

Acoustic discrimination between harbor porpoises and delphinids by using a simple two-band comparison

Saho Kameyama^{a)}

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

Tomonari Akamatsu

National Research Institute of Fisheries Engineering, Hasaki, Kamisu, Ibaraki 314-0408, Japan

Ayhan Dede and Ayaka Amaha Öztürk

Faculty of Fisheries, Istanbul University, Ordu Cad, No. 200, Laleli, Istanbul, Turkey

Nobuaki Arai

Field Science Education and Research Center, Kyoto University, Kyoto 606-8502, Japan

(Received 29 April 2013; revised 11 May 2014; accepted 6 June 2014)

A simple discrimination method between Delphinidae and Phocoenidae based on the comparison of the intensity ratios of two band frequencies (130 and 70 kHz) is proposed. Biosonar signals were recorded at the Istanbul Strait (Bosphorus) in Turkey. Simultaneously, the presence of the species was confirmed by visual observation. Two types of thresholds of two-band intensity ratios, fixed and dynamic threshold, were tested for identification. The correct detection and false alarm rates for porpoises were 0.55 and 0.06 by using the fixed threshold and 0.74 and 0.08 by using the dynamic threshold, respectively. When the dynamic threshold was employed, the appropriate threshold changed depending on the mix ratio of recorded sounds from both Delphinidae and Phocoenidae. Even under biased mix ratios from 26% to 82%, the dynamic threshold worked with >0.80 correct detection and <0.20 false alarm rates, whereas the fixed threshold did not. The proposed method is simple but quantitative, which can be applicable for any broadband recording system, including a single hydrophone with two frequency band detectors.

© 2014 Acoustical Society of America. [<http://dx.doi.org/10.1121/1.4884763>]

PACS number(s): 43.80.Ka [JJF]

Pages: 922–929

I. INTRODUCTION

Passive acoustic monitoring has been widely used in cetacean research (reviewed by Mellinger *et al.*, 2007), especially to determine the presence and abundance of species by using towed hydrophone arrays (e.g., Barlow and Taylor, 2005) and the long-term trend by using fixed hydrophone systems (Verfuss *et al.*, 2007; Kimura *et al.*, 2010). Such monitoring is used not only for detecting the presence of animals but also in identifying the phonating species. For example, baleen whale calls are appropriate targets to identify the location and species because of their stereotype structure. A combination of low-frequency pulsed calls, or A calls, and low-frequency tonal calls, or B calls, of blue whales in the Pacific Ocean can be easily discriminated from ambient noise in offshore calm waters (Oleson *et al.*, 2007). These call features also show geographic differences between populations in the Atlantic Ocean and those in the Pacific Ocean (McDonald *et al.*, 2006; Gavrilov *et al.*, 2012). Only male humpback whales and fin whales generate sequences of calls, which are considered useful for mate selection during their breeding season (Payne and McVay, 1971; Darling and Berube, 2001; Croll *et al.*, 2002). The sequence of specific

call types of these whales can be differentiated from other sound sources.

In the case of dolphins, whistles are considered as acoustic keys for species identification. A previous study showed that the female bottlenose dolphin calls showed the same contour on a sonogram for about 12 years (Sayigh *et al.*, 1990; Quick and Janik, 2012). Oswald *et al.* (2007) developed a real-time odontocete call classification algorithm for nine species of dolphins, based on the basic acoustic features in whistles. However, vocal learning in odontocetes, as shown in killer whales (Deecke *et al.*, 2000) and bottlenose dolphins (Quick and Janik, 2012), could cause changes in vocal characteristics, thus making it further difficult to identify the species.

On the other hand, the spectrum and waveform of biosonar sounds in Delphinidae and Phocoenidae show clear differences and are relatively stable within the same individual and the same species. Most of Delphinidae use broadband signals. Its biosonar sound contains energy that encompasses a wide range of frequencies, ranging from 23 to 134 kHz (Au, 1993), and its power spectrum level shows a relatively flat shape. In contrast, the Phocoenidae, including the harbor porpoise, uses narrowband high-frequency signals with peak energies of approximately 130 kHz (Madsen *et al.*, 2005; Villadsgaard *et al.*, 2007). Soldevilla *et al.* (2008) suggested that even among delphinids, classifying Risso's dolphin and the Pacific white-sided dolphin is possible based on clicks.

^{a)}Author to whom correspondence should be addressed. Electronic mail: kamesaho@bre.soc.i.kyoto-u.ac.jp

The T-POD (timing porpoise detector) and its successor, C-POD (cetacean porpoise detector), were developed to differentiate the harbor porpoise from the delphinids by using these acoustic characteristics of clicks. This equipment has been extensively used in field studies (Verfuss *et al.*, 2007; Philpott *et al.*, 2007; Todd *et al.*, 2009; Dahne *et al.*, 2013). T-POD and C-POD are efficient systems for long-term fixed operations. However, they have device-dependent algorithms (Kyhne *et al.*, 2008; Verfuss *et al.*, 2013) and different versions of T-POD or C-POD show variable detection performances.

In recent years, several types of passive acoustic monitoring devices have been developed and installed in the ocean to monitor odontocetes and other marine animals (Sousa-Lima *et al.*, 2013). For quantitative comparisons among these observations, simple acoustic characteristics are needed for species or family identification. Ideally, the method should not depend on device type and should discriminate species with degree of precision.

In this study, we conducted the simplest two-band frequency intensity comparison (130 and 70 kHz) for the detection of odontocetes in the Istanbul Strait (Bosphorus) in Turkey. This strait is suitable for the evaluation of a species discrimination method because three odontocete species are simultaneously found here: Harbor porpoise (*Phocoena phocoena*), short-beaked common dolphin (*Delphinus delphis*), and bottlenose dolphin (*Tursiops truncatus*). Herein, we show that harbor porpoises can be differentiated from delphinids, the latter two species, by using the two-band intensity ratio with correct detection and false alarm rates. In addition, we propose a model that can quantify how the proposed method works under various mix ratios of sounds from each family.

