

## Article

# A Case Study of Whistle Detection and Localization for Humpback Dolphins in Taiwan

Ching-Tang Hung , Wei-Yen Chu, Wei-Lun Li, Yen-Hsiang Huang, Wei-Chun Hu and Chi-Fang Chen \*

Department of Engineering Science and Ocean Engineer, National Taiwan University, Taipei City 10216, Taiwan; d08525005@ntu.edu.tw (C.-T.H.); r06525012@ntu.edu.tw (W.-Y.C.); r05525004@ntu.edu.tw (W.-L.L.); r05525002@ntu.edu.tw (Y.-H.H.); d03525001@ntu.edu.tw (W.-C.H.)

\* Correspondence: chifang@ntu.edu.tw; Tel.: +886-2-3355-5735

**Abstract:** In recent years, Taiwan's government has focused on policies regarding offshore wind farming near the Indo-Pacific humpback dolphin habitat, where marine mammal observation is a critical consideration. The present research developed an algorithm called National Taiwan University Passive Acoustic Monitoring (NTU\_PAM) to assist marine mammal observers (MMOs). The algorithm performs whistle detection processing and whistle localization. Whistle detection processing is based on image processing and whistle feature extraction; whistle localization is based on the time difference of arrival (TDOA) method. To test the whistle detection performance, we used the same data to compare NTU\_PAM and the widely used software PAMGuard. To test whistle localization, we designed a real field experiment where a sound source projected simulated whistles, which were then recorded by several hydrophone stations. The data were analyzed to locate the moving path of the source. The results show that localization accuracy was higher when the sound source position was in the detection region composed of hydrophone stations. This paper provides a method for MMOs to conveniently observe the migration path and population dynamics of cetaceans without ecological disturbance.

**Keywords:** marine mammal; whistle detection; time difference of arrival; underwater acoustic; underwater sound sensing; ocean sound measurement



**Citation:** Hung, C.-T.; Chu, W.-Y.; Li, W.-L.; Huang, Y.-H.; Hu, W.-C.; Chen, C.-F. A Case Study of Whistle Detection and Localization for Humpback Dolphins in Taiwan. *J. Mar. Sci. Eng.* **2021**, *9*, 725. <https://doi.org/10.3390/jmse9070725>

Academic Editors: Michel André, Christine Erbe and Giuseppa Buscaino

Received: 30 April 2021  
Accepted: 28 June 2021  
Published: 30 June 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Currently, most of Taiwan's raw materials for energy production, including coking coal, fuel coal, crude, and liquefied natural gas [1], are imported and have a large and immediate impact on the environment. Therefore, the government has actively developed green energy, including offshore wind farms [2], but most sites overlap with Indo-Pacific humpback dolphin reservation zones. The noise from pile driving during construction may impact marine mammals and cause auditory injury, ranging from temporary threshold shift (TTS) to permanent threshold shift (PTS) in hearing [3]. To minimize the noise-induced impact on cetaceans caused by construction and the operation of wind turbines, establishing a marine mammal detection mechanism is a priority. The traditional method to detect cetaceans is visual, whereby marine mammal Observers (MMOs) work from vehicles, using the naked eye to search for cetaceans, an operation that is expensive and offers only a low probability of success; moreover, it is limited to daylight hours. Underwater acoustics provide an alternative technique to detect marine mammals, and the cetacean call can be used as a specific characteristic of detection. We used passive acoustic monitoring (PAM) to develop an algorithm and NTU\_PAM to monitor cetacean calls followed by motion tracking. In addition to overcoming the weaknesses of the visual method, NTU\_PAM can show the correlation between the results of the visual method and PAM.

Cetaceans produce two major types of cetacean calls [4,5]: (1) the “whistle” is a continuous, narrow-band, and frequency-modulated signal that is thought to be a form of social communication; (2) the “click” is considered a bio-sonar and is a short, broadband, and directional impulse signal used to navigate, detect, and identify objects. In marine mammal research, PAM has proved a useful tool. For example, (1) Spaulding et al. [6] built a near-real-time buoy system to automatically detect North Atlantic right whale calls in Cape Cod Bay and near the Boston Harbor. When the buoy system detects a whale call, an alarm signal is transmitted and the call is recorded. (2) Linnenschmidt et al. [7] equipped an acoustic data logger on a porpoise to record clicks and determine the relationship among the click, movement, and diving behavior. (3) Akamatsu et al. [8] used an underwater pulse event recorder (A-TAG) to record clicks and analyze critical parameters such as interclick interval (ICI).

Previous studies of cetacean whistle detection have been vigorous. Gannier et al. [9] developed Seafox software to extract whistle characteristics (length, beginning frequency, ending frequency, maximum frequency, minimum frequency, etc.) on a time spectrogram and used a regression tree to classify five dolphin species. Lai [10] used the mel-frequency cepstral coefficient to simulate human auditory features, namely the critical band and auditory masking, and to extract the characteristics of the whistle. The whistle characteristics were then used in a support vector machine (SVM) to identify the cetacean species. Caldwell and Caldwell [11–13] hypothesized that signature whistle variations, which dolphins emit and which carry information, are required so distinctive whistles can be used to identify individual dolphins. Datta and Sturtivant [14] considered two whistle features on a spectrogram (overall contour shape and detailed contour structure difference) as parameters of the signature whistle and grouped whistles using the hidden Markov model (HMM) method. Bahoura and Simard [15] used an artificial neural network to classify blue whale calls. The above research is based on supervised machine learning methods requiring numerous sets of clean training data, manually labeling the calls, and building the model. These models are only suitable for specific or regional species.

To avoid the disadvantages of labeling, training, and specific targeting, Gillespie et al. [16] developed a whistle detector based on image processing on a spectrogram which is implemented as the Whistle and Moan Detector module in PAMGuard. PAMGuard software includes a user-friendly, human-machine interface and modules for data processing and marine mammal detection [17] and has been widely used for real-time marine mammal monitoring. Lin [18] devised a non-targeted algorithm on the MATLAB platform that helps users grasp the position of whistles across many audio files, making further processing convenient. Lin et al. [19] first denoised the spectrogram and then detected the whistle characteristics. Gillespie’s and Lin’s methods include four main steps: (1) spectrogram, (2) image processing, (3) whistle feature extraction, and (4) combination of the whistle data points. We applied the same pattern to develop the whistle detection algorithm. A similar concept is applied in steps 2–4, but the detailed methods are different. We also compared NTU\_PAM and PAMGuard, which is regarded as a standard of whistle detection.

