



**HAL**  
open science

## Efficient artifacts filter by density-based clustering in long term 3D whale passive acoustic monitoring with five hydrophones fixed under an Autonomous Surface Vehicle

Maxence Ferrari, Marion Poupard, Pascale Giraudet, Ricard Marxer, Jean-Marc Prévot, Thierry Soriano, Hervé Glotin

### ► To cite this version:

Maxence Ferrari, Marion Poupard, Pascale Giraudet, Ricard Marxer, Jean-Marc Prévot, et al.. Efficient artifacts filter by density-based clustering in long term 3D whale passive acoustic monitoring with five hydrophones fixed under an Autonomous Surface Vehicle. OCEANS 2019, Jun 2019, Marseille, France. pp.39, 10.1109/OCEANSE.2019.8867416 . hal-02313922

**HAL Id: hal-02313922**

**<https://hal.science/hal-02313922>**

Submitted on 11 Oct 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Efficient artifacts filter by density-based clustering in long term 3D whale passive acoustic monitoring with five hydrophones fixed under an Autonomous Surface Vehicle

Maxence Ferrari

*Univ. Toulon, Aix Marseille Univ.*

CNRS, LIS, DYNI, Marseille, France

LAMFA, CNRS, Amiens Univ., France

SMIoT Toulon, France

maxence.ferrari@univ-tln.fr (corr. author)

Marion Poupard

*U. Toulon, AMU, CNRS*

LIS, DYNI, Marseille, France

SMIoT Toulon, France

Biosong SARL, France

marion.poupard@univ-tln.fr

Pascale Giraudet

*U. Toulon, AMU, CNRS*

LIS, DYNI, Marseille, France

SMIoT Toulon, France

pascale.giraudet@univ-tln.fr

Ricard Marxer

*U. Toulon, AMU, CNRS*

LIS, DYNI, Marseille, France

ricard.marxer@lis-lab.fr

Jean-Marc Prévot

*U. Toulon*

DSIUN, Computer Science dpt.

jmp@univ-tln.fr

Thierry Soriano

*U. Toulon*

COSMER, France

thierry.soriano@univ-tln.fr

Hervé Glotin

*U. Toulon, AMU, CNRS*

LIS, DYNI, Marseille, France

SMIoT Toulon, France

glotin@univ-tln.fr (corr. auth.)

**Abstract**—Passive underwater acoustics allows for the monitoring of the echolocation clicks of cetaceans. Static hydrophone arrays monitor from a fixed location, however, they cannot track animals over long distances. More flexibility can be achieved by mounting hydrophones on a mobile structure. In this paper, we present the design of a small non-uniform array of five hydrophones mounted directly under the Autonomous Surface Vehicle (ASV) Sphyrna (also called an Autonomous Laboratory Vehicle) built by SeaProven in France. This configuration is made challenging by the 40cm aperture of the hydrophone array, extending only two meters below the surface and above the thermocline, thus presenting various artifacts. The array, fixed under the keel of the drone, is numerically stabilized in yaw and roll using the drone's Motion Processing Unit (MPU). To increase the accuracy of the 3D tracking computed from a four hour recording of a Sperm Whale diving several kilometers away, we propose an efficient joint filtering of the clicks in the Time Delay of Arrival (TDoA) space. We show how the DBSCAN algorithm efficiently removes any outlier detection among the thousands of transients, and yields to coherent high definition 3D tracks.

**Index Terms**—Drone, ASV, Passive Acoustic Monitoring, Cetacean Survey, Abyss Monitoring, 3D Tracking, Long Term Survey, Transient Analysis, Weak Signal Detection, DBSCAN, Clustering, Autonomous Laboratory Vehicle, Tortuosity, Sperm Whale, *Physeter macrocephalus*, Biosonar, High Sampling Rate, Scientific Microsystems for the Internet of Things (SMIoT).

## I. INTRODUCTION

In order to assess population status of cetacean and to model their behavior, or to prevent ship collisions, passive acoustic monitoring provides a viable option. Existing techniques use different methods to record and localize the echolocation clicks of cetaceans. Static hydrophone arrays, using underwater buoys, provide stable recordings and observations of foraging [1]–[3] but can only monitor a fixed location, making



Figure 1. The ASV Sphyrna, 17m long. It is a proa (South Pacific design), consisting of two unequal length parallel hulls. It is extremely stable, even at high wind force.

it difficult to track animals over longer distances. A bottom-mounted array with small aperture (2m) and high sampling rate has been deployed for 3D localization of multiple whales [4], [5]. More flexibility can be achieved by mounting hydrophones on a vessel. In another experiment [6], a wide-aperture towed array of two hydrophones was used for tracking in 2D dive profiles of Sperm Whale, taking advantage of surface-reflected paths, but not allowing 3D localization. In [7], a first attempt of a moving deep 6-hydrophone array resulted only into range estimation of the cetacean. In this paper, we designed a high sampling rate sound card and robust passive acoustic algorithm for a small non-uniform array of 5 hydrophones mounted directly under an autonomous surface vehicle (ASV) called Sphyrna. This ASV has been built by Sea Proven in France (Fig. 1, demo<sup>1</sup>).