II. MATERIALS AND METHODS

A. Acoustic event recorder with two-band intensity

We used a stationary acoustic data-logger, A-tag (ML200-ASII; Marine Micro Technology, Inc., Saitama, Japan; Fig. 1) for passive acoustic monitoring. It consisted of two hydrophones placed 60 cm apart, a central processing unit (CPU) (PIC18F6620; Microchip Technology, Inc.,

Chandler, AZ), a 128 MB flash memory, and 2 UM1 batteries. The CPU, flash memory, and batteries were housed in a waterproof aluminum case. A passive two-pole bandpass filter of 55–235 kHz, which includes the peak frequency of the harbor porpoise click sound (129–145 kHz; Villadsgaard *et al.*, 2007) was used. The A-tag is an event data logger that records sound pressure and the difference in the arrival time of sound between two hydrophones when the sound is higher than the preset threshold (138 dB peak-to-peak re 1 μ Pa). The A-tag does not store the waveform of a sound.

For comparison of two-band intensity, the A-tag was equipped with two hydrophones of different sensitive frequency bands to fit the difference in the characteristics of clicks between Phocoenidae and Delphinidae. The primary and the secondary hydrophones were most sensitive at 130 and 70 kHz, respectively (Fig. 1). In addition, both hydrophones can receive sounds even outside its most sensitive resonance frequency, thus it can receive the low frequency component of the narrowband clicks of Phocoenidae. The A-tag records a received intensity when the received level at the primary hydrophone exceeds the pre-fixed detection threshold level (138 dB peak-to-peak re 1 μ Pa). To avoid insufficient triggering, the received signal intensity at the secondary hydrophone was amplified 6 dB more than that of the primary hydrophone because the spectrum intensity of the harbor porpoise sounds around 70 kHz is smaller than that of the delphinids. Note that this amplification does not affect the threshold of data storage because data storing does not happen without the trigger of the primary hydrophone. The received sound pressure by the two hydrophones and the differences in arrival time were recorded every 0.5 ms.

B. Acoustic data collection

We deployed the A-tag in the middle of the Istanbul Strait (Fig. 2; 41.05' 835" N, 29.03' 264" E) between April 12 and June 1, 2012, when the three species were expected to appear in the strait according to Öztürk *et al.* (2009). An iron pipe and a basket that were fixed beside the pier located at the bank of the Istanbul Strait were used to set the A-tag approximately 35 cm deep from the bottom. The tide level change was within 34 cm in this area (Alpar and Yüce, 1998). This is smaller compared to the depth of the A-tag, which was not

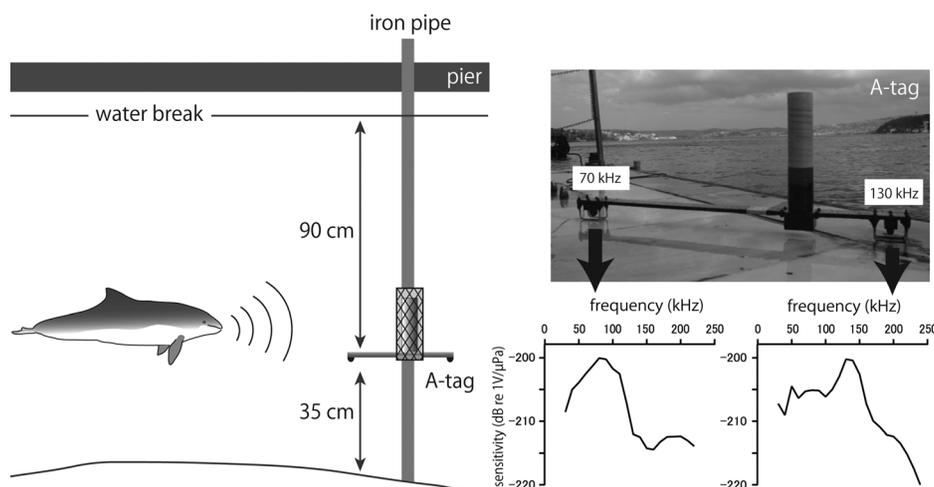


FIG. 1. Left: The A-tag was fixed using an iron pipe and a basket. Right: A photo of the A-tag (upper) and the frequency responses of each hydrophone (lower). Each hydrophone has different frequency sensitivity: The primary hydrophone is the most sensitive at 70 kHz and the secondary hydrophone is the most sensitive at 130 kHz.

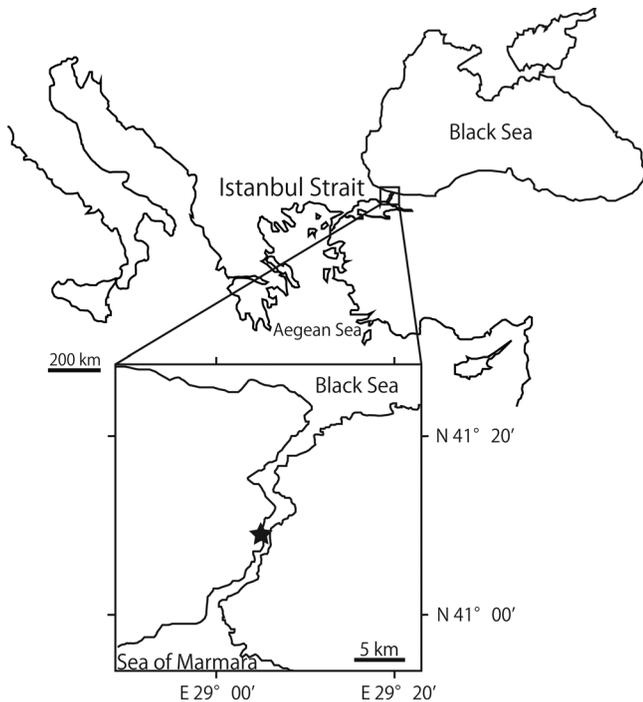


FIG. 2. A map of the acoustic observation site. The Istanbul Strait connecting the Black Sea to the Sea of Marmara is a part of the Turkish Straits system. The A-tag was deployed halfway between the two seas in the Istanbul Strait (indicated using star).

exposed to air at any time. The depth at the deployed point was $125 + 34$ cm depending on the tide level (Fig. 1).

