Comparison of stationary acoustic monitoring and visual observation of finless porpoises

Satoko Kimura\textsuperscript{a)}
Graduate School of Informatics, Kyoto University, 606-8501 Kyoto, Japan

Tomonari Akamatsu
National Research Institute of Fisheries Engineering, Hasaki, Kashima, Ibaraki 314-0421, Japan

Kexiong Wang,\textsuperscript{b)} Ding Wang, Songhai Li, and Shouyue Dong
Institute of Hydrobiology, Chinese Academy of Sciences, Wahan 430072, People's Republic of China

Nobuaki Arai
Graduate School of Informatics, Kyoto University, 606-8501 Kyoto, Japan

(Received 8 May 2008; revised 6 October 2008; accepted 13 October 2008)

The detection performance regarding stationary acoustic monitoring of Yangtze finless porpoises \textit{Neophocaena phocaenoides asiaeorientalis} was compared to visual observations. Three stereo acoustic data loggers (A-tag) were placed at different locations near the confluence of Poyang Lake and the Yangtze River, China. The presence and number of porpoises were determined acoustically and visually during each 1-min time bin. On average, porpoises were acoustically detected 81.7 ± 9.7\% of the entire effective observation time, while the presence of animals was confirmed visually 12.7 ± 11.0\% of the entire time. Acoustic monitoring indicated areas of high and low porpoise densities that were consistent with visual observations. The direction of porpoise movement was monitored using stereo beams, which agreed with visual observations at all monitoring locations. Acoustic and visual methods could determine group sizes up to five and ten individuals, respectively. While the acoustic monitoring method had the advantage of high detection probability, it tended to underestimate group size due to the limited resolution of sound source bearing angles. The stationary acoustic monitoring method proved to be a practical and useful alternative to visual observations, especially in areas of low porpoise density for long-term monitoring. © 2009 Acoustical Society of America. [DOI: 10.1121/1.3021302]

PACS number(s): 43.80.Ka, 43.80.Jz, 43.66.Hg [WWA] Pages: 547–553

I. INTRODUCTION

Visual observation of surfaced cetaceans is well established and has been applied widely to species ranging from small odontocetes to large baleen whales. However, data gathered using visual surveys are limited to daytime since this is the only time visual observation is possible. Weather conditions such as fog and glare also have considerable effects on the visibility of animals. Patient long-term visual observation can be very costly, particularly under the very low-density conditions of endangered species.

The Yangtze finless porpoise (\textit{Neophocaena phocaenoides asiaeorientalis}), a freshwater porpoise subspecies unique to the Yangtze River in China, is a typical example. In the early 1990s, the population size was estimated at approximately 2700 individuals (Zhang \textit{et al.}, 1993). By 2006, estimates had decreased to as low as 1800 over the porpoise's entire distribution range (Zhao \textit{et al.}, 2008). Recent genetic studies have confirmed that populations of the Yangtze finless porpoise are scattered throughout the habitat area (Zheng \textit{et al.}, 2005), so monitoring them requires a great deal of effort. In addition, the Yangtze finless porpoise is one of the most difficult species to observe visually due to the turbid river's low visibility (less than 1 m) and the porpoise's lack of dorsal fin and rostrum.

Use of acoustic monitoring can avoid some major difficulties related to visual observation. Researchers have recently applied stationary acoustic monitoring methods to observe many species of aquatic mammals in various water systems. These methods are considered to be suitable for long-term automatic monitoring. The underwater sounds produced by aquatic animals can be used to monitor various characteristics of a species, including presence, behavior, and distribution (Nishimura and Conlon, 1994; Janik, 2000; Janik \textit{et al.}, 2000; van Parijs \textit{et al.}, 2002; Au and Benoit-Bird, 2003; Ichikawa \textit{et al.}, 2006; Tsutsui \textit{et al.}, 2006). For example, researchers have used the T-POD (a passive acoustic porpoise or dolphin detector system) to monitor harbor porpoises and bottlenose dolphins (Thomsen \textit{et al.}, 2005; Philpott \textit{et al.}, 2007). The T-POD system can detect the presence and sensing effort of echolocating animals, indicated by the detection rate of clicks per hour or day (Thomsen \textit{et al.}, 2005; Verfuß \textit{et al.}, 2007), the number of minutes containing clicks (Carstensen \textit{et al.}, 2006), the click characteristics of animals (Philpott \textit{et al.}, 2007), and the length of interclick intervals (Leeney \textit{et al.}, 2007). However, unlike visual observation, the T-POD system is not suitable for counting the