Tracking cetaceans is another recent primary research subject. Janik et al. [20] deployed three hydrophones to form a two-dimensional, triangular array in Beaulieu Firth, Northern Scotland, U.K. The interhydrophone distances were 208, 513, and 506 m. An artificial sound was then projected at a depth of 1 m. The time difference of signal arrival for each pair of hydrophones became markers to conduct localization of a sound source. Wang et al. [21] deployed a two-dimensional, cross-shaped array consisting of five hydrophones from the side of the boat at a depth of 1 m in Pearl River Estuary, China, and Beibu Gulf of Guangxi, China. The inter-hydrophone distances were 1.47, 1.54, 2.08, and 2.18 m. The boat followed

the dolphin group at a close distance to receive the dolphin call, and they used the difference in arrival time of a sound at each hydrophone pair to localize the targets. Wiggins et al. [22] deployed a tracking high-frequency acoustic recording package (HARP) [23] consisting of four hydrophones at 3 m above the seafloor offshore of Southern California to track beached whales and dolphins. Wiggins et al. [24] also deployed four HARP's offshore of Southern California to track whistling dolphins. Both of Wiggins's methods used the TDOA method. Building on the demonstrated effectiveness of TDOA for tracking and localization, we utilized four hydrophone stations to form a kilometer-scale array for tracking the source based on TDOA.

We designed an experiment that simulated different whistle types in the real field and developed four PAM stations to track the artificial source. Four stations were deployed near Taichung Harbor to record the simulated calls. After processing the detected algorithm, finding the whistle time, and tracking the source, we compared the results from the algorithm and the moving path of the boat carrying the sound source. In this study, we developed an algorithm that does not require a trained model for the automatic detection of the whistle. The algorithm is based on the time length and frequency band of the whistle feature. Furthermore, the automatic detection algorithm and localization method were combined as NTU\_PAM. NTU\_PAM can work as an auxiliary tool for MMO during the daytime, and it can function as the main monitoring tool at night.

## 2. Whistle Detector Algorithm

Passive acoustic monitoring has been used widely in marine monitoring to amass longitudinal data and requires high-efficiency algorithms to assist researchers in finding the required file segments. We developed a whistle detector algorithm, which was then improved according to Li's prototype algorithm [25]. The algorithm can detect any creature producing a whistle and the whistle's detected frequency range, depending on the species. There are six main processes in the algorithm:

1. Transfer time-series data to a spectrogram by short-time Fourier transform (STFT);
2. Remove the noise on the time axis of the spectrogram;
3. Remove the salt and pepper noise in the spectrogram;
4. Find the data point that satisfies the condition of the power spectral density (PSD) and signal-to-noise ratio (SNR);
5. Extract data points using the features of whistles;
6. Cluster data points into different whistles.

A flow chart of the algorithm is shown in Figure 1. In order to present whistles clearly on the spectrogram, some processes are based on image processing. Each process will be described in detail. Figure 2 shows each step of the results.

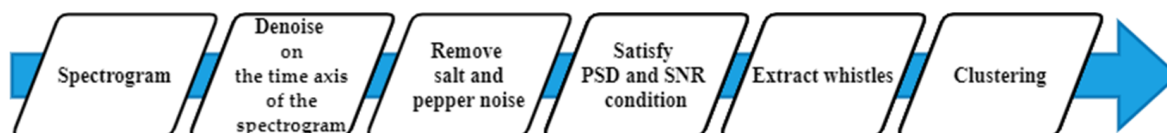
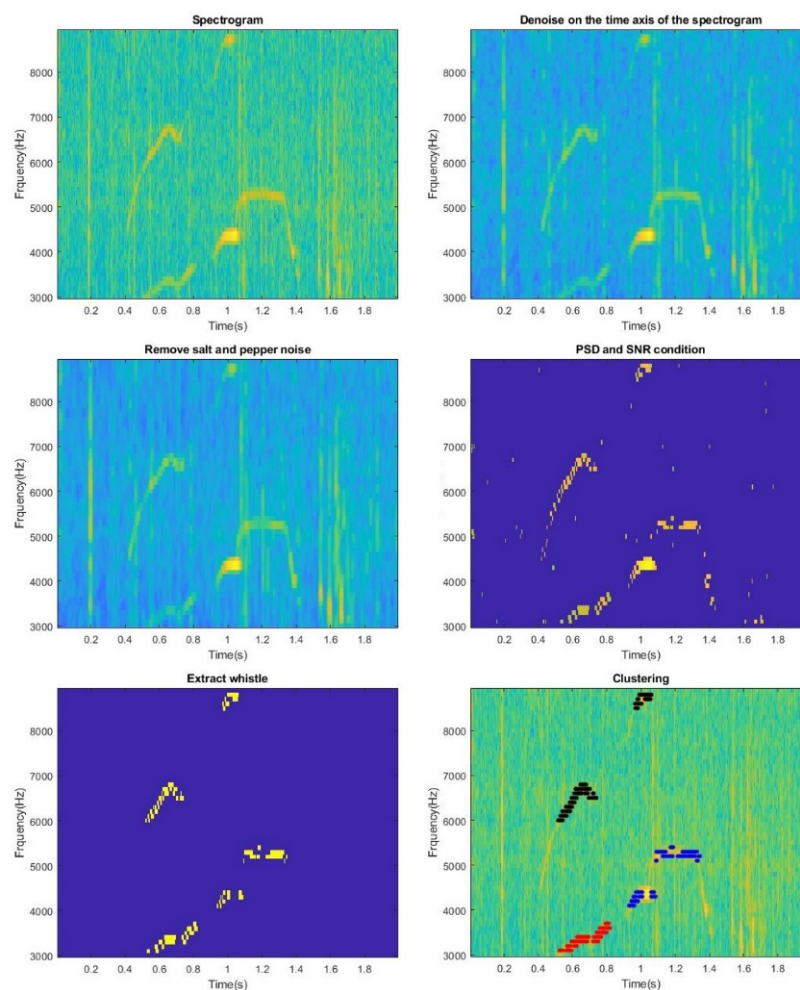


Figure 1. Whistle detector algorithm flow chart.



**Figure 2.** Each step of the results.

### 2.1. Spectrogram

We used the STFT [26], which adds a window function to obtain the frequency domain information changed by the time domain. This establishes a frame to slide on the time domain signal and extracts the signal in the frame, which convolves with the window function to perform the Fourier transform. This information is used to produce the spectrogram. The window function is the Hamming window [27], the frame length is 0.01 s, and the overlap is 90%. The STFT formula is shown in Equation (1), where  $w(t)$  is window function and  $x(t)$  is raw data.

$$S_{t,f} = \int_{-\infty}^{+\infty} w(t - \tau)x(\tau)e^{-jft}d\tau \quad (1)$$

### 2.2. Denoising on the Time Axis of the Spectrogram

Whistle length is long compared to impulse noise; therefore, we use the moving average method to remove impulse noise on the spectrogram. Every 20 points on the time axis of each single frequency band are averaged to build a new spectrogram; the formula is shown in Equation (2), where  $S_{t,f}$  is the original spectrogram and  $S'_{t,f}$  is the new spectrogram after denoising.

$$S'_{t,f} = \frac{1}{20} \sum_{n=0}^{19} S_{t-n,f} \quad (2)$$



### 2.3. Removing Salt and Pepper Noise

A median filter, often used in image processing and a technique for nonlinear signal processing, was used to remove salt and pepper noise [28]. The median of every 3-by-3 matrix on the spectrogram is calculated. The formula is shown in Equation (3), where  $S'_{t,f}$  is the spectrogram after the denoising and  $S''_{t,f}$  is the spectrogram after using the median filter.