This setup is made challenging by the small aperture and the

<sup>1</sup><http://sabiod.org/seeabyss>

recording of sound close to the surface, above the thermocline. The recordings are made at 16 bits@600 kHz with our available high-resolution sound card JASON (Fig. 3), designed in our scientific platform SMIoT [8]. In this paper we propose a method for efficient filtering of the transients and other noises generated in these extreme conditions, in order to monitor deep foraging cetaceans like the Sperm Whale (*Physeter macrocephalus*, Pm) which spends more than 70% of its time down to 1 km. In the following we will illustrate our algorithms on echolocation pulses of Pm that it is using for orientation and prey localization.

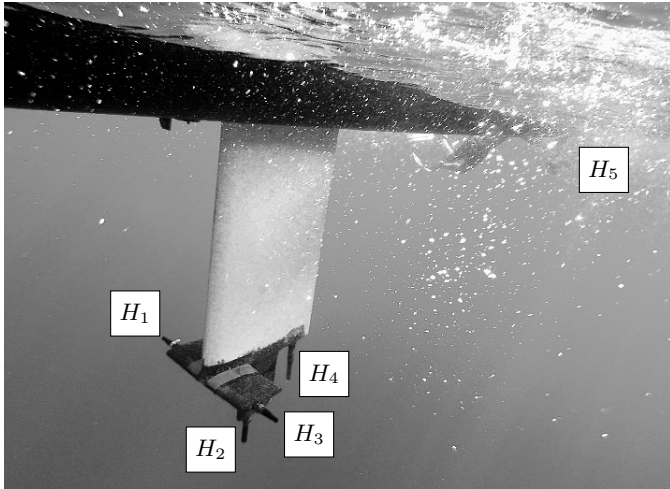


Figure 2. Layout of the 4+1 hydrophone array. The inter-hydrophone distances under the keel are 35, 59, 59, 63, 63, 70 cm. Hydrophone  $H_5$  is placed at the stern, 7 m away.

To the best of our knowledge, this is the first demonstration that cetaceans can be tracked precisely in spatially (3D) over time from a moving, near-surface hydrophone array. This mobile observatory opens up new possibilities to analyze the movements and behaviors of cetaceans, either when expressing their natural behavior or when disturbed by human activities nearby. Thus, it may assist in the design of mitigation and conservation programs.

## II. METHOD

To estimate the 3D localization of the animal, we compute the Time Delay of Arrival (TDoA) on signal chunks resulting from an automatic sperm whale click detector. TDoA estimation is cross-correlation based. The clicks are extracted using a simple spike detector on an enhanced signal (or noise-reduced bandpass-filtered signal).

The cross-correlation based method we apply to compute TDoA also yields values linked to the energy of the click and its coherence between all pairs of channels. However, after this phase, echoes from the surface and from objects in the water column still exist. Thus, we cluster the clicks by tracks in the time-TDoA space with the DBSCAN algorithm (Density-Based Spatial Clustering of Applications with Noise [9]). DBSCAN is a density-based clustering non-parametric algorithm: given a set of points in some space, it groups

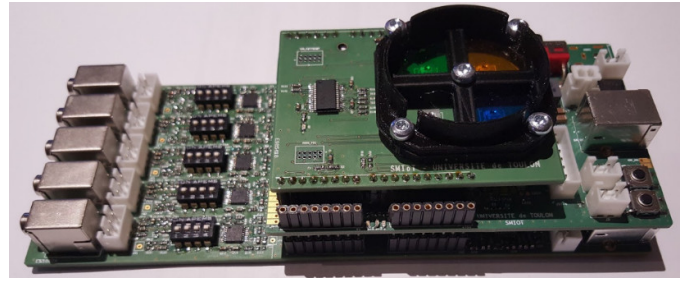


Figure 3. The JASON sound card (from <http://smiot.univ-tln.fr>), up to 5 x 2 MHz sampling rate at 16 bits resolution, placed into the drone (the built-in luxmeter was not used in this work).

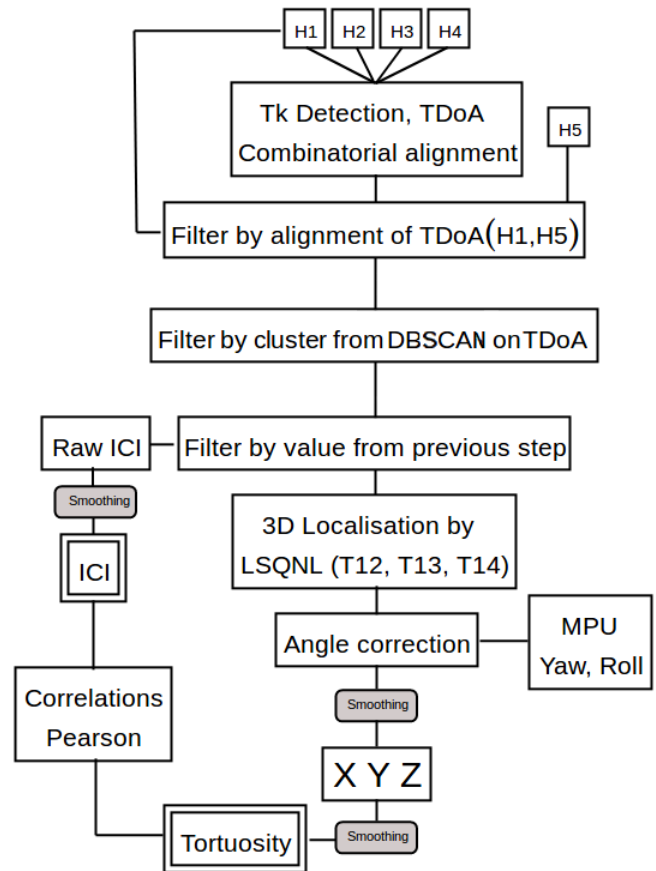


Figure 4. Flowchart of our method.

together points that are closely packed together (points with many nearby neighbors), marking as outliers points that lie alone in low-density regions (whose nearest neighbors are too far away). This clustering procedure contributes to discarding spurious clicks (e.g. false positives from the spike detector) and groups together tracks of clicks facilitating the analysis of the cetacean behavior. We finally estimate 3D positions of the whale from the TDoAs using a nonlinear solver. The tracks are stabilized according to the yaw and roll from the ship's Motion Processing Unit (MPU).