The maximum detection range of the A-tag was calculated using the following parameter: Maximum source sound pressure level (Source SPL), which was estimated at 227 dB peak-to-peak re $1 \mu\text{Pa}$ for bottlenose dolphins (Au 1993; Simard *et al.*, 2010) and 205 dB peak-to-peak re $1 \mu\text{Pa}$ for harbor porpoise (Villadsgaard *et al.*, 2007); the minimum received SPL of the present A-tag (Received SPL: 138 dB peak-to-peak re $1 \mu\text{Pa}$); and the absorption coefficient of the Istanbul Strait (α). α was calculated by using the Francois and Garrison model (Francois and Garrison, 1982a,b) at 0.024 dB/m. The maximum detection range, d , was calculated using Eq. (1),

$$\text{Source SPL} = \text{Received SPL} + 20 \log_{10}(d) + \alpha d. \quad (1)$$

C. Visual observation

Visual observations were conducted simultaneously with the acoustic observations. During visual observations, the time, species, distance, angle from the observation site, group size, and swimming direction were recorded on each sighting. The time and species were used for analysis in this study. Two observers conducted observations from the bank side of a pier where the A-tag was deployed. Two to four persons participated in the visual observations. Two persons were on-duty, covering 90° each, thus ensuring the monitoring of a 180° sector from the observation point to the opposite shore. One observation interval lasted 30 min, and the position was changed alternately. If more than two observers were available, the remaining observers rested during off-duty periods.

D. Off-line data analysis

We applied an off-line filter using Igor Pro 6.2.2 (WaveMetrics, Inc., Lake Oswego, OR) for data collection to exclude false positives that were generated by background noise. First, we excluded the data that were not triggered by both hydrophones before applying the filtering. The off-line filter consisted of three steps. In the first step, the surface or bottom reflections, which could be recorded within 2.5 ms at most from the direct path sound, were excluded. We defined the boundary of different click trains at 200 ms pulse intervals (Akamatsu *et al.*, 2007). If the click sequences were separated by more than 200 ms, each sequence was considered a different click train. In the second step, the minimum number of clicks in a click train was defined. If there were only a few clicks in a click train, it was highly likely that these were background noise. We extracted the click trains that contained six or more pulses in the train (Akamatsu *et al.*, 2010; Kimura *et al.*, 2010). In the third step, the click trains that showed irregular changes in inter-click intervals were eliminated. The inter-click intervals of biosonar sounds in delphinids and porpoises should gradually change (Akamatsu *et al.*, 1998; Li *et al.*, 2009). If the coefficient of variation of the inter-click interval in a click train was below 0.3, these click trains may contain more than two animals' sonar signals or background noise that might have randomly changed the inter-click interval (Kimura *et al.*, 2010). Finally, manual checks were conducted to exclude false positives generated by noise caused by large ships approaching the device or the occasional contamination of artificial sonar sounds. These noises did not occur frequently but masked the data for several 10 s segments. The period of noise contamination was excluded from further analyses.

The ratio of received sound pressure by the two hydrophones (130 kHz divided by 70 kHz) was calculated for each click in a click train. A two-band ratio was defined as the averaged ratios of all clicks in a click train.

Acoustic datasets that clearly matched the visual observation results were selected for analysis. If we sighted one of the three species—harbor porpoise, short-beaked common dolphin, or bottlenose dolphin—we considered that the click trains recorded ± 10 min from the sighting were generated by the sighted animals. We excluded the click trains that were recorded when both harbor porpoise and delphinids were sighted within ± 10 min periods. The click trains without visual confirmation were not used, because it was difficult to determine the species that produced the click trains. Finally, we obtained the two-band ratio of each click train with species confirmation.

E. Species discrimination model

Histograms of the two-band ratios were drawn for the two families. These distributions could be overlapped because some of the two-band ratios of harbor porpoises could be the same or even smaller than those of Delphinidae and vice versa. A simple definition of an appropriate threshold of two-band ratios to discriminate species is the intersection point of two distributions [see Fig. 3(a) observed distribution]. Hereinafter, we refer the intersection point of

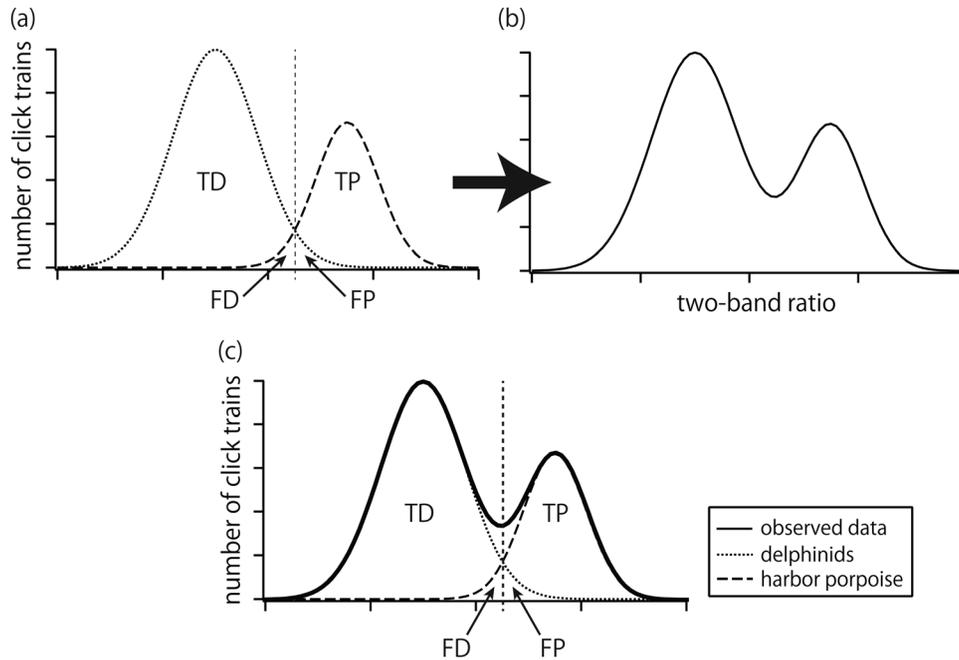


FIG. 3. The diagram of the two-band ratio with the correct detection of delphinids (TD) and the harbor porpoise (TP) and the false alarm for delphinids (FD) and the harbor porpoise (FP) and how to estimate the appropriate dynamic threshold from the observed acoustic data based on grand truth distribution. TD and FD are the number of click trains up to the threshold, and FP and TP are the rest of them. (a) The dynamic threshold with the correct detection rate and false alarm rate in each mix ratio were calculated in this study by using grand truth distribution. (b) Bimodal distribution of the delphinids and harbor porpoise were generated from acoustic observations. The mix ratio of each family is unknown in these data. (c) To estimate the appropriate mix ratio of (b), (a) can be used as the validation function by minimizing the differential of (a) and (b).