$$S''_{t,f} = \text{median}(S'_{t+i,f+j}); i, j = -1, 0, 1 \quad (3)$$

### 2.4. Satisfying PSD and SNR Conditions

Since a whistle is a narrow frequency band signal, with the occurrence of a whistle, its PSD is much larger than that of the point whose frequency is very close to the whistle. The definition of SNR in this study is shown in Equation (4). If the PSD is larger than the PSD threshold and the SNR is larger than the SNR threshold simultaneously at a data point, the data point will be replaced by one. If this is not the case, the data point will be replaced by zero. The formula is shown in Equation (5). The new spectrogram  $B_{t,f}$  is a binary image. The default value of the SNR threshold and the PSD threshold are 6 dB and 40 dB (re 1  $\mu\text{Pa}^2/\text{Hz}$ ), respectively.

$$\text{SNR} = \frac{2S''_{t,f}}{(S''_{t,f+1} + S''_{t,f-1})} \quad (4)$$

$$B_{t,f} = \begin{cases} 1, & \text{SNR}_{t,f} \geq \text{SNR}_{\text{threshold}} \ \& \ S''_{t,f} \geq \text{PSD}_{\text{threshold}} \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

### 2.5. Extracting the Whistle

As mentioned in Section 2.4, the whistle is a narrow frequency band and a continuous signal. In this method, the nearby data points whose value is one are connected and labeled as a segment. Next, two conditions are set: the frequency bandwidth threshold and the time length threshold. Lastly, the segments whose frequency bandwidth is smaller than the frequency bandwidth threshold and whose time length is longer than the time length threshold are retained. The binary image  $B_{t,f}$  will be refreshed as a new image  $B'_{t,f}$ . The default values of frequency bandwidth threshold and time length threshold are 300 Hz and 0.06 seconds, respectively.

### 2.6. Clustering

The k-means method [29] is used to cluster the data points in  $B'_{t,f}$ . According to the difference of frequency and time, some of the whistle segments from Section 2.5 and above are merged. If the time interval of two segments is smaller than 0.3 seconds and the difference of frequency between two segments is smaller than 1 kHz simultaneously, two segments will be considered as one whistle segment. After merging, the k (number of clusters) is decided by the new number of segments. Each data point automatically combines into k whistles by calculating Euclidean distance of frequency and time index in  $B'_{t,f}$ . Each whistle's start time, end time, start frequency, and end frequency are recorded after k-means.

## 3. Localization Method

TDOA was used to track the whistle. We devised an experiment to track the moving path of the artificial source by a whistle detector algorithm and TDOA.

### 3.1. Time Difference of Arrival (TDOA)

TDOA is often used in signal source positioning [30]. It only requires the received signal time and the speed that the signal travels. Once the signal is received at the two

receiving stations, the difference in arrival time can be used to draw the hyperbola of possible location by the equation shown in Equations (6) and (7). If we have three receiving stations, least two hyperbolas are produced, as shown in Figure 3, and their intersection will be the signal source location. To realize this hypothesis, the receiving stations must be time-synchronized.

$$\sqrt{(x-x_1)^2 + (y-y_1)^2} - \sqrt{(x-x_3)^2 + (y-y_3)^2} = c(t_1 - t_3) \quad (6)$$

$$\sqrt{(x-x_2)^2 + (y-y_2)^2} - \sqrt{(x-x_3)^2 + (y-y_3)^2} = c(t_2 - t_3) \quad (7)$$

where  $t_1$ ,  $t_2$ , and  $t_3$  are the times when the same signal arrives at different hydrophones;  $(x, y)$  is the position of the unknown signal source; and  $c$  is the sound speed from the local sound speed profile.

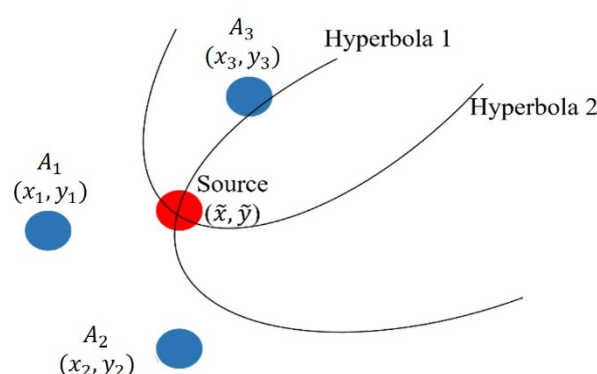


Figure 3. TDOA schematic.

### 3.2. Taichung Harbor TDOA Experimental Configuration

We deployed four hydrophone stations near Taichung Harbor, an area where Indo-Pacific humpback dolphins are extremely active [31,32]. The locations of the hydrophones are shown in Figure 4, and the exact latitude and longitude are shown in Table 1. The Beaufort Sea state was below 3, and the ambient noise is illustrated in Figure 5 as a percentile level. The highest PSD was around 95 dB (re  $1 \mu\text{Pa}^2/\text{Hz}$ ) from 60–70 Hz on L50, possibly produced by shipping noise, and the PSD from 3 kHz–10 kHz was around 65 dB (re  $1 \mu\text{Pa}^2/\text{Hz}$ ).

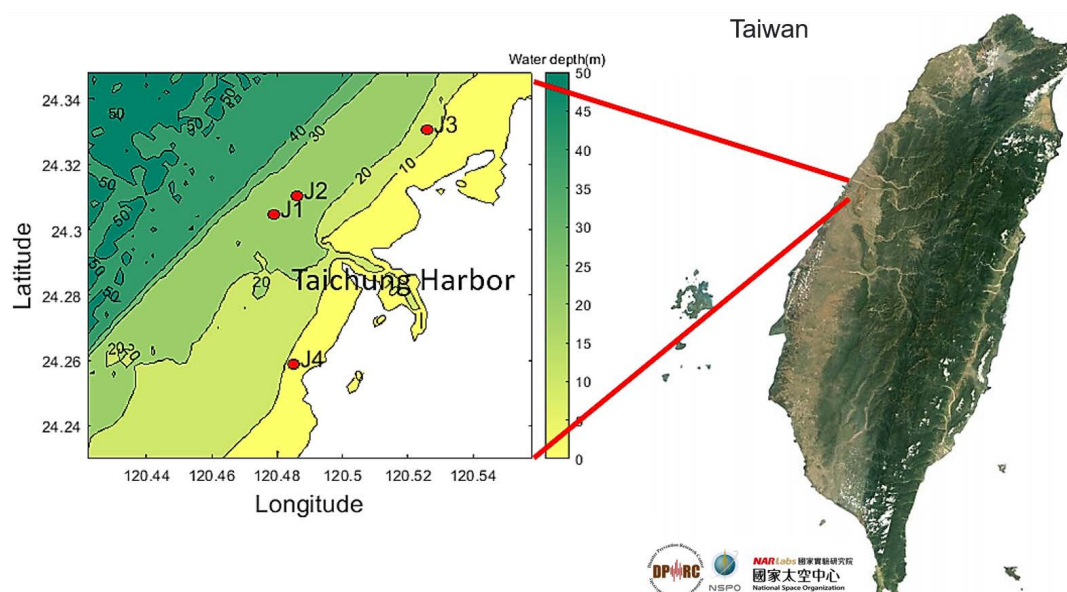
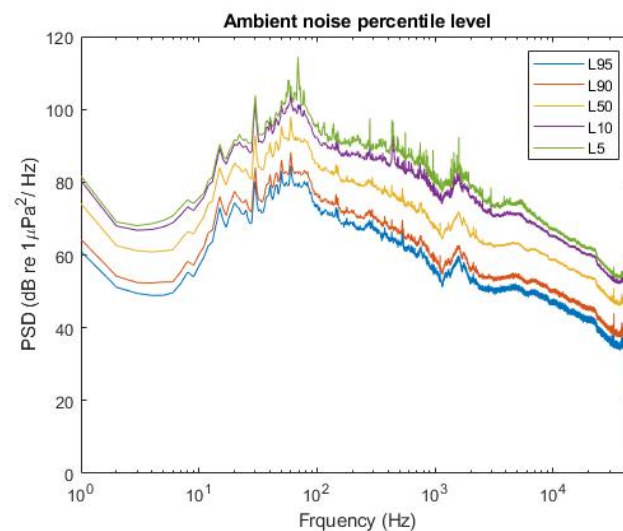