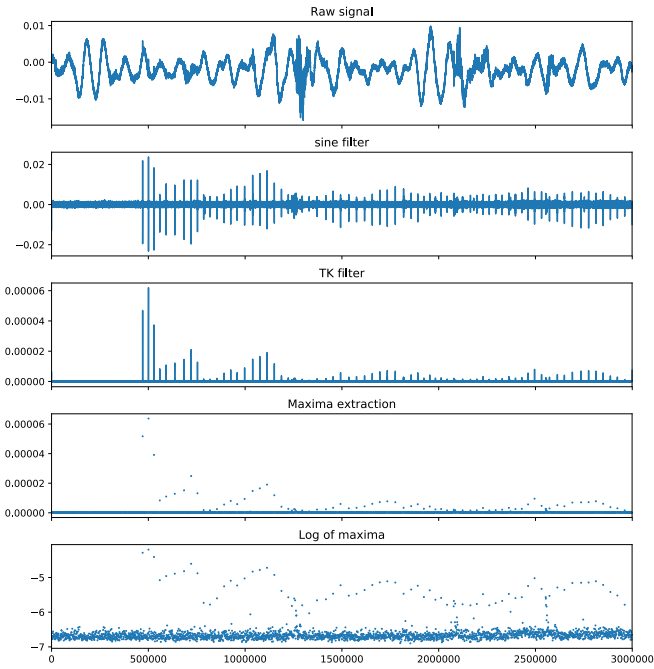


Figure 5. The four steps of the proposed click detector. Note that for visual purposes, the sampling rate have been reduced by a factor six.

### III. TDOA COMPUTATION

#### A. Click detection

We designed a click detector composed of five simple steps illustrated in Fig. 5. We cross-correlate the signal with one period of a 12.5 kHz sine wave which acts as a band-pass filter (bandwidth of echolocation clicks is 10–15 kHz [10]). We then apply a Teager-Kaiser filter [11], [12] and extract the local maxima in 20 ms windows (twice the largest inter-pulse interval of 10 ms [13]). For each 1 minute audio file we compute the mean and standard deviation of the maxima values in decibels (dB), and only keep samples over three times the standard deviation [14].

For audio recordings with a large amount of clicks, we must first cluster the local maxima using a 2-component Gaussian Mixture Model (GMM), and only keep samples that are over three times the standard deviation of the component with the lower mean. This latter component is assumed to be the one capturing spurious clicks.

#### B. Time Delay of Arrival (TDoA) estimation

Let  $\tau_{ij}$  denote the TDoA between hydrophones  $H_i$  and  $H_j$ , and let  $x_i$  be the signal recorded by  $H_i$ . We propose an approach similar to steered response power (SRP) [15], but in our case, we process multiplications instead of additions. We thus improve the TDoA computation by combining multiple cross-channel information. For  $N$  channels, the method starts by computing the cross-correlation  $c_{ij}$  between all the  $\binom{N}{2}$  possible pairs of signals  $(x_i, x_j)$ . Then, the method subtracts the minimum of  $c_{ij}$  to it. With  $N$  channels, there are only

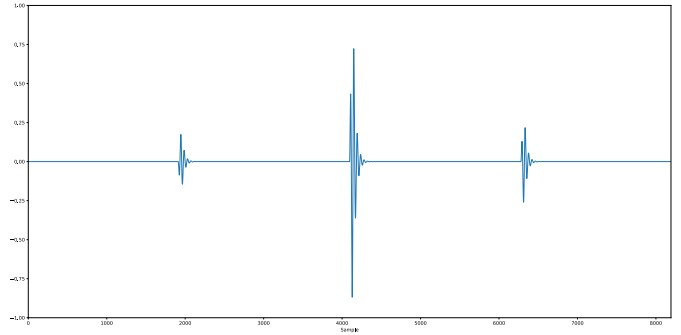


Figure 6. Synthetic click used for the SNR - TDoA simulation illustrated in Fig. 7.

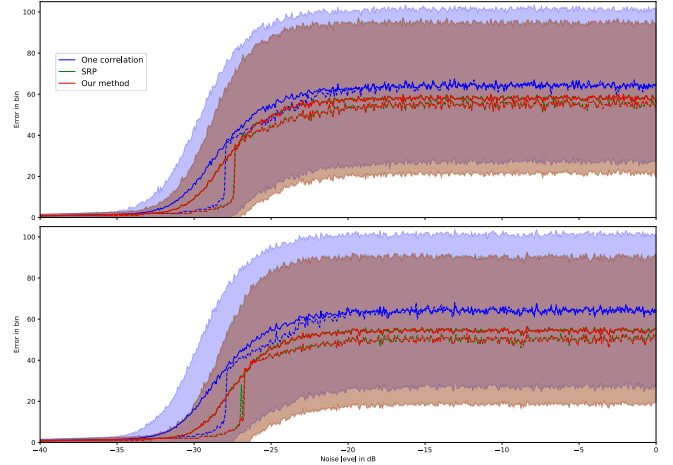


Figure 7. TDoA error with 3 channels (*top*) and TDoA error with 4 channels (*bottom*) computed on simulated clicks at various SNR. The full lines are the mean error, the dashed lines are the median of the error, and the filled region corresponds to one standard deviation across the mean.