two distributions of observed data as the fixed threshold. To determine the intersection point, we used a 0.1 bin histogram. This simple method to discriminate species also includes some false alarm [indicated as FD and FP in Fig. 3(a)]. The shape of the two-band ratio distribution is assumed to be the same for each species regardless of the data size if enough data was used to develop the distribution. However, the intersection point of each distribution clearly changes depending on the mix ratio of the detected number of click trains in each species. For example, if the portion of click trains recorded from delphinids is higher than that from porpoises, the fixed threshold should shift toward the harbor porpoise side. To adjust this kind of change of mix proportions, we propose a method to calculate the dynamic threshold level of the two-band ratio.

Correct detection and false alarm are the functions of an arbitrary threshold (x). Ideally, the summation of the number of correct detections of delphinids (TD) and that of harbor porpoise (TP) should be maximized. Similarly, the number of false alarms of delphinid (FD) and that of harbor porpoise (FP) should be minimized. Thus, we define the value of the index of classification efficiency as shown in Eq. (2):

$$\text{Index of classification efficiency} = \frac{\text{TD}(x) + \text{TP}(x)}{\text{FP}(x) + \text{FD}(x)}. \quad (2)$$

To identify the maximum value of index of classification efficiency according to threshold (x), the differential of the index of classification efficiency by threshold (x) should be zero, as shown in the following Eq. (3),

$$\frac{\Delta \text{ Index of classification efficiency}}{\Delta \text{ threshold } (x)} = 0. \quad (3)$$

The dynamic threshold is when threshold (x) satisfies Eq. (3). In reality, the function of $\text{TD}(x)$, $\text{TP}(x)$, $\text{FD}(x)$, and $\text{FP}(x)$ are given numerically based on the acoustic identification compared to the visual ground truth data. Therefore, the index of classification efficiency was also calculated numerically.

F. Model validation

The proposed models of fixed and dynamic thresholds were validated using observed data with changing proportions of click trains in each species. Mix distributions of the two-band ratio, which contains the variable ratio of each family from 0% to 100%, were created. The portions of click trains of harbor porpoise and delphinids were β and $(100 - \beta)$, where $0 < \beta < 100$. We calculated the dynamic threshold that maximizes the index of classification efficiency [see Eqs. (2) and (3)] in each mix ratio. The fixed and dynamic thresholds were provided as a function of the mix ratio between numbers of click trains, β .

Applying fixed and dynamic thresholds to the calculated mix distribution, we could evaluate the correct detection and false alarm rates of each model as follows [Eqs. (4) and (5)]:

$$\text{Correct detection rate} = \frac{\text{TP}}{(\text{TP} + \text{FD})}, \quad (4)$$

$$\text{False alarm rate} = \frac{\text{FP}}{(\text{TD} + \text{FP})}. \quad (5)$$

Equations (4) and (5) are for harbor porpoise. P and D should be inverted for delphinids.

For quantitative evaluation, Matthews correlation coefficient (MCC) was introduced (Matthews, 1975; Baldi *et al.*, 2000), as shown in Eq. (6). MCC was used in machine learning to evaluate the quality of binary classifications that can be considered both correct detection and false alarm rate using one value. When FD and FP are 0, the value becomes 1, which indicates perfect classification. On the other hand, when either TD or TP is zero, the value becomes zero, which indicates that the classification failed,

$$MCC = \frac{(TP \times TD) - (FP \times FD)}{\sqrt{(TP + FP)(TP + FD)(TD + FP)(TD + FD)}}. \quad (6)$$

Once the threshold was derived, it was possible to estimate the mix ratio β of two families and the appropriate threshold. The acoustic data can only provide a bimodal distribution of the two-band ratio of both harbor porpoise and delphinids [Fig. 3(b)]. Assuming that the distribution of the two-band ratio of each family is nearly stable, the distribution obtained in the present study can be used as ground truth [Fig. 3(a)]. A difference in the observed distribution of the two-band ratio and a linear combination of ground truths of the two families were used as a validation function to estimate the appropriate threshold, as shown in Eq. (7). The image of threshold estimation is shown in Fig. 3(c),

$$\text{Difference} = \int | \text{observed} - ((100 - \beta) \text{delphinids} + \beta * \text{harbor porpoise}) |. \quad (7)$$

The “difference” is the difference between observed data, which is the data of sound pressure ratio observed in other experiments, and the grand truth data, which we used to calculate the threshold in this study. This difference should be minimized to estimate the closest model (i.e., appropriate β) to the observed data. The “observed,” “delphinids,” and “harbor porpoise” represent the number of click trains in each β that is the mix coefficient of each family’s distribution.

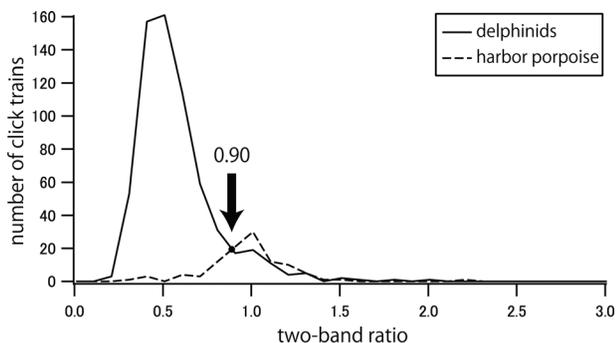


FIG. 4. Distribution of the two-band ratio of the observed click trains. The solid line represents the delphinids and the dotted line depicts the harbor porpoise. An arrow indicates the fixed threshold, and its value was 0.90.