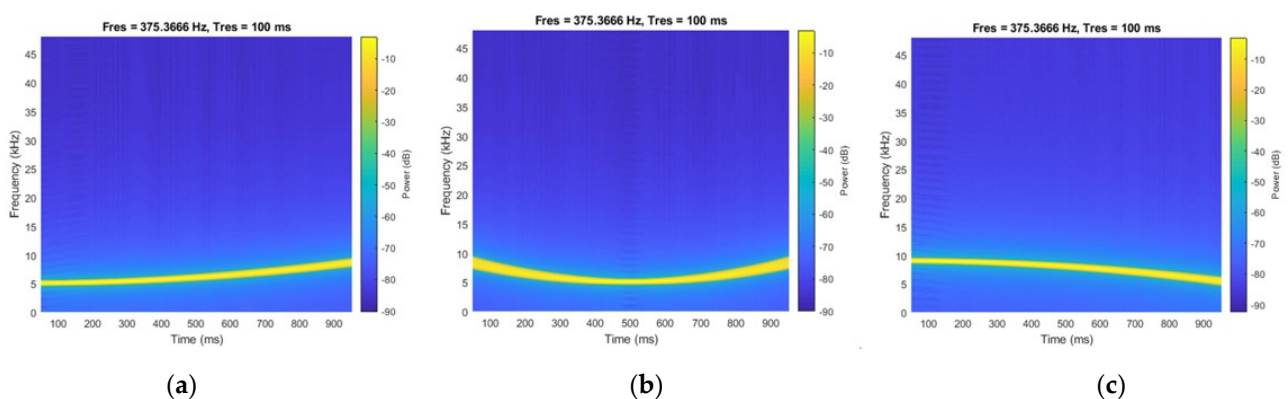
Figure 4. Hydrophone station locations.

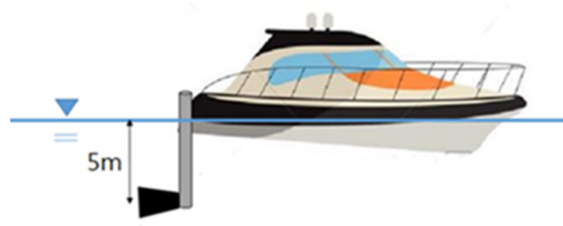
**Table 1.** Latitude and longitude of hydrophone stations.

Station	Latitude (N)	Longitude (E)	Depth (m)
J1	24.3305°	120.4788°	29.1
J2	24.3101°	120.4861°	28.7
J3	24.3305°	120.5259°	8.0
J4	24.2588°	120.4851°	11.0

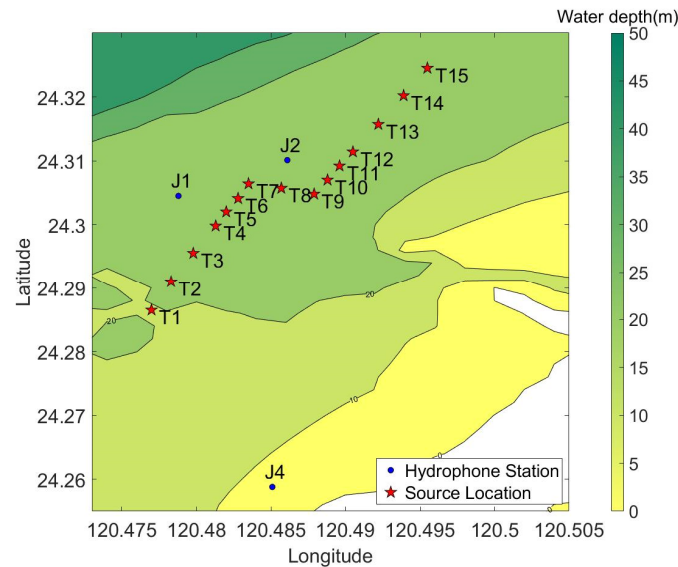
**Figure 5.** Ambient noise percentile level:  $L_n$  is the noise level exceeding  $n\%$  of the measurement time, i.e.,  $L_{50}$  is the noise level exceeding 50% of the measurement time.

The SoundTrap ST500 hydrophone recorder was used at point J3, and three Wildlife Acoustics SM3M hydrophone recorders were used at points J1, J2, and J4. They were deployed using the bottom-mounted method with sampling frequency set to 96 kHz. To achieve time synchronization for all recorders, we produced an impulse signal as a benchmark for correcting the time before deploying. To simulate the whistle of an actual Indo-Pacific humpback dolphin, which features a frequency range of 3–9 kHz, three kinds of artificial sound signals were employed: (a) rising frequency (5–9 kHz), (b) U-type (9–5–9 kHz), and (c) decreasing frequency (9–5 kHz), with a time length of one second, as shown in Figure 6. The source level (SL) was 160 dB (re 1  $\mu$  Pa at 1 m). The underwater acoustic projector SQS-23 was placed at a water depth of 5 m (Figure 7), since Indo-Pacific humpback dolphins often stay about 5 m below sea level [33]. Figure 8 shows where the artificial sound signals were played, every 10 seconds for 10 minutes, in the 15 spots (T1–T15) outside Taichung Harbor.