\textsuperscript{a)}Electronic mail: sk0130@bre.soc.i.kyoto-u.ac.jp
\textsuperscript{b)}Electronic mail: wangk@ihb.ac.cn
specific number of animals because it is a monaural system. Wang et al. (2005) used the stereo acoustic data logger, A-tag (Little Leonardo, Tokyo, Japan), in an oxbow of the Yangtze River that contained an ex situ conservation area for finless porpoises. The researchers found a weak positive linear correlation between the number of recorded signals and the group size of sighted porpoises. The number of signals may be an indicator of the number of individuals in a group, but because the sound production ratio varies between animals, Wang et al. (2005) were not successful in using stationary acoustic data to determine the exact number of porpoises.

Counting the number of finless porpoises from a moving boat is possible using the sound source bearing angle, monitored by a stereo acoustic system (Akamatsu et al., 2008). Acoustic transect observation from a moving platform effectively prevents double-counting of animals, whereas stationary acoustic observations require additional evaluation by comparing with ground truth data such as visually observed number of animals to determine the number of animals. In this study, we used a stereo acoustic monitoring system to conduct a stationary counting of finless porpoises in the channel where Poyang Lake flows into the Yangtze River. We compared the detection performance of a stationary acoustic monitoring system to that of visual observations.

II. MATERIALS AND METHODS

A. Study area

We conducted simultaneous acoustic and visual observations from boats at the confluence of the Yangtze River and Poyang Lake located in the middle reaches of the Yangtze River in South-Central China (Fig. 1). Three stations were used in the study area in April 27–29, 2006 and May 9 and 10, 2007. Data were collected over a summed period of 5 days. Station 0 (29° 45′06″ N, 116° 12′41″ E) was located at the point where the lake joined the main channel of the Yangtze River. Station 1 (29° 44′34″ N, 116° 12′10″ E) was located at the mouth of the lake approximately 1300 m upstream from Station 0. Station 2 (29° 44′02″ N, 116° 11′47″ E) was situated between two bridges and was located approximately 1100 m upstream from Station 1. During observation, boats at each station were fixed using double anchors to minimize drifting. Each boat engine was completely stopped. Water depth was approximately 3 m at all stations.

B. Acoustic data logger

We used stereo acoustic data loggers, A-tag (Little Leonardo Ltd., Tokyo, Japan, in 2006; Marine Micro Technology, Saitama, Japan, in 2007), for the acoustic observations. An A-tag is an event data logger that records sound pressure and the difference in time arrival between two hydrophones. It does not record the waveforms of received sound.

An A-tag consists of a stereo hydrophone, preamplifier with bandpass filter, CPU (PIC18F6620), flash memory (128 Mbytes), and lithium battery cell (CR2). The hydrophones had a sensitivity of MHP-140 (Marine Micro Technology) –201 dB (1 V/μPa) and a resonant frequency of 130 kHz, similar to the dominant frequency of finless porpoise sonar signals. This setting reduced noise outside the sensitive band of the hydrophone at sound reception. Hydrophone sensitivity was calibrated using an acoustical measurement tank (10 m in width, 15 m in length, and 10 m in depth) at the Fisheries Research Agency in Ibaraki, Japan. The ultrasonic sound transmission system used in calibration consisted of a function generator (NF1930A, NF Corp., Tokyo, Japan) and a transducer (B&K8103, Bruel & Kjaer, Naerum, Denmark); the system generated a 10-cycle tone burst for any frequency. A-tags were also confirmed to be able to record sounds made by free-ranging porpoises in an ex situ oxbow of the Yangtze River (Akamatsu et al., 2005a).

Each A-tag had two hydrophones, approximately 170 mm apart, which were used to identify the sound source direction. Electronic bandpass filters at the preamplification stage were adjusted to 70–300 kHz (in 2006) or 55–235 kHz (in 2007) to match the frequency band of Yangtze finless porpoise sonar signals, which ranges from 87 to 145 kHz and averages at 125 ± 6.92 kHz (Li et al., 2005a). The acoustic data logger recorded sound pressure at the primary and secondary hydrophones, as well as the difference in sound arrival times between the two hydrophones, every 0.5 ms (2 kHz event sampling frequency). The three data sets and the absolute time were recorded automatically only when the received sound pressure was greater than the trigger level of the primary hydrophone. Otherwise, no data were stored to conserve memory capacity. An A-tag can record information up to 30–40 h, depending on the number of pulses stored.