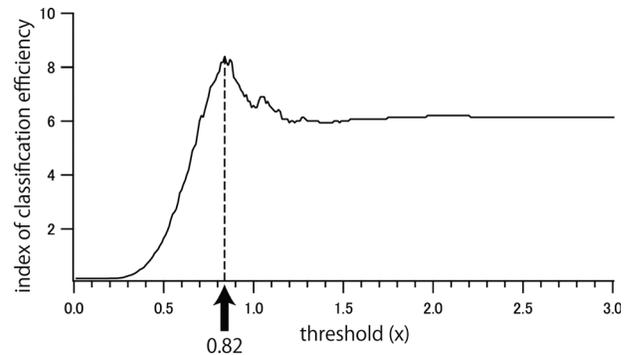


FIG. 5. Distribution of the index of classification efficiency calculated from each threshold (x) of the observed two-band ratio as shown in Fig. 4. The determined dynamic threshold is indicated by a black arrow, and its value was 0.82.

III. RESULTS

The recordings of 639 click trains of delphinids and 104 click trains of harbor porpoises with visual species confirmation were obtained after off-line filtering. The distribution of the two-band ratios is presented in Fig. 4. The fixed threshold (0.90) is represented by an arrow in Fig. 4. The index of classification efficiency was depicted according to Eq. (2) and the arbitrary threshold (x) in Fig. 5. The graph has a simple one-peak shape, and the dynamic threshold was defined as 0.82 that maximizes the index of classification efficiency. The dynamic threshold was slightly lower than the fixed threshold level of 0.90. The receiver operating characteristics (ROC) curve was used to describe and compare the performance of binary classification (Fig. 6). We changed the threshold in 0.01 steps from 0 to 3 and calculated the correct detection rate and false alarm rate. When the fixed threshold of 0.90 was used, the harbor porpoise showed a correct

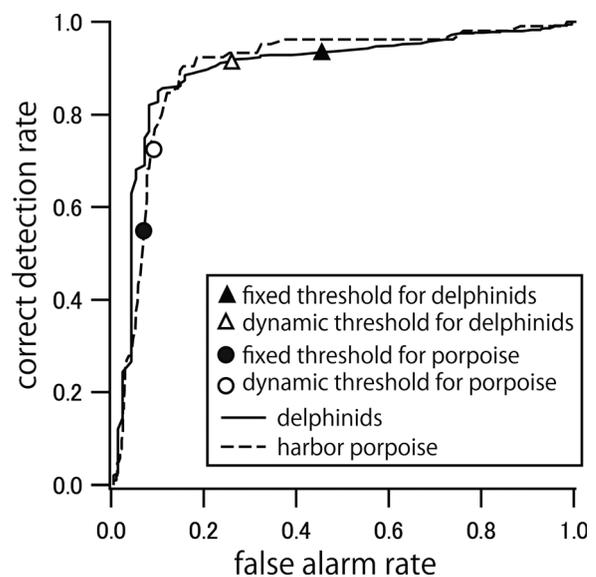


FIG. 6. The ROC curve of the observed two-band ratio. The solid line represents the delphinids and the dotted line depicts the harbor porpoise. Black markers indicate the correct detection rate and false alarm rate using the fixed threshold, whereas the white markers indicate those of the dynamic threshold. Triangles indicate delphinids and circles indicate the harbor porpoise, as calculated using the fixed threshold and the dynamic threshold, respectively.

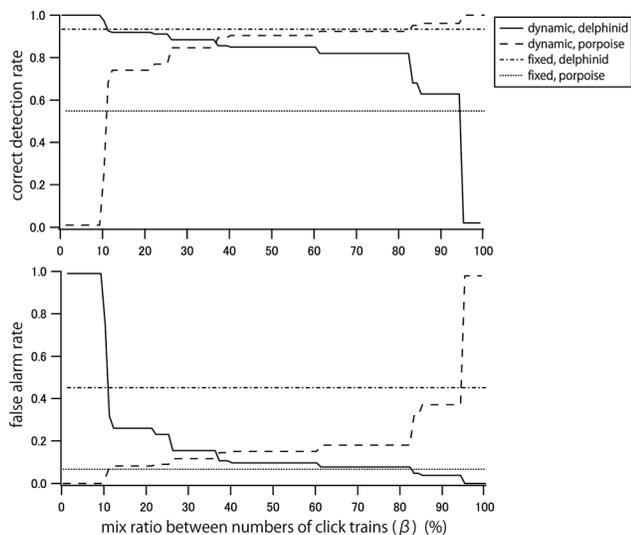


FIG. 7. The correct detection rate (upper graph) and false alarm rate (lower graph) of the delphinids and harbor porpoise as provided by the dynamic threshold and the fixed threshold in each mix ratio between the numbers of click trains, which can also be called mix ratio β .

detection and false alarm rate of 0.55 and 0.06, respectively. When the dynamic threshold of 0.82 was used, the harbor porpoise showed a correct detection and false alarm rate of 0.74 and 0.08, respectively. On the other hand, delphinids showed a correct detection and false alarm rate of 0.93 and 0.45, respectively, when the fixed threshold was used and 0.92 and 0.26, respectively, when the dynamic threshold was used. The present mixed families data correspond to $\beta = 14$, suggesting that the detection of click trains from delphinids were dominant. In this situation, false alarm rate of delphinids was large due to incorrect categorization of harbor porpoise sounds as those of the delphinids.

We changed the mix ratio β by conducting a numerical experiment. The calculated dynamic threshold is shown in Fig. 7 as a solid line. The dynamic threshold changed from 0.24 to 3.00 depending on the ratio between the number of click trains of harbor porpoises and delphinids β .