**Figure 6.** (a) Rising frequency; (b) U-type; (c) decreasing frequency.



**Figure 7.** Schematic of the installation position of the projector.



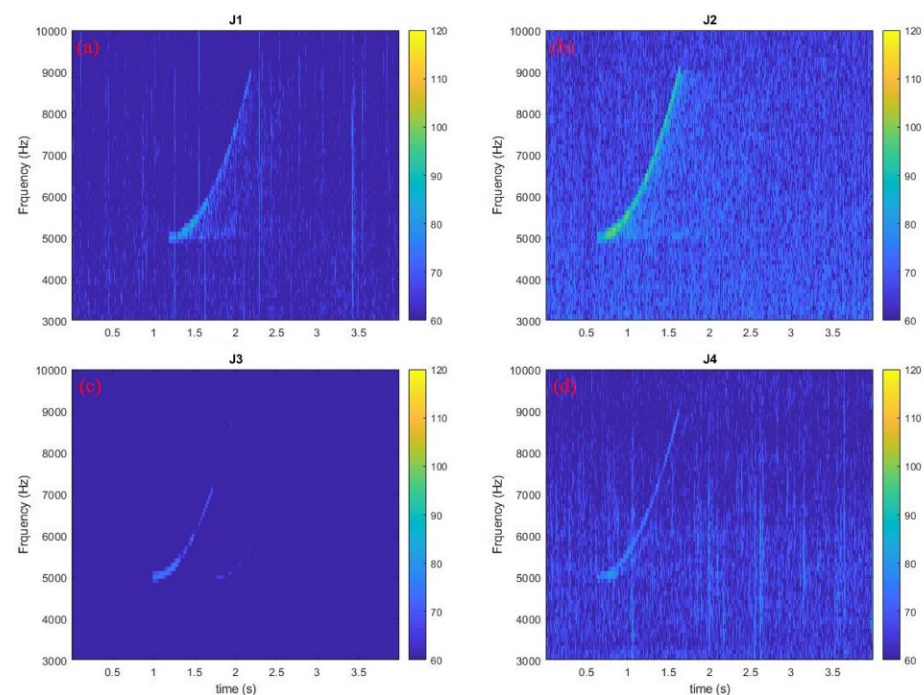
**Figure 8.** Locations of the signal sources.

### 3.3. Experimental Data Analysis Method

In this experiment, the SNR of the received signal was larger than 10 dB, exceeding the NTU\_PAM-recommended SNR threshold of 6 dB. The signals recorded by each of the hydrophones at the four stations when the source was at point T10 are shown in Figure 9. To find the artificial whistle within the sound file, NTU\_PAM was used to extract information, namely the start and end times from the raw data of the four hydrophones. However, the extracted time information was not precise enough for TDOA. For increased accuracy, the raw data of the start and end times of the whistle were directly analyzed without being processed by the algorithm. The time of the J2 station was considered as the central time, and cross-correlation analysis with the full frequency band raw data of the central station and three other stations was performed to determine the time difference, as shown in Equations (8) and (9), where  $X_2$  is J2 station's whistle raw data;  $X_o$  is the three other stations' whistle raw data;  $R$  is the result of cross-correlation; and  $td$  is the time difference, which was used to obtain the location of the signal source by the TDOA method.

$$R(\tau) = \int X_2(t)X_o(t - \tau)dt \quad (8)$$

$$td = \max(R(\tau)) \quad (9)$$



**Figure 9.** Received signal of each station when the source was at point T10. (a) J1 station; (b) J2 station; (c) J3 station; (d) J4 station.

## 4. Results

### 4.1. Comparison with PAMGuard

As mentioned, PAMGuard software is widely used in the field of marine mammal observation. In this research, the performance of NTU\_PAM and the Whistle and Moan Detector module of PAMGuard were compared using the same hardware (an i9-9900 CPU from Intel Corporation with 64 GB of memory). The test audio is a two-minute sound file, rich in whistles and with a sampling frequency of 96 kHz, recorded near the sea area of Yunlin, Taiwan [34]. We manually confirmed that the file contained a total of 33 whistles.

When the PAMGuard Whistle and Moan Detector's parameters were set at a window length of 2048 data points (0.02 s) and 1024 data points (0.01 s), and when the overlap ratios were 50% and 90%, the NTU\_PAM's recommended window length was 0.01 s with an overlap ratio of 90% and SNR set to 6 dB. As shown in Table 2, PAMGuard with settings of window length at 1024 data points, 90% overlap ratio, and 6 dB SNR shows the closest result of the 47 detected whistles to the manually confirmed 33 whistles. A total of 30 whistles were detected by NTU\_PAM.

**Table 2.** Comparison of results.

Algorithm	Parameters	Detected Numbers
Manually confirmed	-	33
PAMGuard	Window = 2048 Overlap = 50% SNR = 6 dB	79
PAMGuard	Window = 2048 Overlap = 90% SNR = 6 dB	50



Table 2. Cont.

Algorithm	Parameters	Detected Numbers
PAMGuard	Window = 1024 Overlap = 50% SNR = 6 dB	91
PAMGuard	Window = 1024 Overlap = 90% SNR = 6 dB	47
NTU_PAM	Window = 0.01 s Overlap = 90% SNR = 6 dB	30

#### 4.2. Experimental Results

At least three signal receiving stations were used to calculate TDOA. When the intersection of the hyperbolic curves is plural, the center point is taken as the final judgment location. To verify localization accuracy, GPS data from the experimental ship bearing the sound source were compared to results from TDOA.

In the series of graphs in Figure 10, the blue dot is the hydrophone station position (J1, J2, and J4), the red dot is the signal source position of the experimental ship's GPS record, and the yellow star is the TDOA positioning result. The results from the first experiment testing the rising frequency (5–9 kHz) signal are shown in Figure 10a. The positioning accuracy was higher when the sound source was nearer to the center positions J1 and J2 from the group of hydrophone stations. The nearest positioning points T4 to T11 showed an average positioning error of 24.7 m, and the overall positioning error was 143.5 m, which was affected by the lower accuracy of the outer point.