$N - 1$  independent TDoAs (for instance  $\{\tau_{i1}\}_{i \in [2, N]}$ ) because of the relation  $\tau_{ij} = \tau_{ik} + \tau_{kj}$ . In our setup there are in fact three independent TDoAs. If the sound is recorded on all channels the maximum of each cross-correlation  $c_{ij}$  is at  $\tau_{ij}$ . However, this does not hold if the recordings are noisy. That is why both methods try to maximize the values of all the cross correlation regarding a set of independent TDoAs  $\{\tau_{i1}\}_{i \in [2, N]}$ .

In short our method finds the TDoAs by computing:

$$\tau_{21}, \tau_{31}, \dots, \tau_{n1} = \arg \max_{\hat{\tau}_{21}, \hat{\tau}_{31}, \dots, \hat{\tau}_{n1}} \prod_{i=3}^n \prod_{j=2}^{i-1} (x_i \otimes x_j)(\hat{\tau}_{i1} - \hat{\tau}_{j1}) - \min(x_i \otimes x_j), \quad (1)$$

where  $\otimes$  denotes cross-correlation.

In Fig. 7 we show the results of both methods to compute  $\tau_{21}$  on a signal in three or four channels with various SNR dB levels of white noise. We also compared them to  $\tau_{21}$  obtained with the cross-correlation between  $x_1$  and  $x_2$ . This figure shows the absolute error in bins of 2048 examples for each

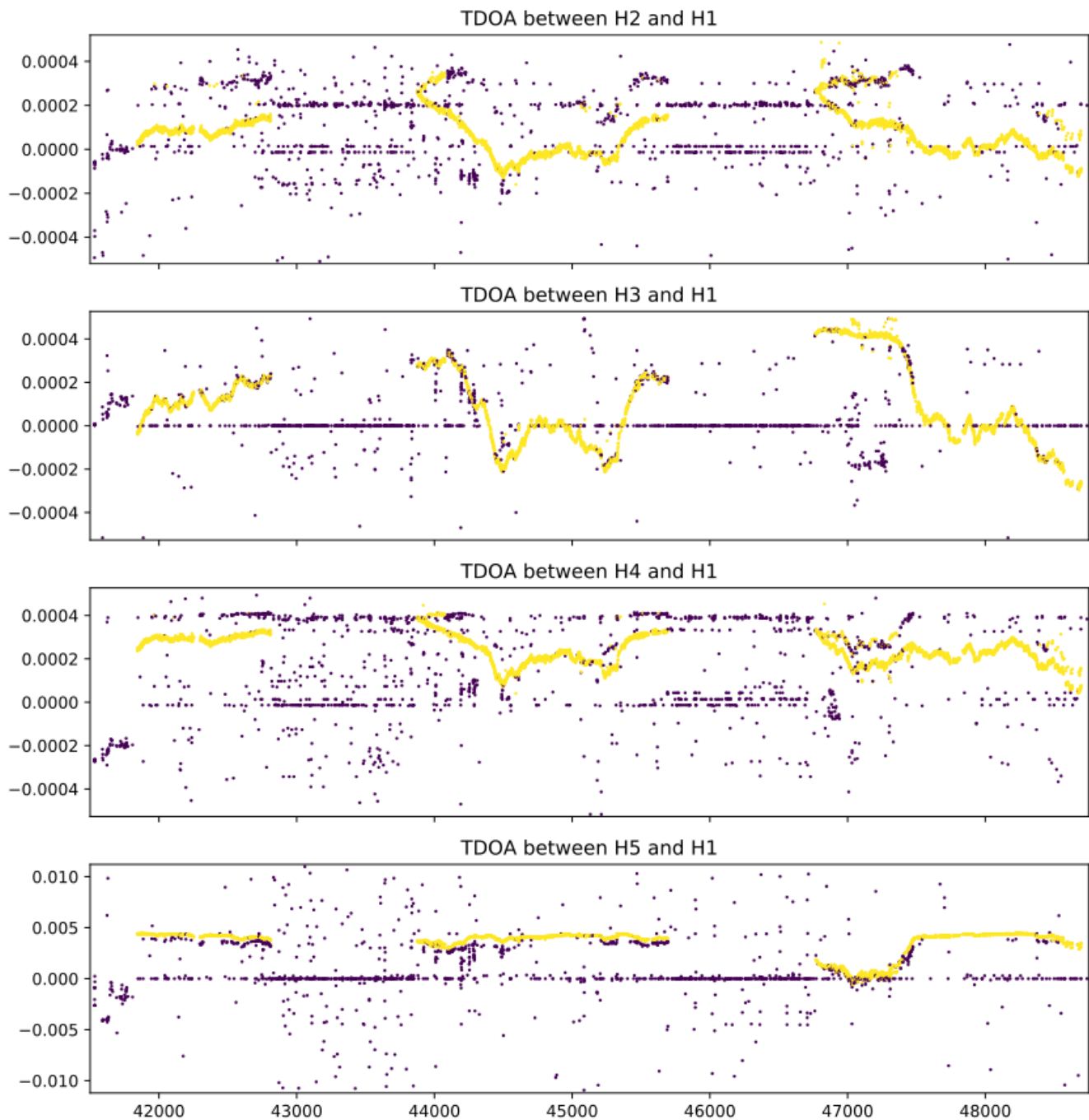


Figure 8. TDoA during 3h of recording. The yellow points are the points kept after the DBSCAN clustering.