The correct detection and false alarm rates proposed by the dynamic threshold and fixed threshold for the created distribution are shown in Fig. 8. In the fixed threshold, the

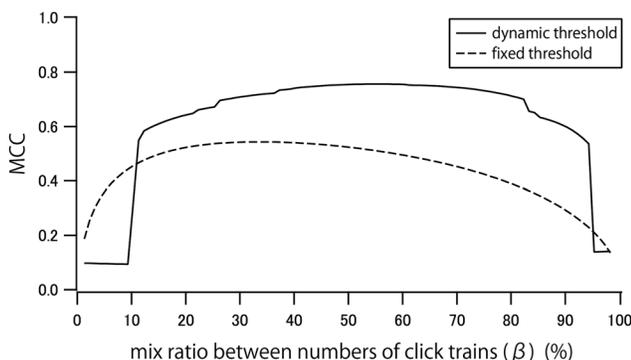


FIG. 8. The MCC value provided by the dynamic threshold (solid line) and the fixed threshold (dotted line) in each mix ratio between the numbers of click trains, which can also be called mix ratio β . The value of MCC of the dynamic threshold is higher than that of the fixed threshold, except when the mix ratio between the number of click trains is below 10% or above 95%.

correct detection and false alarm rates have a stable value. The correct detection rate was 0.93 for the delphinids and 0.55 for the harbor porpoise. On the other hand, the correct detection rate changed from 0.02 to 1 for the delphinids and from 0.01 to 1 for the harbor porpoise when the dynamic threshold was used (Fig. 7). Similarly, the false alarm rate calculated using the fixed threshold was 0.45 for delphinids and 0.07 for harbor porpoise. The false alarm rate calculated using the dynamic threshold changed from 0 to 0.99 for the delphinids and from 0 to 0.98 for the harbor porpoise. Although the correct detection rate was high, the false alarm rate was also high when β was below 10% in the delphinids, and the opposite for the harbor porpoise. In addition, both correct detection rate and false alarm rate were low when β was $>90\%$ in the delphinids and the opposite for the harbor porpoise. Between 26% and 82% of β , the dynamic threshold could identify both the delphinids and the harbor porpoise with over 0.80 of correct detection rate and below 0.20 of the false alarm rate.

The comparison of MCC between the dynamic and fixed threshold is shown in Fig. 8. The MCC that was calculated using the dynamic threshold was higher than that calculated using the fixed threshold between 10% and 95% of β .

The maximum detection range of the harbor porpoise was 524 m and that of the bottlenose dolphin was more than 1000 m, which was calculated using the same method as that used for the harbor porpoise.

IV. DISCUSSION

In this study, we proposed the simplest quantitative species identification method by using the two-band ratio (130 kHz/70 kHz). As shown in previous studies [acoustic characteristics of biosonar sounds listed in Au (1993) and Wartzok and Ketten (1999)], delphinids use broadband signals of similar order intensities at both 70 and 130 kHz. On the other hand, harbor porpoises use narrowband high-frequency signals that have peak intensity at around 130 kHz. These general characteristics are similar no matter where these species live; therefore, the comparisons between two-band ratios would be robust and can be performed using any calibrated hydrophone that is capable of differentiating the two frequencies. Even a single hydrophone with two different bandpass filters could be utilized for this purpose.

However, we have to take into consideration that delphinids also produce a variety of clicks of different spectrum shapes such as on- and off-axis and occasionally produce narrowband high-frequency signals during extreme cases (Au, 1993; Simon *et al.*, 2010). This is why the statistical evaluation of discrimination performance is needed, as described in this paper. Perfect discrimination of delphinids and the harbor porpoise is impossible because the distribution of the two-band ratio of delphinids partly overlapped with that of the harbor porpoise (Fig. 4). Quantitative discrimination performance should be described with detection probability.

We examined the detection performance of both the fixed and the dynamic thresholds, which depend on the mix ratio of recorded sounds from both families. The fixed threshold can be calculated as the simplest constant value of

the ratio of two families from the observed data. However, the proposed dynamic threshold showed a greater performance in discriminating families than that by using the fixed threshold (Figs. 6–8).

The classification performance was affected by the mix ratio of the detected number of click trains in each family in this study. The performance to discriminate each family was examined by using the ROC curve (Fig. 6). The ROC curve of family identification showed that the false alarm rate was lower and the correct detection rate was higher in the upper left area of the graph. Therefore, the upper left area can be considered as the optimal threshold. The result showed that the dynamic threshold (white markers) could identify the harbor porpoise and delphinids better than that by using the fixed threshold (black markers) even under the same mix ratio, which was visually confirmed as the ground truth distribution of two-band ratio in the present study. The white markers were closer together than the black markers in the upper left area in the ROC curves of the harbor porpoise and delphinids (Fig. 6). This ROC curve was not available without visual confirmation of the species. Determining the discrimination criteria without ground truth data was the arbitrary choice and was not quantitative in most cases.

By definition, the dynamic threshold changes according to the mix ratio (Fig. 9). The difference between the fixed threshold and the dynamic threshold was large when the mix ratio was strongly biased. This is the compensation to minimize false alarm by the dynamic threshold. Large numbers of false alarms could be detected using the fixed threshold because the left or right tail of the distribution could be included (Fig. 4). The dynamic threshold did not change smoothly due to the limited sample size of the ground truth data, especially in the harbor porpoise (Fig. 4).

The MCC value, which can take into account both correct detection and false alarm rates as one parameter to evaluate classification efficiency, of the dynamic threshold was higher than that of the fixed threshold (Fig. 8). This indicates that the dynamic threshold could distinguish each family more accurately than the fixed threshold for various mix ratios of the two families. The fixed or dynamic threshold is not applicable for extremely biased mix ratio of families (Figs. 7 and 8). The dynamic threshold could not be used when the mix ratio between numbers of click trains is over

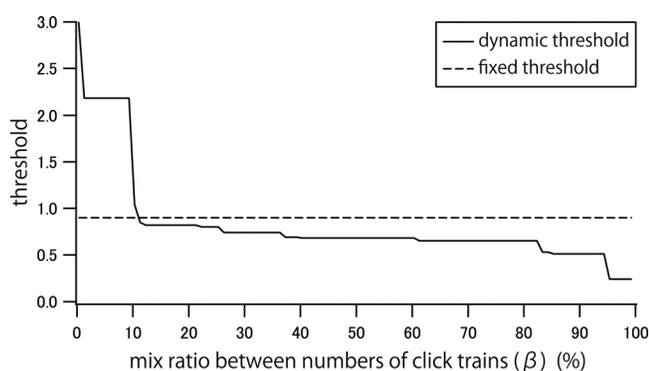


FIG. 9. The dynamic threshold (solid line) and fixed threshold (dotted line) in each mix ratio between numbers of click trains, which can also be called mix ratio β . The dynamic thresholds change depending on the mix ratio.