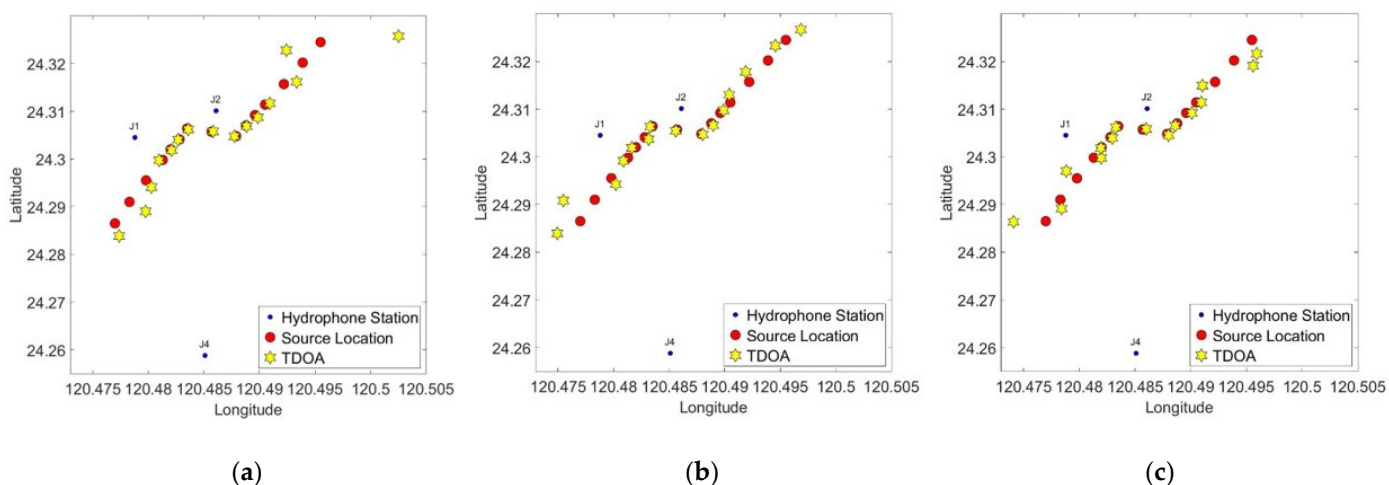
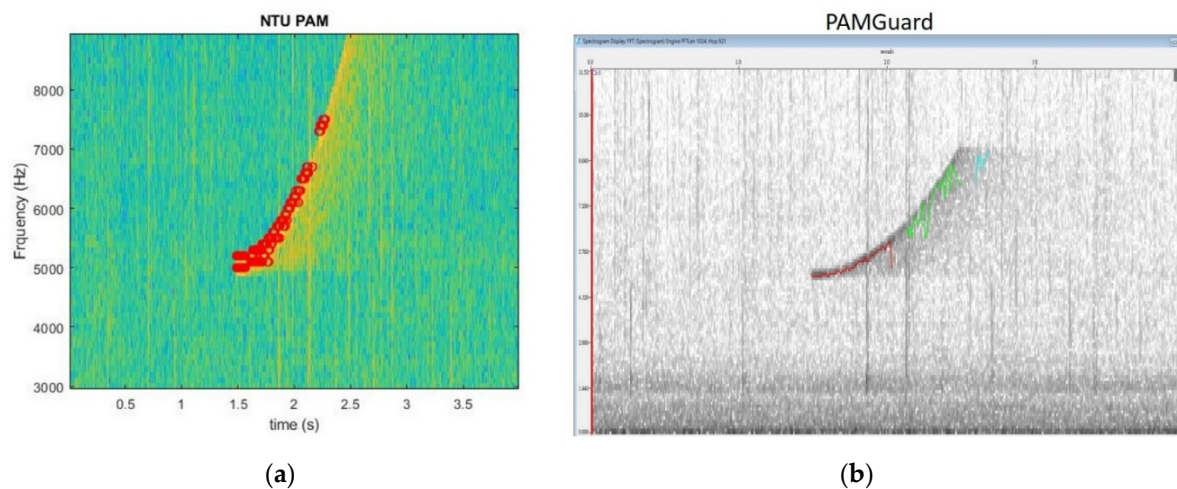


Figure 10. (a) Result of rising frequency signal; (b) result of decreasing frequency signal; (c) result of U-shaped signal.

The second experiment was the decreasing frequency (9–5 kHz) signal, and its positioning trend was similar to the rising frequency signal (Figure 10b). It also showed higher positioning accuracy when the signal source was close to the J1 and J2 stations. The average positioning error of T4 to T11 was 44.8 m, larger than that of the rising frequency signal, and the overall positioning error was 145.9 m. Finally, the U-shaped (9–5–9 kHz) signal displayed a similar trend as the aforementioned signals (Figure 10c). The average positioning error of T4 to T11 was 39.6 m, but the overall positioning error was the smallest of the three signals at 116.1 m.

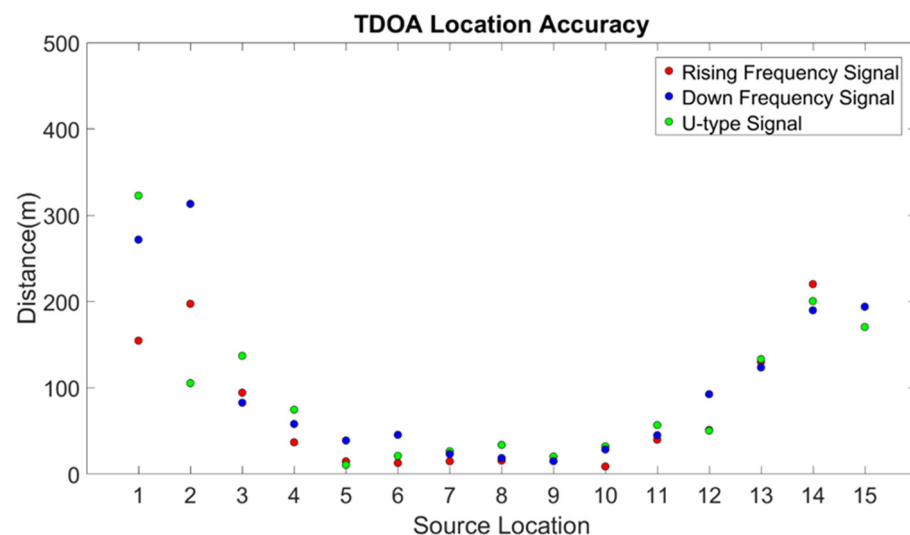
## 5. Discussion

In the comparison between PAMGuard and NTU\_PAM, the results were close to the number of whistles that was manually confirmed and showed that both performed well on whistle detection. The reason for the different numbers detected may be that PAMGuard is a real-time auxiliary tool mainly provided to visual method researchers for detecting the occurrence of a call; as such, it only needs a few window lengths of data to detect the whistle. As to the amount of audio data required, NTU\_PAM needs one second or more of data to build a spectrogram and to initiate processing. However, PAMGuard may, at times, break one call into several calls, as shown in Figure 11. According to the results, NTU\_PAM is suitable for processing measurements captured over a longer duration, and it proves as robust as PAMGuard.

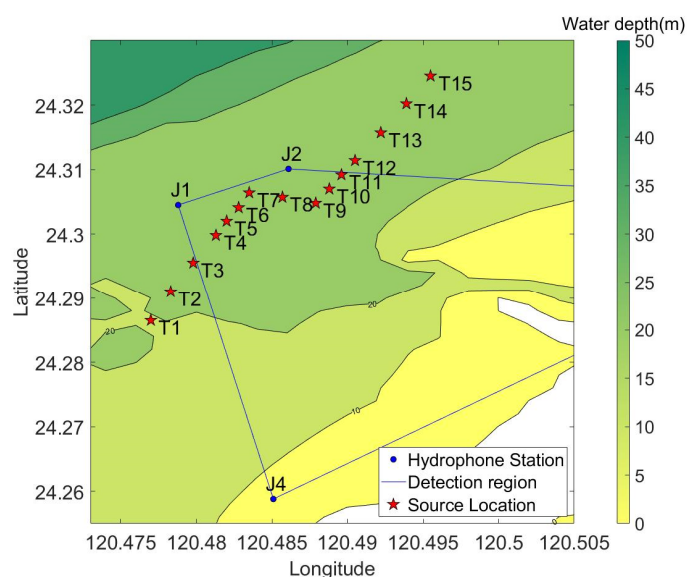


**Figure 11.** Output results of rising frequency signals: (a) NTU\_PAM; (b) PAMGuard with 1024 data points window length, 90% overlap ratio and 6 dB SNR.