level of noise. The TDoA search was done between -128 and +128 samples. The source signal was synthesized using an adaptation of a dolphin click generation model [16], which is used here to generate the pulse P1 of the sperm whale click:

$$U(t) = U_0 \operatorname{ReLU} \left( \frac{a^{\operatorname{ReLU}(t/T-m)} - a^{t/T}}{1 - a^m} \right), \quad (2)$$

where  $m$  is the number of periods to reach the maximum of amplitude,  $a$  is the reflectance factor, and  $T$  is the period of

the carrier waveform. The dolphin-specific parameters  $m = 5$  and  $a = 0.84$  have been changed to  $m = 1$  and  $a = 0.25$  to approximate sperm whales pulses. A carrier frequency of 12,5 kHz was used. The denominator has also been changed from  $1 - a$  to  $1 - a^m$  in order to have a maximum magnitude of  $U_0$ . The equation (2) describes the envelope of one pulse. To better mimic sperm whale clicks [13], two other pulses of relative amplitudes -0.2 and 0.3 were added with an Inter

Pulse Interval (IPI) of 4 ms.

We then divide the TDoA estimation into two parts. First, we compute the three independent TDoAs  $\tau_{21}, \tau_{31}, \tau_{41}$  of the keel antenna ( $H_1$  to  $H_4$ ), then we compute the remaining TDoA  $\tau_{51}$ . Compared to estimating all TDoAs jointly, this reduces the time complexity, and accounts for our non-uniform array, with  $H_5$  placed further apart. For the keel antenna, we first compute cross-correlations between all six possible hydrophone pairs, and make them positive by subtracting their minimal values. We then search for the combination of  $\tau_{21}, \tau_{31}, \tau_{41}$  that maximizes their product. For the remaining  $\tau_{51}$ , we use the same method, but hold  $\tau_{21}, \tau_{31}, \tau_{41}$  fixed, only searching over one dimension.

#### IV. FILTERING

##### A. Filtering on TDoA

By splitting the TDoA computation in two, we obtain two values for each click. One resulting from the multiplication of the cross-correlations between the hydrophones of the keel antenna, and the other from the multiplication of the cross-correlation between the last hydrophone and each of the other hydrophones of the keel antenna 9.

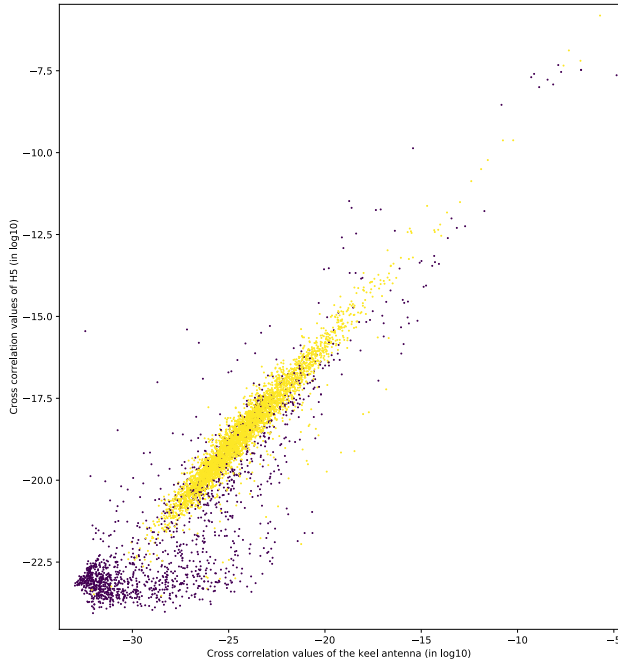


Figure 9. Scatter plot of the two cross correlation values. Same 3h of recording as in the Fig. 8.

The Fig. 8 shows the four independent TDoAs  $\tau_{21}, \tau_{31}, \tau_{41}, \tau_{51}$  during a 3-hour recording session where the sperm whale we observed did three dives. Horizontal tracks of points in this figure were caused by artifacts. In order to distill the actual clicks from the echoes and false positives, we apply DBSCAN clustering on the time-TDoA space and keep clusters linked to the main track. The metric used was the euclidean distance. The optimal value for epsilon is set

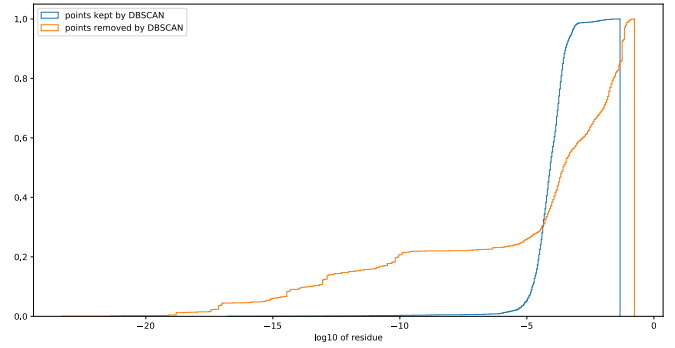


Figure 10. Cumulative density histogram of the residue of the localization algorithm.

