showed (i) the reliable performance of the trained long-call recognition algorithm up to distances of 700 m, and (ii) the possibility to triangulate long call positions up to 200 m outside the microphone grid, with a mean triangulation error of 58 m.

Based on the results of this validation, we could divide the area that the ALS covered into (i) an area where acoustic location was possible up to 200 m outside the microphone grid, and (ii) a surrounding area, where only rough localization estimations were possible (see table III).

In this validation study, we did not attempt to identify the calling males in the acoustic landscape. However, Lameira and Wich [2008] showed that acoustic identification of long calls was possible up to a distance of 300 m from the sound source. Thus, individual identification should be possible within the triangulation area and might still be possible at

TABLE II. Comparison of the Accuracy of Long Call Triangulations Originating from Different Areas Relative to the Microphone Grid

Comparison of different areas	Mean difference (m)	Bootstrap (9999 resamples)			
				Confidence interval 95%	
		Std. Error	P-value	lower	upper
in[a] versus edge[b]	−16.5	8.64	0.084	−32.6	1.4
in[a] versus out[c]	−293.4	68.64	0.016	−425.4	−162.2
edge[b] versus out[c]	−276.9	69.10	0.020	−408.7	−146.0

[a]long calls that occurred inside the microphone grid
[b]long calls that occurred < 200 m outside the microphone grid
[c]long calls that occurred > 200 m outside the microphone grid

distances between 300 and 400 m (74% of the long calls fall within < 300 m, 91% fall within < 400 m of the closest microphone), given the much larger sample size of long calls that were recorded (via focal follows and the ALS), and consequently the far superior statistical power in the discriminant function analysis.

Factors That Account for Errors in the Acoustic Location Results

There are several explanations for the relatively high triangulation errors. First, sound propagation is affected by distance, leading to lower accuracy when distances between microphones are high. Second, localization accuracy is affected by habitat type: forest habitats lead to higher inaccuracy in acoustic location compared to open fields [Andersson and McGregor, 1999; Bradbury and Vehrencamp, 1998; McGregor et al., 1997; Mennill et al., 2012]. One way to improve accuracy is to increase the height of the receiver [Dabelsteen, 2005]. Therefore we installed microphones at a height of 10–15 m. Third, we defined triangulation error as the distance between the acoustic location of Sound Finder and the position of a corresponding long call recorded on a handheld GPS unit during an individual focal follow. At the Tuanan field site, handheld GPS devices show an error of approximately 8–12 m, and observers take GPS points at ground level at a (horizontal) distance of ca. 5–20 m from the focal subject. It is difficult to know the proportion of the triangulation errors caused by errors in the acoustic location established by the ALS vs. those that caused by GPS errors (error of GPS device and observer position relative to focal subject). Finally, rain and wind may affect the accuracy of results. Fortunately, long calls rarely occur during rain and especially not during hard rain. In the Tuanan population, observational data suggest that <5% of long calls occur during rain (unpubl. data). Strong wind conditions at Tuanan are rare and brief, usually occurring right before thunderstorms (pers. obs.). Moreover, closed habitats are generally less affected by atmospheric irregularities [Brown and Handford, 2000]. Therefore, wind probably had only a minor effect on our results.

Is the Acoustic Location System Accurate Enough?

Although the localization errors of ca. 50 m may appear large, and certainly are large compared to other studies [McGregor et al., 1997; Mennill et al., 2012; Wilson et al., 2013], they are actually modest in relation to orangutans' mean inter-individual distances and daily mobility. Orangutans social organization is referred to as semi-solitary: individuals are widely dispersed, except for mother–infant pairs. Simultaneous follows of different sex–age classes in Suaq Balimbing revealed average distances between adult females and dominant or other flanged males of 550–730 m, and between flanged males and

TABLE III. Coverage of ALS Divided into Two Areas

	Triangulation Area 450 ha	Recording Area 900 ha
All long calls recorded	yes	Prob. yes
Long call localization	yes	only rough estimate
Long call – ID[a]	yes	only presence of LCs
Use	ID + localization	No ID + rough localization, recognizing responses to LCs

unflanged males of ca. 500 m [Mitra et al., 2009; Utami et al., 2009]. In addition, mean daily travel distances of adult orangutans in Tuanan range from 700 to 900 m (unpubl. data). The farther away a conspecific's position to a long calling male, the smaller is the angle for a straight approach of this individual to the long call source. Thus, a triangulation error of 50 m on average—which accounts for 7–9% of the average inter-individual distances and 5–7% of the mean distance traveled by orangutans—has a minor effect on conclusions concerning ranging responses.