Peak-to-peak source levels for this species were 163.7–185.6 dB, referred to 1 μPa (Li et al., 2006), and the sound pressure level off the beam axis at 90° reached a maximum of 162 dB (Akamatsu et al., 2005b). Transmitted sound pressure levels can be highly variable, but off-axis signals still reached significant levels during this study, and it was possible to observe them using the data logger. We set the detection threshold level of the data logger at 135.3 dB. Our calibration experiment revealed that each A-tag had a slightly different threshold level, but the threshold level of 135.3 dB was higher than any one A-tag threshold. Our off-line analysis used recorded pulses greater than 135.3 dB (5.85 Pa). We allowed a maximum of 50.3 dB propagation loss for detecting signals. Assuming a simple spherical propagation model based on the freshwater values set out by Fisher and Simmons (1977) (absorption coefficient of
the direction of movement watched for 1 h and rested for 30 min; eye height was ap-
observed from the same anchored boat. Observers each covered a 90° arc from the boats. Observers
neous visual observations from the same anchored boat; four
rent direction to monitor the direction of porpoise movement
between the river and the lake. The primary hydrophone of
the A-tag was directed upstream of the site
Poyang Lake
stream
D. Visual observations
During acoustic observations, we conducted simulta-
neous visual observations from the same anchored boat; four
observers each covered a 90° arc from the boats. Observers
watched for 1 h and rested for 30 min; eye height was ap-
proximately 2 m above the water surface. When porpoises
were sighted, the observer recorded the minimum group size,
the direction of movement (upstream or downstream), and
the distance and bearing angle from the bow of the survey
boat. These parameters were the same as those measured by
the stereo acoustic data logger, with the exception of dis-
tance. To ensure that these results could be compared with
those obtained by acoustic detection, we only recorded visual
observation data detected within 300 m, similar to the acous-
tical detection range. The minimum group size was defined
as the number of the animals that respired successively
within a few seconds because this species has an average
shallow dive time of 4.86 ± 4.72 s (Akamatsu et al., 2002).
For the purposes of analysis, groups separated by more than
1 min were considered to be different sightings because this
species has an average deep dive time of 70.9 ± 22.9 s (Aka-
matsu et al., 2002). Currents and winds affected the direction
of the observation boat; this parameter was identical with
the direction of the data logger. The direction of the boat’s bow
was used as a reference to synchronize data collected
through acoustic and visual observations.

C. Acoustic observations
We used a bamboo rod to fix the acoustic data logger at
a 1-m depth from the side of each anchored boat. In 2006,
we fixed two A-tags underwater from boats at Stations 1 and
2, and we fixed an additional A-tag at Station 0 in 2007. The
stations were spaced more than 1000 m apart, well outside
the A-tag detection range of 290 m. This design ensured that
the observations at each station were independent. We as-
sumed no simultaneous detection of individual animals. The
two hydrophones of each A-tag were set parallel to the cur-
rent direction to monitor the direction of porpoise movement
between the river and the lake. The primary hydrophone of
the A-tag was directed upstream of the site (Poyang Lake
side), and the secondary hydrophone was directed down-
stream (the Yangtze River side).

E. Acoustic signal processing
We eliminated contamination from noise and reflection
and calculated the interclick intervals and relative angles of
sound sources using a custom-made program developed us-
ing IGOR PRO 5.03 (WaveMetrics, Lake Oswego, OR). Relative angles to the sound source were calculated using the
difference in time arrival between the two hydrophones.
Sample data shown in Fig. 2 illustrate sound pressure,
relative bearing angle, and the interclick interval of porpoise
clicks. We were able to track porpoises easily because they
phonated frequently. As shown in the figure, interclick inter-
vals and sound pressure levels changed smoothly (Akamatsu et al., 2005c), while background or boat noise caused ran-
domly changing patterns in the interclick interval and sound
pressure. We were unable to use frequency information to
exclude noise because A-tags do not record waveform. In-
stead, we used interclick intervals to discriminate signals
from noise. We excluded any successive clicks greater than
twice or less than half the previous interclick intervals (Aka-

The multipath propagation in the Yangtze River can
cause echolocation signals to have a multipulse structure (Li
et al., 2005b). In this shallow freshwater system, reflected
signals came just after the direct path signal. Because the
animals had a very shallow depth, the surface reflection had
an angle similar to the direct path signal; this resulted in the
echo’s very short delay time. Pulses within 2 ms after a di-
rect path pulse were eliminated during offline signal process-
ing. Since the mean minimum lag time to process returning
echoes inside an animal brain is 2.5 ms (Au, 1993), por-
poises’ sound is considered to be not excluded in this pro-
cessing.