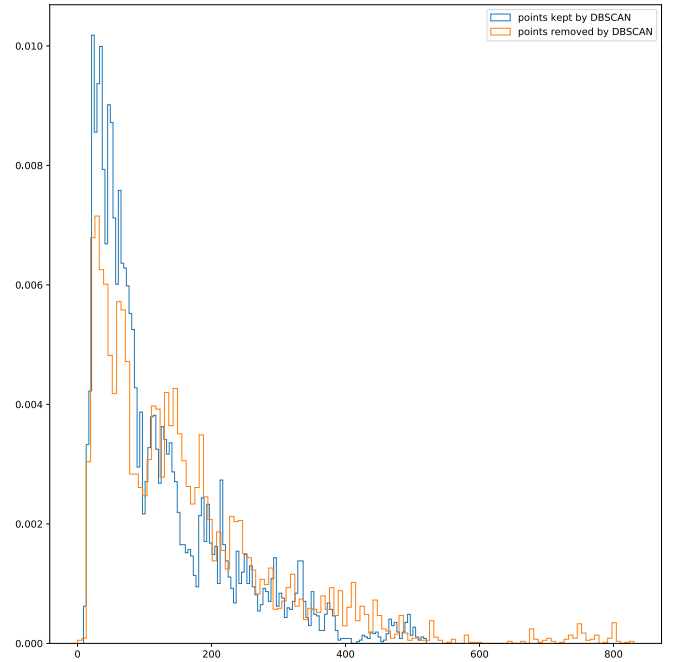


Figure 11. Moving standard deviations in a 40 sec window for the set of points kept by DBSCAN (blue) versus for points removed by DBSCAN (red).

according to the computed number of clusters in the epsilon function. We then determined the clusters for five epsilon values that were in middle of the plateau depicting the evolution in number of clusters. Finally we chose the values of epsilon that clustered the track with a minimal number of clusters (4 main clusters for the three dives and two small clusters at the end of the last track) eliminating most of the spurious and echo clicks. This gives a better result than filtering by keeping points that belong to the Gaussian with the highest energy (points that were in the upper right quadrant of  $(-29, -22)$  of Fig. 9).

The only echoes that we could not eliminate with DBSCAN were the surface echoes (for example in Fig. 8, TDoA between  $H_2$  and  $H_1$ , the yellow points that are above 0.00022 s), since their tracks originate from the same point as the main track,

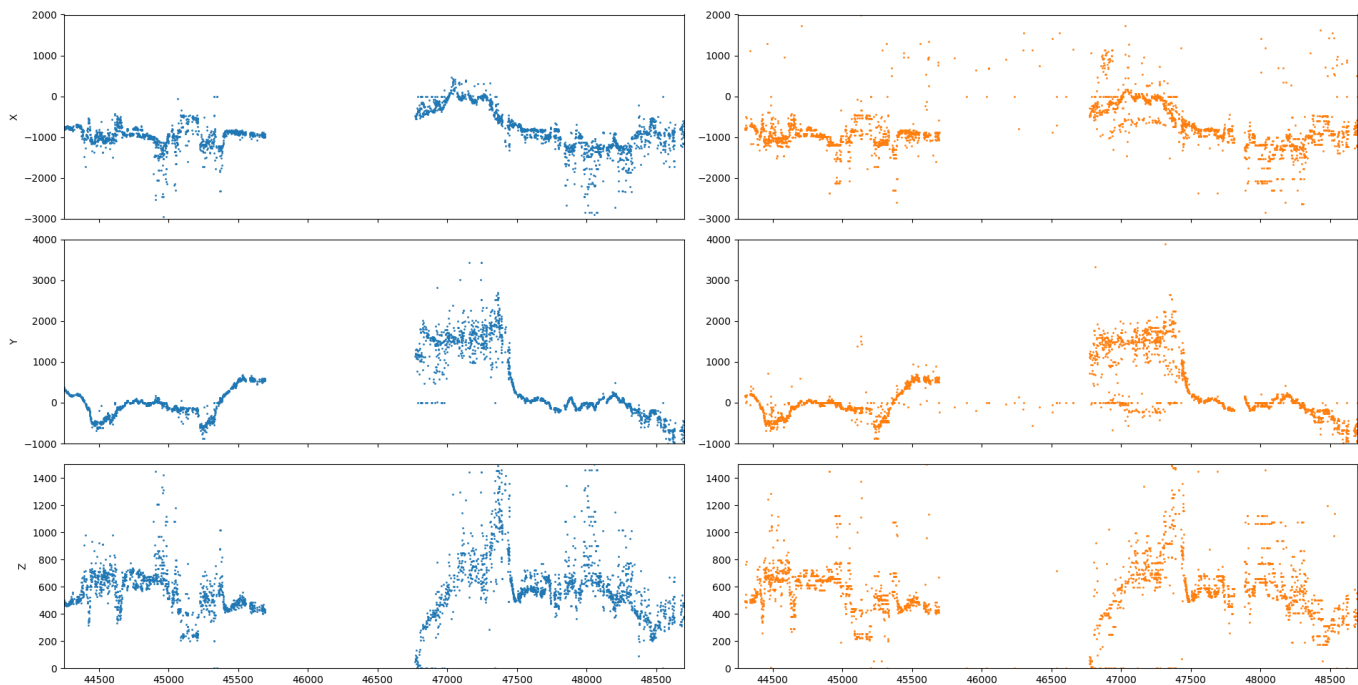


Figure 12. Raw coordinate of the Sperm Whale during the last two dives, in the keel antenna reference space. The blue points are from the DBSCAN filtering. The orange points are from the baseline [17].

which is the start of the dive.