This study shows that an ALS, at this microphone density and number, is an effective tool to record and map all orangutan long-call occurrences in space and time in an area of 450–900 ha. One could argue that even at larger inter-microphone distances, viable results would be possible, but then microphone or recorder failure would have a much larger impact on data analysis. The ALS could enable us to monitor male presence in the study area as well as its ecological and social correlates. Therefore, several research questions can be addressed with the data acquired by the ALS. First, we will be able to track, over time and through space, the abundance of calling males in the area after identification is accomplished (in relation to ecological conditions and/or female reproductive state) as well as their ranging responses (male-male relationships). Second, we can monitor vocal responses to long calls by either focal flanged males or calling males recorded by the ALS (male-male relationships). Finally, we can use individual focal follows to document the ranging responses to long calling males by different sex–age classes (relationships between different sex–age classes). Thus, by combining an ALS with conventional individual focal follows, we can obtain a more complete picture of orangutan social organization.

While our study showed that an ALS works to detect orangutan long calls, we think that this method could easily apply to other primates emitting loud calls. This technology might especially enrich the study of species that are relatively difficult to observe or show a nocturnal lifestyle.

ACKNOWLEDGEMENTS

We gratefully acknowledge the Indonesian Institute of Science (LIPI), the Indonesian State Ministry for Research and Technology (RisTek), the Director General Departemen Kehutanan (PHKA), Departamen Dalam Negri, the local government in Central Kalimantan, the BKSDA Palangkaraya, the Bornean Orangutan Survival Foundation (BOSF), and MAWAS in Palangkaraya for their permission and support to conduct this research.

We also thank Universitas Nasional (UNAS) for support and collaboration and particularly Sri Suci Utami. Many thanks go to our local field team in Tuanan; Pak Rahmadt, Tono, Idun, Abuk, Kumpo, Suwi, Suga, and Ibu Fitriah. For their contribution to data collection we thank all local and foreign students and researchers. Thanks for the fruitful discussions and sharing the knowledge about the set-up of remote recorders to Hjalmar Kühl and Oliver Wagner and to the team of Wildlife Acoustics for their professional support. Many thanks go to Alison Ashbury for proofreading the manuscript. We were financially supported by the University of Zurich, the Leakey Foundation, the Basler Stiftung für biologische Forschung, and the Stiftung Dr. Joachim de Giacomi. We also thank the anonymous reviewer and Alex Piel for comments on this manuscript.

REFERENCES

Agranat I. 2009. Automatically identifiying animal species from their vocalizations. Concord, MA: Wildlife Acoustics Inc.

Andersson S, McGregor PK. 1999. Animal communication: what is the signal to noise ratio. Trends in Ecology & Evolution 14: 174–175.

Blumstein DT, Mennill DJ, Clemins P, et al. 2011. Acoustic monitoring in terrestrial environments using microphone arrays: applications, technological considerations, and prospectus. Journal of Applied Ecology 48: 758–767.

Bradbury JW, Vehrencamp SL. 1998. Principles of Animal Communication. Sunderland, MA: Sinauer Associates.

Brown TJ, Handford P. 2000. Sound design for vocalizations: quality in the woods, consistency in the fields. The Condor 102: 81–92.

Buxton RT, Jones IL. 2012. Measuring nocturnal seabird activity and status using acoustic recording devices: applications for island restoration. Journal of Field Ornithology 83: 47–60.

Collier TC, Blumstein DT, Girod L, Taylor CE. 2010. Is alarm calling risky? Marmots avoid calling from risky places. Ethology 116: 1171–1178.

Dabelsteen T. 2005. Public, private or anonymous? Facilitating and countering eavesdropping. In: McGregor PK editor, Animal Communication Networks. Cambridge: Cambridge University Press. p 38–62.

Delgado RA, Lameira AR, Ross MD, et al. 2009. Geographic variation in orangutan long calls. In: Wich SA Utami Atmoko SS Mitra Setia T van Schaik CP editors, Orangutans: geographic variation in behavioral ecology and conservation. Oxford: Oxford University Press. p 215–224.

Delgado RA. 2003. The function of adult male long calls in wild orangutans (Pongo pygmaeus) [Dissertation]. Durham, NC: Duke University.

Fitzsimmons LP, Foote JR, Ratcliffe LM, Mennill DJ. 2008. Eavesdropping and communication networks revealed through playback and an acoustic location system. Behavioral Ecology 19: 824–829.

Foote JR, Fitzsimmons LP, Mennill DJ, Ratcliffe LM. 2008. Tied to the nest: male black-capped chickadees decrease dawn chorus movement behaviour when their mate is fertile. Animal Behaviour 76: 1227–1233.

Galdikas BFM. 1983. The orangutan long call and snag crash at Tanjung Puting Reserve. Primates 24: 371–384.

Galdikas BMF. 1985a. Adult male sociality and reproductive tactics among orangutans at Tanjung Puting. Folia Primatologica 45: 9–24.