In the localization experiment, the TDOA method proved useful for localizing the whistle source. Figure 12 plots the errors of the three different types of signals at each spot and indicates that the error is small when the source is inside the region of the four hydrophone recorders (points T4–T11); when outside the region (points T1–T3 and T12–T15), location was only approximate (Figure 13). The results of this experiment indicate strengths in using the NTU\_PAM for successful tracking of cetaceans.



**Figure 12.** Distribution of localization errors.



**Figure 13.** Detection region consisted of hydrophones.

## 6. Conclusions

In this research, we devised and developed the NTU\_PAM algorithm, which performs whistle detection and whistle localization based on the TDOA method. The results showed NTU\_PAM is able to localize and track the whistle sound source with high accuracy. In the future, MMOs can monitor the moving path of marine mammals via the visual method combined with NTU\_PAM, making it possible to monitor cetaceans without being limited by daylight hours.

**Author Contributions:** Conceptualization, C.-T.H., W.-L.L., Y.-H.H., W.-Y.C. and C.-F.C.; methodology, C.-T.H., W.-L.L., Y.-H.H., W.-Y.C. and C.-F.C.; software, C.-T.H., W.-L.L. and W.-Y.C.; formal analysis, W.-Y.C. and W.-C.H.; writing—original draft preparation, C.-T.H. and W.-Y.C.; writing—review and editing, C.-T.H. and C.-F.C.; supervision, C.-F.C.; project administration, C.-F.C.; funding acquisition, C.-F.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Ministry of Science and Technology, Taiwan (MOST 109-2221-E-002-198-MY3).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Request from the corresponding author of this article.

**Acknowledgments:** The authors would like to thank the Formosa Plastics Group and the Formosa Petrochemical Corporation, 1141946-00, for the 2 minutes of data. The authors would like to thank the Ministry of Science and Technology, Taiwan, for the funding; and Professor Lien-Siang Chou for help with marine mammal knowledge.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Short Biography of Authors



**Ching-Tang Hung** received his B.S. degree from Department of Hydraulic and Ocean Engineering, National Cheng Kung University and M.S. degree from Department of Engineering Science and Ocean Engineering, National Taiwan University. He is currently a Ph.D. candidate in Department of Engineering Science and Ocean Engineering, National Taiwan University. His research interests include underwater acoustic, whistle detection and automated unmanned surface vehicle system.



**Yen-Hsiang Huang** received his B.S. degree from the Department of Mechanical Engineering, National Chun Hsing University and M.S. degree from Department of Engineering Science and Ocean Engineering, National Taiwan University. He is currently working as firmware engineer in Taiwan. His research interest including underwater acoustics, USV system integration and strong skill in C/C++ programming.



**Wei-Yen Chu** received his B.S. degree from Department of Marine Engineering, National Taiwan Ocean University and M.S. degree from Department of Engineering Science and Ocean Engineering, National Taiwan University. He is currently working in the semiconductor manufacturing area in Taiwan. His research interests include underwater acoustics, signal processing, whistle detection, localization and simulation.



**Wei-Chun Hu** is a Ph.D. candidate at the Department of Engineering Sciences and Ocean Engineering, National Taiwan University, Taiwan. His research focuses specifically on the soundscape and the propagation of underwater noise. Recent participated publications can be found in Ecological Indicators and Entropy journal.



**Wei-Lun Li** received his B.S. degree from Department of Hydraulic and Ocean Engineering, National Cheng Kung University and M.S. degree from Department of Engineering Science and Ocean Engineering, National Taiwan University. He is currently working as research assistant, Institute of hydrobiology, Chinese Academy of science. His research interests include underwater acoustic and underwater detection of marine mammals.



**Chi-Fang Chen** received her Ph.D. in the Department of Ocean Engineering, Massachusetts Institute of Technology in 1991, and started her career as the faculty member of the Department of Naval Architecture of National Taiwan University from 1991 till now. (Department of Naval Architecture was renamed as Department of Engineering Science and Ocean Engineering in 2000). Her research expertise and interests are underwater acoustics and underwater acoustic propagation. She is conducting passive acoustic monitoring (PAM) in recognizing sounds from different species in the ocean which includes *Sousa Chinensis* in Taiwan waters. She also has interests in autonomous ocean sensing, and has supervised two master's theses in AUV, and is now supervising five graduate students in autonomous surface vehicle study.

## References

1. Bureau of Energy, Ministry of Economic Affairs (Ed.) *Bureau of Energy Annual Report*; Bureau of Energy, Ministry of Economic Affairs: Taiwan, 2018.
2. Department of Information Services, Executive Yuan (Ed.) *Four-Year Wind Power Promotion Plan*; Department of Information Services, Executive Yuan: Taiwan, 2019.
3. Bailey, H.; Senior, B.; Simmons, D.; Rusin, J.; Picken, G.; Thompson, P. Assessing underwater noise levels during pile-driving at an offshore windfarm and its potential effects on marine mammals. *Mar. Pollut. Bull.* **2010**, *60*, 888–897. [[CrossRef](#)] [[PubMed](#)]
4. Au, W.W.L. Hearing in Whales and Dolphins: An Overview. In *Hearing by Whales and Dolphins*; Springer: Berlin/Heidelberg, Germany, 2000; pp. 1–42.
5. Oswald, J.N.; Rankin, S.; Barlow, J. To Whistle or Not to Whistle? Geographic Variation in the Whistling Behavior of Small Odontocetes. *Aquat. Mamm.* **2008**, *34*, 288–302. [[CrossRef](#)]
6. Spaulding, E.; Robbins, M.; Calupca, T.; Clark, C.W.; Tremblay, C.; Waack, A.; Warde, A.; Kemp, J.; Newhall, K. An autonomous, near-real-time buoy system for automatic detection of North Atlantic right whale calls. In Proceedings of the Meetings on Acoustics 157ASA, Portland, Oregon, 18–22 May 2009; Acoustical Society of America: New York, NY, USA, 2009; Volume 6.
7. Linnenschmidt, M.; Teilmann, J.; Akamatsu, T.; Dietz, R.; Miller, L.A. Biosonar, dive, and foraging activity of satellite tracked harbor porpoises (*Phocoena phocoena*). *Mar. Mammal Sci.* **2012**, *29*, E77–E97. [[CrossRef](#)]
8. Akamatsu, T.; Wang, D.; Wang, K.; Naito, Y. Biosonar behaviour of free-ranging porpoises. *Proc. Soc. B Biol. Sci.* **2005**, *272*, 797–801. [[CrossRef](#)] [[PubMed](#)]
9. Gannier, A.; Fuchs, S.; Quèbre, P.; Oswald, J.N. Performance of a contour-based classification method for whistles of Mediterranean delphinids. *Appl. Acoust.* **2010**, *71*, 1063–1069. [[CrossRef](#)]
10. Lai, Y.-C. The Analysis of Identification of Cetaceans' Whistle with Support Vector Machine. Master's Thesis, Department of Engineering Science and Ocean Engineering, National Taiwan University, Taipei, Taiwan, 2014; pp. 1–61.
11. Caldwell, M.C.; Caldwell, D.K. Individualized whistle contours in bottle-nosed dolphins (*Tursiops truncatus*). *Nature* **1965**, *207*, 434–435. [[CrossRef](#)]
12. Caldwell, M.C.; Caldwell, D.D. *Statistical Evidence for Individual Signature Whistles in the Pacific Whitesided Dolphin, Lagenorhynchus obliquidens*; Los Angeles County Museum CA: Los Angeles, CA, USA, 1970.
13. Caldwell, M.C.; Caldwell, D.K.; Miller, J.F. *Statistical Evidence for Individual Signature Whistles in the Spotted Dolphin, Stenella plagiodon*; Los Angeles County Museum Calif: Los Angeles, CA, USA, 1970.
14. Datta, S.; Sturtivant, C. Dolphin whistle classification for determining group identities. *Signal Process.* **2002**, *82*, 251–258. [[CrossRef](#)]