### B. Source localization and post filtering

Once the list of clicks has been finalized, we converted their TDoA into cartesian coordinates using the keel antenna as base, these are then mapped to GPS positions using the yaw, roll and pitch obtained from the MPU. Surface echoes are eliminated at this stage since they lead to positions above the surface. We then remove the remaining noise with a moving median filter with a window size of 20 seconds.

We used the solver detailed in ([12]) to compute the 3D coordinates from the TDoAs. The histograms of the residue of the solutions are given Fig. 10, on the set of points kept using DBSCAN, and those discarded using DBSCAN. They show that the residue of retained clicks lie in a small range, while the others are more spread out, from smaller residues (artifacts) to larger residues (non coherent TDoAs), thus DBSCAN filters spurious transient detections.

### V. TRACKING RESULTS

In this section we compare our results to the 3D track obtained previously without the proposed filtering (<http://sabiod.org/seeabyss>) [17]. The Fig. 12 shows the improvements : the spurious clicks disappeared, so did some of the tracks of echoes.

In order to compare how the added DBSCAN filtering improved the track we obtained, we plotted the density histogram of standard deviation over 40 ms windows. The Fig. 11 shows that our method allows to remove all the positions with a standard deviation above 550 m, and increases by a factor 4/3

the proportion of positions with a standard deviation below 100 m.

### VI. CONCLUSION AND PERSPECTIVES

The DBSCAN clustering we added increased our ability to filter out spurious clicks without removing valid ones. The Fig. 13 shows the final 3D track we obtain after all filtering (including the moving median filter) were applied. These tracks demonstrates a normal foraging of a Pm. It is computed click by click, then it will allow to compute the tortuosity of the predator more precisely than in [17]. Moreover, based on this valuable data set, we offer open data to learn to estimate range or depth of the whale only from partial acoustic observations of these clicks<sup>2</sup>.

### ACKNOWLEDGMENTS

We thank M. Asch, V. Barchasz, V. Gies (SMIoT), S. Paris, N. Boizot, J. Patris, M. Grimaldi, J. Olson (Cetacean Research, USA), and D. B. for their collaboration. We thank the Direction Générale de l'Armement and Région Haut de France for Ferrari's PhD grant. This research has been possible due to Sea Proven, F. De Varenne and his team. We thank 'Marine et Océans' and Lemer Pax, H. Bergère and coll. for logistic help at Port-Cros Nat. Park, and A. Barcelo, M. Peirache (PNPC). We thank 'Fondation Prince Albert II de Monaco', S.E. B. Fautrier, P. Mondielli, TechnoLaval, Pôle 'Information Numérique Prévention et Santé' UTLN, Accobams, Pelagos, Marine Nationale, CNAM LAVAL, TES minister N. Hulot,