F. Number and movement direction of animals
The number and movement direction of animals were
determined manually from click trains. This species usually
produces an interclick interval shorter than 130 ms (Li et al.,
Simultaneous phonation of two individuals swimming close together could be identified through the double different cyclic characteristics of the sound pressure and/or inter-click intervals within a single trace, so we counted these as originating from two porpoises. This phenomenon was relatively easy to discriminate from reflections because reflection sound always involves a separation time after the direct path click.

In contrast, when single periodicity in interclick intervals and/or a smoothly changing sound pressure accompanied close parallel traces, we counted only one porpoise. These parallel traces were caused by an error in the trigger point among multiple wavelengths in a click. The trigger point of primary and secondary hydrophones could differ when the sound pressure at the onset of a click is comparable to the detection threshold level. Among finless porpoises, click amplitude rises gradually. Therefore, the second wave highlight next to the first onset wave tends to be triggered by the secondary hydrophone, even if the first onset was triggered by the primary hydrophone. One wavelength ambiguity of the trigger point occurred, resulting in close parallel traces of a single phonating animal.

Animals were counted visually through observation around the survey boat during the same time bin as they were counted using the acoustic method. If porpoises passed near the observation station, they were likely to be observed once within a 1-min time bin, which is close to the average respiration interval of an adult finless porpoise engaged in deep diving activity.

We also compared the visually and acoustically measured movement directions of animals. In the acoustic method, direction was determined by changes in the bearing angles of received sounds. A change in bearing angle from positive to negative indicated that the porpoise moved from the lake side to the river side, and vice versa. The difference in time arrival between the two hydrophones, correlated with the bearing angles, had a minimum resolution 13.6 μs. When the difference in time arrival was considerably greater than 13.6 μs, the swimming direction was determined to be either upstream toward Poyang Lake or downstream toward the Yangtze River. Otherwise, we did not record a swimming direction. When the trace consisted of only one click train, determining the swimming direction was impossible. When the primary hydrophone of the data logger was triggered but the secondary hydrophone received an insufficient sound level, the difference in time arrival was 0 and indicated as a line at 90°, as shown in Fig. 3. We did not use these measurements to count individuals but used them instead to identify simultaneous phonation of multiple individuals.

III. RESULTS

We obtained 1216 min of effective visually and acoustically measured data at Station 1 and 504 min at Station 2 in 2006, and 464 min at Station 0 in 2007, for an overall total of 2184 min of observations.

In total, 2987 and 591 animals were detected acoustically and visually. At Stations 0, 1, and 2, respectively, animals were detected acoustically in 92.9%, 76.2%, and 76.0% of all time bins, whereas animals were detected visually in 23.5%, 13.1%, and 1.6% of all time bins. On average, porpoises were detected acoustically in 81.7% ± 9.7% and visually in 12.7% ± 11.0% of all observation times; the acoustic detection rate during the total observation time was significantly higher than the visual detection rate (Scheffe’s test, \( P < 0.01 \)). As shown in Fig. 4, both methods detected the most porpoises at Station 0 and the least at Station 2 (Scheffe’s test, \( P < 0.01 \)). Detection rates differed among observation sites.

We monitored the swimming direction of porpoises using bearing angles and compared the results of acoustic observations of swimming direction with visual observations. At each station, the number of positive swimming direction identifications divided by the total observation time was similar for both observation methods (Fig. 5).

Figure 6 shows the number and size of detected groups by time bin. Note that the ordinate is logarithmic. Over the total observation time (2184 min), both methods determined the same group size for only 458 min. The most detected numbers of animals were zero by the visual method (1881 min) and one by the acoustic one (866 min). The
acoustic data logger could count group sizes to a maximum of five individuals, whereas visual observation could count group sizes to a maximum of ten individuals.

IV. DISCUSSION

Stationary acoustic monitoring was effective for counting Yangtze finless porpoises that were echolocating; this method yielded a detection rate seven times higher than visual observation (Fig. 4). The results clearly show that the acoustic method was more effective at detecting the presence of animals than the visual method.