95% or under 10%. In the Istanbul Strait, the delphinids sightings were dominant or equal to those of the harbor porpoise for most of the year (Dede *et al.*, 2008; Öztürk *et al.*, 2009). Therefore, obtaining biased mix ratios such as over 95% or below 10% in a month would be rare.

The maximum detection ranges were consistent with those reported by a previous study, which was conducted using acoustic observation with an A-tag in the same area (Dede *et al.*, 2013). In this study, we extracted the clicks only after visual confirmation. Most of the extracted clicks were at a closer range than the maximum detection range and each click train showed clear acoustic characteristics. The long-range propagation leads to the loss of the higher frequency component. The power spectrum of the delphinids would be shifted to the lower frequency side for the sound that traveled for several hundred meters. Distance dependence of the performance was not examined due to the limited visually confirmed data of families at a long distance in the present study. The proposed procedure of quantitative classification using visually confirmed ground truth data might be applied to target sounds of various biological sources.

ACKNOWLEDGMENTS

We would like to express our thanks to all the members of Biosphere Informatics, Graduate School of Informatics, Kyoto University. We especially thank K. Moriya, T. Sakai, L. Koyama, and H. Mitamura, who helped our analyses, for their kind and clear advices. We also appreciate B. Öztürk for his help in the field. This study was partly supported by CoCoNET, a project funded under the Ocean of Tomorrow Joint Call 2011 of the European Union, the Core Research for Evolutional Science and Technology of JST and Development Program for New Bioindustry Initiatives, the Sasakawa Scientific Research Grant from The Japan Science Society (24-735), Grant-in-Aid for JSPS Fellows (25-6458), and the Turkish Marine Research Foundation.

- Akamatsu, T., Nakamura, K., Kawabe, R., Furukawa, S., Murata, H., Kawakubo, A., and Komaba, M. (2010). "Seasonal and diurnal presence of finless porpoises at a corridor to the ocean from their habitat," *Mar. Biol.* **157**, 1879–1887.
- Akamatsu, T., Teilmann, J., Miller, L. A., Tougaard, J., Dietz, R., Wang, D., Wang, K., Siebert, U., and Naito, Y. (2007). "Comparison of echolocation behaviour between coastal and riverine porpoises," *Deep Sea Res., Part II* **54**, 290–297.
- Akamatsu, T., Wang, D., Nakamura, K., and Wang, K. (1998). "Echolocation range of captive and free-ranging baiji (*Lipotes vexillifer*), finless porpoise (*Neophocaena phocaenoides*) and bottlenose dolphin (*Tursiops truncatus*)," *J. Acoust. Soc. Am.* **104**, 2511–2516.
- Alpar, B., and Yüce, H. (1998). "Sea-level variations and their interactions between the Black Sea and the Aegean Sea," *Est. Coast. Shelf Sci.* **46**, 609–619.
- Au, W. W. L. (1993). *The Sonar of Dolphins* (Springer, New York), pp. 1–277.
- Baldi, P., Brunak, S., Claus, Y. C., Andersen, A. F., and Nielsen, H. (2000). "Assessing the accuracy of prediction algorithms for classification: An overview," *Bioinformatics* **16**, 412–424.
- Barlow, J., and Taylor, B. (2005). "Estimates of sperm whale abundance in the northeastern temperate Pacific from a combined acoustic and visual survey," *Mar. Mammal Sci.* **21**, 429–445.
- Croll, D. A., Clark, C., Acevedo, A., Tershey, B., Flores, S., Gedamke, J., and Urban, J. (2002). "Only male fin whales sing loud songs," *Nature (London)* **417**, 809.