Lameira AR, Wich SA. 2008. Orangutan long call degradation and individuality over distance: a playback approach. International Journal of Primatology 29: 615–625.
Lippold S, Fitzsimmons L, Foote J, Ratcliffe L, Mennill D. 2008. Post-contest behaviour in black-capped chickadees (*Poecile atricapillus*): loser displays, not victory displays, follow asymmetrical countersinging exchanges. Acta Ethologica 11: 67–72.
MacKinnon J. 1979. Reproductive behaviour in wild orangutan populations. In: Hamburg DA McCown ER editors, The Great Apes. Menlo Park, CA: Benjamin/Cummings Publishing Company. p 257–273.
Maciej P, Fischer J, Hammerschmidt K. 2011. Transmission characteristics of primate vocalizations: implications for acoustic analyses. PLoS ONE 6: e23015.
Mathevon N, Dabelsteen T, Blumenrath SH. 2005. Are high perches in the blackcap Sylvia atricapilla song or listening posts? A sound transmission study. The Journal of the Acoustical Society of America 117: 442–449.
McGregor PK, Dabelsteen T, Clark CW, Bower JL, Holland J. 1997. Accuracy of a passive acoustic location system: empirical studies in terrestrial habitats. Ethology Ecology & Evolution 9: 269–286.
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, Burt JM, Fristrup KM, Vehrencamp SL. 2006. Accuracy of an acoustic location system for monitoring the position of duetting songbirds in tropical forest. The Journal of the Acoustical Society of America 119: 2832–2839.
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.
Mitani JC. 1985b. Sexual selection and adult male orangutan long calls. Animal Behaviour 33: 272–283.
Mitra Setia T, Delgado RA, Utami Atmoko SS, Singleton I, van Schaik CP. 2009. Social organization and male-female relationships. In: Wich SA Utami Atmoko SS Mitra Setia T van Schaik CP editors, Orangutans geographic variation in behavioural ecology. Oxford: Oxford University. p 245–253.
Peake TM, Terry AMR, McGregor PK, Dabelsteen T. 2001. Male great tits eavesdrop on simulated male-to-male vocal interactions. Proceedings of the Royal Society of London Series B: Biological Sciences 268: 1183–1187.
Piel AK. 2014. Savanna Sounds: using remote acoustic sensing to study spatiotemporal patterns in wild chimpanzee loud vocalizations in the Issa Valley, Western Tanzania [Dissertation]. San Diego: University of California, p 192.
R Core Team. 2013. R: A Language and Environment for Statistical Computing. 3.0.1 ed. Vienna, Austria.
Rijksen HD. 1978. A field study on Sumatran orang utans (Pongo pygmaeus abelii) : ecology, behaviour and conservation. Wageningen, The Netherlands: H. Veenman and Zonen.
Spillmann B, Dunkel LP, van Noordwijk MA, et al. 2010. Acoustic properties of long calls given by flanged male orangutans (Pongo pygmaeus wurmbii) reflect both individual identity and context. Ethology 116: 385–395.
Thompson ME, Schwager SJ, Payne KB, Turkalo AK. 2010. Acoustic estimation of wildlife abundance: methodology for vocal mammals in forested habitats. African Journal of Ecology 48: 654–661.
Utami SS, Singleton I, van Noordwijk MA, van Schaik CP, Mitra Setia T. 2009. Male-male relationships in orangutans. In: Wich SA Utami Atmoko SS Mitra Setia T van Schaik CP editors, Orangutans Geographigc Variation in Behavioral Ecology and Conservation. Oxford: Oxford University Press. p 235–244.
Waser PM, Brown CH. 1986. Habitat acoustic and primate communication. American Journal of Primatology 10: 135–154.
Watkins WA, Schevill WE. 1972. Sound source location by arrival-times on a non-rigid three-dimensional hydrophone array. Deep Sea Research and Oceanographic Abstracts 19: 691–706.
Wilson DR, Battiston M, Brzustowski J, Mennill DJ. 2013. Sound Finder: a new software approach for localizing animals recorded with a microphone array. Bioacoustics 23: 99–112.
Wrege PH, Rowland ED, Bout N, Doukaga M. 2012. Opening a larger window onto forest elephant ecology. African Journal of Ecology 50: 176–183.
van Schaik CP, Wich SA, Utami SS, Odom K. 2005. A simple alternative to line transects of nests for estimating orangutan densities. Primates 46: 249–254.
van Schaik CP, Damerius L, Isler K. 2013. Wild orangutan males plan and communicate their travel direction one day in advance. PLoS ONE 8: e74896.
van Schaik CP. 1999. The socioecology of fission-fusion sociality in orangutans. Primates 40: 69–86.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web-site.