We were able to detect porpoises frequently using the acoustic method, while only occasionally using visual observation. This was a result of different visual and acoustic cues from porpoises. Porpoises can be recorded acoustically when they produce sonar phonates within a detection range. Yangtze finless porpoises produce sonar click trains every 5–6 s on average (Akamatsu et al., 2005c, 2007), so their frequent phonation resulted in a high detection rate using the acoustic method. In contrast, we were only able to observe the porpoises visually when they surfaced in the turbid water of the Yangtze River. Among adult Yangtze finless porpoises, long dive duration averages 70.9 s (Akamatsu et al., 2002). The acoustic and visual detection methods used cues of differing intervals. In addition, because finless porpoises are small cetaceans, do not have dorsal fins, and exhibit little aerial behavior, they are easily overlooked even when they are near the surface.

The stereo acoustic data logger systems revealed high and low porpoise density areas, which agreed with visual observations (Fig. 4). The average number of detected porpoises was highest at Station 0, which was near the confluence of the lake and the river. This finding is consistent with previous research indicating that Yangtze finless porpoises tend to aggregate in that area (Wei et al., 2003).

Acoustic observations were used successfully to detect the movement direction of porpoises underwater (Fig. 5). The stereo system was more powerful than a monaural system because it could separate sound sources to count the number of animals and also identify their swimming direction. This feature is most suitable for long-term monitoring of porpoise migration by using several A-tags.

The number of time bins in which no porpoises were detected acoustically (447 min) was about one-quarter the number in which none were detected visually (1881 min). Therefore, the acoustic method missed fewer porpoises than the visual method. A towed acoustic survey also resulted in a large ratio of individual animals that were missed (Akamatsu et al., 2008). The acoustic method could detect group sizes to a maximum of five individuals but tended to underestimate the size of larger groups (Fig. 6). When the group size was fewer than four individuals, the acoustic method detected porpoises in 1728 1-min bins, approximately four times the number of bins in which porpoises were detected visually (286 min).

The inability of the acoustic method to detect more than five individuals in a 1-min bin was probably due to the limited resolution of the stereo acoustic data logger’s bearing angle. The short distance between the two hydrophones (170 mm) was a possible cause for this limitation. A longer baseline should improve the bearing angle resolution. Other
possible causes may have been the small source level of porpoises (including off-axis sounds), alternate phonations of multiple individuals near each other, or eavesdropping to maintain silence. In addition, previous research has suggested that large groups may not vocalize as much as small groups (Götz et al., 2006). Several porpoises swimming together within the detection range phonate alternately, but our passive acoustic system was unable to differentiate them. Yangtze finless porpoises, however, often swim alone or in very small groups and in areas in which visual observations may not be practical, as indicated in Fig. 6. The results indicate that our acoustic monitoring system would have a limited application to species that form larger groups. We used a conservative criterion when counting the number of porpoises to avoid double-counting.

In conclusion, the stationary acoustic monitoring system using stereo acoustic data loggers performed more efficiently than the visual method, especially in areas of low-density echolocating animals. Stationary acoustic observation is suitable for use in areas in which porpoises appear infrequently and form small groups, where visual observation may not be practical. The acoustic system appears to be powerful at monitoring porpoises in a narrow channel such as a river system. In future research, we will monitor porpoise migration using multiple acoustic monitoring systems.

ACKNOWLEDGMENTS

We thank all members of the Laboratory of Conservation Biology of Aquatic Animals at the Institute of Hydrobiology, Chinese Academy of Sciences, the National Research Institute of Fisheries Engineering, the Japan Fisheries Resource Conservation Association, Biosphere Informatics, Graduate School of Informatics, Kyoto University, and the Fisheries Administration Bureau of Hukou Jiangxi Province. This study was partly supported by a Grant-in-aid for Scientific Research (B) (19405005) from the Japanese Research and Development Program for New Bio-Industry Initiatives, the National Natural Science Foundation of China (30730018 and 30570233), the Chinese Academy of Sciences (Present Fund), the Ocean Park Conservation Foundation of Hong Kong, and the Sasagawa Scientific Research Grant from The Japan Science Society (19-740M).


The population of finless porpoise in the middle and lower reaches of Yangtze River,” Acta Theriologica Sinica 16, 490–496.