<sup>2</sup><http://sabiod.org/learnclicks>

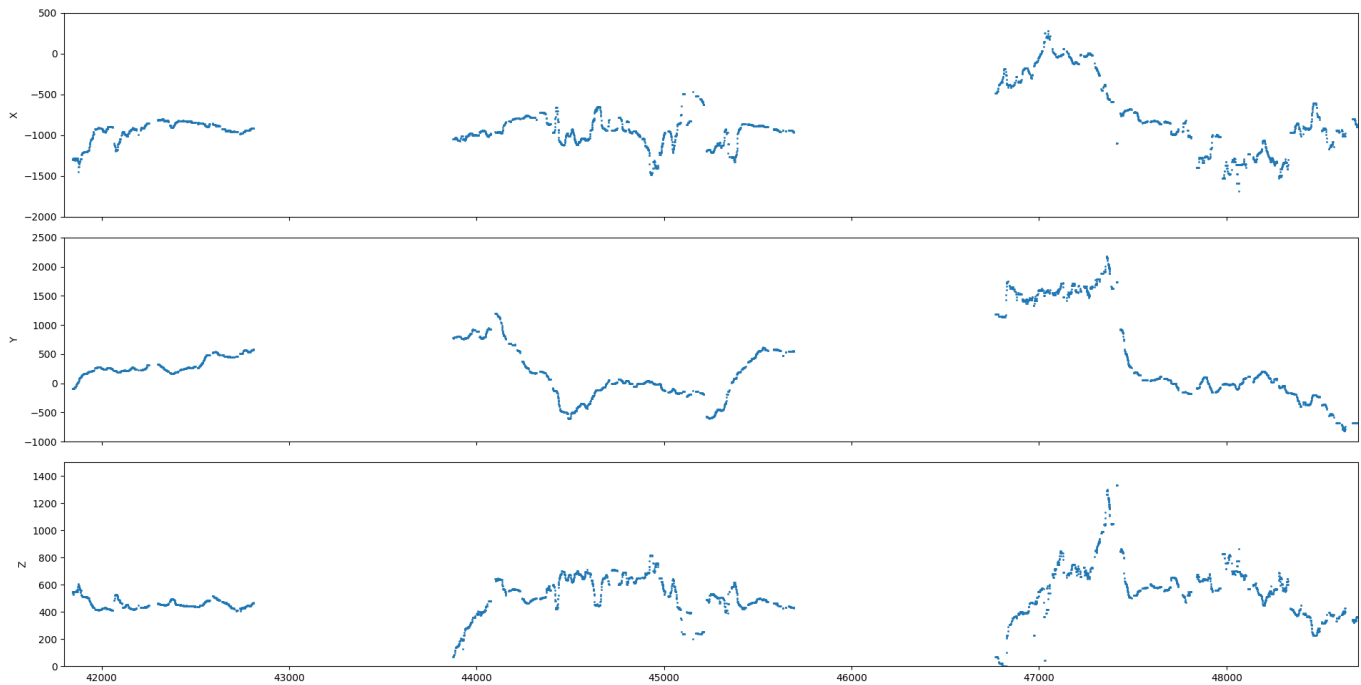


Figure 13. The computed track after DBSCAN of the three recorded dives, time in seconds. X is Northing (m), Y Easting (m), Z Depth (m). Z axes is directed downward.

and F. & V. Sarano of ONG Longitude 181. This project is part of SABIOD.org and EADM MADICS CNRS bio-acoustic research groups. It has partly been funded by FUI 22 Alysound, ANR-18-CE40-0014 SMILES, ANR-17-MRS5-0023 NanoSpike, and MARITTIMO european GIAS projects.

## REFERENCES

- [1] Y. Simard and N. Roy, "Detection and localization of blue and fin whales from large-aperture autonomous hydrophone arrays: A case study from the St. Lawrence estuary," *Canadian Acoustics*, vol. 36.1, no. 1, pp. 104–110, 2008.
- [2] N. Roy, Y. Simard, and C. Gervaise, "3D tracking of foraging belugas from their clicks: Experiment from a coastal hydrophone array," *Applied Acoustics*, vol. 71, no. 11, pp. 1050–1056, 2010.
- [3] H. Glotin et al., "Bombyx stereophonic sonobuoy: a permanent survey of cetacean in Pelagos sanctuary," *Pelagos Research Report*, vol. <http://sabioid.org/bombyx>, 2017.
- [4] F. Bénard-Caudal, P. Giraudet, and H. Glotin, "Whale 3D monitoring using astrophysic NEMO ONDE two meters wide platform with state optimal filtering by rao-blackwell monte carlo data association," *App. Acoustics*, vol. 71, no. 11, pp. 994–999, 2010.
- [5] P. Favali, F. Chierici, G. Marinaro, G. Giovanetti, et al., "NEMO-SN1 abyssal cabled observatory in the western ionian sea," *J. Oceanic Eng.*, vol. 38, no. 2, pp. 358–374, 2013.
- [6] A. Thode, "Tracking sperm whale (p. macrocephalus) dive profiles using a towed passive acoustic array," *J. of the Acoustical Society of America*, vol. 116, no. 1, pp. 245–253, 2004.
- [7] W. Zimmer, "Range estimation of cetaceans with compact volumetric arrays," *J. of Acoustical Society of America*, vol. 134, no. 3, pp. 2610–2618, 2013.
- [8] M. Fourniol, V. Gies, V. Barchasz, E. Kussener, H. Barthelemy, R. Vauché, and H. Glotin, "Low-power wake-up system based on frequency analysis for environmental internet of things," in *Int. Conf. on Mechatronic, Embedded Systems, App.* IEEE, 2018, pp. 1–6.
- [9] Martin Ester, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *KDD'96 Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*. 1996, pp. 226–231, AAAI Press.
- [10] P.-T. Madsen, R. Payne, N. Kristiansen, M. Wahlberg, I. Kerr, and B. Møhl, "Sperm whale sound production studied with ultrasound time/depth-recording tags," *J. of Exp. Biology*, vol. 205, no. 13, pp. 1899–1906, 2002.
- [11] V. Kandia and Y. Stylianou, "Detection of sperm whale clicks based on the Teager–Kaiser energy operator," *Applied Acoustics*, vol. 67, pp. 1144–1163, 2006.
- [12] H. Glotin, F. Caudal, and P. Giraudet, "Whale cocktail party: real-time multiple tracking and signal analyses," *Canadian acoustics*, vol. 36, no. 1, pp. 139–145, 2008.
- [13] R. Abeille, Y. Doh, P. Giraudet, H. Glotin, J.-M. Prévot, and C. Rabouy, "Estimation robuste par acoustique passive de l'intervalle-inter-pulse des clics de physeter macrocephalus: méthode et application sur le parc national de Port-Cros," *Journal of the Scientific Reports of Port-Cros National Park*, vol. 28, 2014.
- [14] F. Pukelsheim, "The three sigma rule," *The American Statistician*, vol. 48, no. 2, pp. 88–91, 1994.
- [15] J. DiBiase, H. Silverman, and M. Brandstein, "Robust localization in reverberant rooms," in *Microphone Arrays*, pp. 157–180. Springer, 2001.
- [16] N. Dubrovsky and L. Giro, "Modeling of the click-production mechanism in dolphins," in *Echolocation in Bats and Dolphins.*, p. 68. University of Chicago Press, 2004.
- [17] M. Poupard, M. Ferrari, J. Schluter, R. Marxer, P. Giraudet, V. Barchasz, V. Giés, G. Pavan, and H. Glotin, "Real-time passive acoustic 3d tracking of deep diving cetacean by small non-uniform mobile surface antenna," in *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2019, pp. 8251–8255.