- Dahne, M., Verfuss, U. K., Brandecker, A., Siebert, U., and Benke, H. (2013). "Methodology and results of calibration of tonal click detectors for small odontocetes (C-PODs)," *J. Acoust. Soc. Am.* **134**, 2514–2522.
- Darling, J. D., and Berube, M. (2001). "Interaction of singing humpback whales with other males," *Mar. Mammal Sci.* **17**, 570–584.
- Dede, A., Öztürk, A. A., Akamatsu, T., Tonay, A. M., and Öztürk, B. (2013). "Long-term passive acoustic monitoring revealed seasonal and diurnal presence patterns of cetaceans in the Istanbul Strait," *J. Mar. Biol. Ass. U.K.* DOI:10.1017/S0025315413000568.
- Dede, A., Öztürk, A. A., and Tonay, A. M. (2008). "Cetacean surveys in the Istanbul (Bosphorus) Strait in 2006," in *Proceedings of 22nd Annual Conference of the European Cetacean Society*, Egmond aan Zee, the Netherlands, pp. 235–236.
- Deecke, V. B., Ford, J. K. B., and Spong, P. (2000). "Dialect change in resident killer whales: Implications for vocal learning and cultural transmission," *Anim. Behav.* **60**, 629–638.
- Francois, R. E., and Garrison, G. R. (1982a). "Sound absorption based on ocean measurements: Part I: Pure water and magnesium sulfate contributions," *J. Acoust. Soc. Am.* **72**, 896–907.
- Francois, R. E., and Garrison, G. R. (1982b). "Sound absorption based on ocean measurements: Part II: Boric acid contribution and equation for total absorption," *J. Acoust. Soc. Am.* **72**, 1879–1890.
- Gavrilov, A. N., McCauley, R. D., and Gedamke, J. (2012). "Steady inter and intra-annual decrease in the vocalization frequency of Antarctic blue whales," *J. Acoust. Soc. Am.* **131**, 4476–4480.
- Kimura, S., Akamatsu, T., Li, S., Dong, S., Dong, L., Wang, K., Wang, D., and Arai, N. (2010). "Density estimation of Yangtze finless porpoises using passive acoustic sensors and automated click train detection," *J. Acoust. Soc. Am.* **128**, 1435–1445.
- Kyhn, L., Tougaard, J., Teilmann, J., Wahlberg, M., Jørgensen, P. B., and Bech, N. I. (2008). "Harbour porpoise (*Phocoena phocoena*) static acoustic monitoring: Laboratory detection thresholds of T-PODs are reflected in field sensitivity," *J. Mar. Biol. Ass. U.K.* **88**, 1085–1091.
- Li, S., Akamatsu, T., Wang, D., and Wang, K. (2009). "Localization and tracking of phonating finless porpoises using towed stereo acoustic dataloggers," *J. Acoust. Soc. Am.* **126**, 468–475.
- Madsen, P. T., Carder, D. A., Bedholm, K., and Ridgway, S. H. (2005). "Porpoise clicks from a sperm whale nose—Convergent evolution of 130 kHz pulses in toothed whale sonars?," *Bioacoustics* **15**, 195–206.
- Matthews, B. W. (1975). "Comparison of the predicted and observed secondary structure of T4 phage lysozyme," *Biochim. Biophys. Acta* **405**, 442–451.
- McDonald, M. A., Hildebrand, J. A., and Mesnick, S. L. (2006). "Biographic characterization of blue whale song worldwide: Using song to identify populations," *J. Cetacean Res. Manage.* **8**, 55–65.
- Mellinger, D. K., Stafford, K. M., Moore, S. E., Dziak, R. P., and Matsumoto, H. (2007). "An overview of fixed passive acoustic observation methods for cetacean," *Oceanogr.* **20**, 36–45.
- Oleson, E. M., Calambokidis, J., Burgess, W. C., McDonald, M. A., LeDuc, C. A., and Hildebrand, J. A. (2007). "Behavioral context of call production by eastern North Pacific blue whales," *Mar. Ecol. Prog. Ser.* **330**, 269–284.
- Oswald, J. N., Rankin, S., Barlow, J., and Lammers, M. O. (2007). "A tool for real-time acoustic species identification of delphinid whistles," *J. Acoust. Soc. Am.* **122**, 587–595.
- Öztürk, A. A., Dede, A., Tonay, M. A., and Öztürk, B. (2009). "Cetacean surveys in the Istanbul (Bosphorus) Strait in 2007–2008," in *Proceedings of 23rd Annual Conference of the European Cetacean Society*, Istanbul, Turkey, p. 29.
- Payne, R. S., and McVay, S. (1971). "Songs of humpback whales," *Science* **173**, 585–597.
- Philpott, E., Englund, A., Ingram, S., and Rogan, E. (2007). "Using T-PODs to investigate the echolocation of coastal bottlenose dolphins," *J. Mar. Biol. Ass. U.K.* **87**, 11–17.
- Quick, N. L., and Janik, V. M. (2012). "Bottlenose dolphins exchange signature whistles when meeting at sea," *Proc. R. Soc. London, Ser. B.* **279**, 2539–2545.
- Sayigh, L. S., Tyack, P. L., Wells, R. S., and Scott, M. D. (1990). "Signature whistles of free-ranging bottlenose dolphins *Tursiops truncatus*: stability and mother-offspring comparisons," *Behav. Ecol. Sociobiol.* **26**, 247–260.
- Simard, P., Hibbard, A. L., McCallister, K. A., Frankel, A. S., Zeddies, D. G., Sisson, G. M., Gowans, S., Forsys, E. A., and Mann, D. A. (2010). "Depth dependent variation of the echolocation pulse rate of bottlenose dolphins (*Tursiops truncatus*)," *J. Acoust. Soc. Am.* **127**, 568–578.
- Simon, M., Nuutila, H., Reyes-Zamudio, M. M., Ugarte, F., Verfuß, U. K., and Evans, P. G. H. (2010). "Passive acoustic monitoring of bottlenose dolphin and harbour porpoise, in Cardigan Bay, Wales, with implications for habitat use and partitioning," *J. Mar. Biol. Assoc. U. K.* **90**, 1539–1545.
- Soldevilla, M. S., Henderson, E. E., Campbell, G. S., Wiggins, S. M., Hildebrand, J. A., and Roch, M. A. (2008). "Classification of Risso's and Pacific white-sided dolphins using spectral properties of echolocation clicks," *J. Acoust. Soc. Am.* **124**, 609–624.
- Sousa-Lima, R. S., Norris, T. F., Oswald, J. N., and Fernandes, F. P. (2013). "A review and inventory of fixed autonomous recorders for passive acoustic monitoring of marine mammals," *Aquat. Mamm.* **39**, 23–53.
- Todd, V. L. G., Pearse, W. D., Tregenza, N. C., Lepper, P. A., and Todd, I. B. (2009). "Diel echolocation activity of harbour porpoises (*Phocoena phocoena*) around North Sea offshore gas installations," *ICES J. Mar. Sci.* **66**, 734–745.
- Verfuss, U. K., Dahne, M., Gallus, A., Jabbusch, M., and Benke, H. (2013). "Determining the detection thresholds for harbor porpoise clicks of autonomous data loggers, the Timing Porpoise Detectors," *J. Acoust. Soc. Am.* **134**, 2462–2468.
- Verfuss, U. K., Honnef, C. G., Meding, A., Dahne, M., Mundry, R., and Benke, H. (2007). "Geographical and seasonal variation of harbour porpoise (*Phocoena phocoena*) presence in the German Baltic Sea revealed by passive acoustic monitoring," *J. Mar. Biol. Ass. U. K.* **87**, 165–176.
- Villadsgaard, A., Wahlberg, M., and Tougaard, J. (2007). "Echolocation signals of wild harbor porpoises, *Phocoena phocoena*," *J. Exp. Biol.* **210**, 56–64.
- Wartzok, D., and Ketten, D. R. (1999). "Marine mammal sensory systems," *Biology of Marine Mammals*, edited by J. E. Reynolds III and S. A. Rommel (Smithsonian Institution Press, Washington), pp. 117–175.