15. Bahoura, M.; Simard, Y. Blue whale calls classification using short-time Fourier and wavelet packet transforms and artificial neural network. *Digit. Signal Process.* **2010**, *20*, 1256–1263. [[CrossRef](#)]
16. Gillespie, D.; Caillat, M.; Gordon, J.; White, P. Automatic detection and classification of odontocete whistles. *J. Acoust. Soc. Am.* **2013**, *134*, 2427–2437. [[CrossRef](#)] [[PubMed](#)]
17. Gillespie, D.; Mellinger, D.; Gordon, J.; McLaren, D.; Redmond, P.; McHugh, R.; Trinder, P.; Deng, X.; Thode, A. PAMGUARD: Semiautomated, open source software for real-time acoustic detection and localisation of cetaceans. *J. Acoust. Soc. Am.* **2008**, *30*, 54–62. [[CrossRef](#)]
18. Lin, T.-H. The Application of Passive Acoustic Monitoring for Studying Indo-Pacific Humpback Dolphin Behavior and Habitat Use off Western Taiwan. Master's Thesis, Institute of Ecology and Evolutionary Biology, National Taiwan University, Taipei, Taiwan, 2013; pp. 1–150.
19. Lin, T.-H.; Chou, L.-S.; Akamatsu, T.; Chan, H.-C.; Chen, C.-F. An automatic detection algorithm for extracting the representative frequency of cetacean tonal sounds. *J. Acoust. Soc. Am.* **2013**, *134*, 2477–2485. [[CrossRef](#)] [[PubMed](#)]
20. Janik, V.M.; Parijs, S.M.; Thompson, P.M. A two-dimensional acoustic localization system for marine mammals. *Mar. Mammal Sci.* **2006**, *16*, 437–447. [[CrossRef](#)]
21. Wang, Z.-T.; Au, W.W.; Rendell, L.; Wang, K.-X.; Wu, H.-P.; Wu, Y.-P.; Liu, J.-C.; Duan, G.-Q.; Cao, H.-J.; Wang, D. Apparent source levels and active communication space of whistles of free-ranging Indo-Pacific humpback dolphins (*Sousa chinensis*) in the Pearl River Estuary and Beibu Gulf, China. *PeerJ* **2016**, *4*, 1695. [[CrossRef](#)] [[PubMed](#)]
22. Wiggins, S.M.; McDonald, M.A.; Hildebrand, J.A. Beaked whale and dolphin tracking using a multichannel autonomous acoustic recorder. *J. Acoust. Soc. Am.* **2012**, *131*, 156–163. [[CrossRef](#)]
23. Wiggins, S.M.; Hildebrand, J.A. High-frequency Acoustic Recording Package (HARP) for broad-band, long-term marine mammal monitoring. In Proceedings of the 2007 Symposium on Underwater Technology and Workshop on Scientific Use of Submarine Cables and Related Technologies, Tokyo, Japan, 17–20 April 2007; Institute of Electrical and Electronics Engineers (IEEE): Piscataway, NJ, USA, 2007; pp. 551–557.
24. Wiggins, S.M.; Frasier, K.E.; Henderson, E.E.; Hildebrand, J.A. Tracking dolphin whistles using an autonomous acoustic recorder array. *J. Acoust. Soc. Am.* **2013**, *133*, 3813–3818. [[CrossRef](#)] [[PubMed](#)]
25. Li, W.-L. Study of Dolphin Whistle Detection. Master's Thesis, Department of Engineering Science and Ocean Engineering, National Taiwan University, Taipei, Taiwan, 2018; pp. 1–73.
26. Griffin, D.; Lim, J. Signal estimation from modified short-time Fourier transform. *IEEE Trans. Acoust. Speech, Signal Process.* **1984**, *32*, 236–243. [[CrossRef](#)]
27. Podder, P.; Khan, T.Z.; Khan, M.H.; Rahman, M.M. Comparative Performance Analysis of Hamming, Hanning and Blackman Window. *Int. J. Comput. Appl.* **2014**, *96*, 1–7. [[CrossRef](#)]
28. Chan, R.H.; Ho, C.-W.; Nikolova, M. Salt-and-pepper noise removal by median-type noise detectors and detail-preserving regularization. *IEEE Trans. Image Process.* **2005**, *14*, 1479–1485. [[CrossRef](#)] [[PubMed](#)]
29. Steinley, D. K-means clustering: A half-century synthesis. *Br. J. Math. Stat. Psychol.* **2006**, *59*, 1–34. [[CrossRef](#)] [[PubMed](#)]
30. Kaune, R. In Accuracy studies for TDOA and TOA localization. In Proceedings of the 2012 15th International Conference on Information Fusion, Singapore, 9–12 July 2012; IEEE: Piscataway, NJ, USA, 2012; pp. 408–415.
31. Chou, L.-S.; Ding, J.-J.; Lin, H.-J.; Suen, J.-P. *Population Ecology and Estuary Habitat Monitoring for Chinese White Dolphin (Sousa chinensis) (II)*; Forestry Bureau, Council of Agriculture, Executive Yuan: Taiwan, 2019.
32. Chou, L.-S.; Lin, H.-J.; Suen, J.-P. *Population Ecology and Estuary Habitat Monitoring for Chinese white Dolphin (Sousa Chinensis)*; Forestry Bureau, Council of Agriculture, Executive Yuan: Taiwan, 2017.
33. Kuo, L.-H. Statistical Analysis of Whistle and Ambient Noise of Summer in the Habitat of Indo-Pacific Humpback Dolphin (*Sousa chinensis*). Master's Thesis, Department of Engineering Science and Ocean Engineering, National Taiwan University, Taipei, Taiwan, 2014; pp. 1–99.
34. Chi-Fang, C.; Wei-Chun, H.; Chi-Hung, L. *Population Ecology for Indo-Pacific Humpback Dolphins (Sousa chinensis) in Yunlin, Taiwan*; Forestry Bureau, Council of Agriculture, Executive Yuan: Taiwan, 2021.