

# SCIENTIFIC REPORTS



OPEN

## Sperm whale long-range echolocation sounds revealed by ANTARES, a deep-sea neutrino telescope

Received: 08 November 2016

Accepted: 14 February 2017

Published: 12 April 2017

M. André<sup>1</sup>, A. Caballé<sup>1</sup>, M. van der Schaar<sup>1</sup>, A. Solsona<sup>1</sup>, L. Houégnigan<sup>1</sup>, S. Zaugg<sup>1</sup>, A. M. Sánchez<sup>1</sup>, J. V. Castell<sup>1</sup>, M. Solé<sup>1</sup>, F. Vila<sup>1</sup>, D. Djokic<sup>1</sup>, S. Adrián-Martínez<sup>2</sup>, A. Albert<sup>3</sup>, M. Anghinolfi<sup>4</sup>, G. Anton<sup>5</sup>, M. Ardid<sup>2</sup>, J.-J. Aubert<sup>6</sup>, T. Avgitas<sup>7</sup>, B. Baret<sup>7</sup>, J. Barrios-Martí<sup>8</sup>, S. Basa<sup>9</sup>, V. Bertin<sup>6</sup>, S. Biagi<sup>10</sup>, R. Bormuth<sup>11,12</sup>, M. C. Bouwhuis<sup>11</sup>, R. Bruijn<sup>11,13</sup>, J. Brunner<sup>6</sup>, J. Busto<sup>6</sup>, A. Capone<sup>14,15</sup>, L. Caramete<sup>16</sup>, J. Carr<sup>6</sup>, S. Celli<sup>14,15</sup>, T. Chiarusi<sup>17</sup>, M. Circella<sup>18</sup>, A. Coleiro<sup>7</sup>, R. Coniglione<sup>10</sup>, H. Costantini<sup>6</sup>, P. Coyle<sup>6</sup>, A. Creusot<sup>7</sup>, A. Deschamps<sup>19</sup>, G. De Bonis<sup>14,15</sup>, C. Distefano<sup>10</sup>, I. Di Palma<sup>14,15</sup>, C. Donzaud<sup>7,20</sup>, D. Dornic<sup>6</sup>, D. Drouhin<sup>3</sup>, T. Eberl<sup>5</sup>, I. El Bojaddaini<sup>21</sup>, D. Elsässer<sup>22</sup>, A. Enzenhöfer<sup>6</sup>, K. Fehn<sup>5</sup>, I. Felis<sup>2</sup>, L. A. Fusco<sup>17,23</sup>, S. Galatà<sup>7</sup>, P. Gay<sup>7,24</sup>, S. Geißelsöder<sup>5</sup>, K. Geyer<sup>5</sup>, V. Giordano<sup>25</sup>, A. Gleixner<sup>5</sup>, H. Glotin<sup>26,39,40</sup>, R. Gracia-Ruiz<sup>7</sup>, K. Graf<sup>5</sup>, S. Hallmann<sup>5</sup>, H. van Haren<sup>27</sup>, A. J. Heijboer<sup>11</sup>, Y. Hello<sup>19</sup>, J. J. Hernandez-Rey<sup>8</sup>, J. Hößl<sup>5</sup>, J. Hofestädt<sup>5</sup>, C. Hugon<sup>4,28</sup>, G. Illuminati<sup>8,14,15</sup>, C. W. James<sup>5</sup>, M. de Jong<sup>11,12</sup>, M. Jongen<sup>11</sup>, M. Kadler<sup>22</sup>, O. Kalekin<sup>5</sup>, U. Katz<sup>5</sup>, D. Kießling<sup>5</sup>, A. Kouchner<sup>7,26,39,40</sup>, M. Kreter<sup>22</sup>, I. Kreykenbohm<sup>29</sup>, V. Kulikovskiy<sup>10,30</sup>, C. Lachaud<sup>7</sup>, R. Lahmann<sup>5</sup>, D. Lefèvre<sup>31,41</sup>, E. Leonora<sup>25,32</sup>, S. Loucatos<sup>7,33</sup>, M. Marcelin<sup>9</sup>, A. Margiotta<sup>17,23</sup>, A. Marinelli<sup>34,35</sup>, J. A. Martínez-Mora<sup>2</sup>, A. Mathieu<sup>6</sup>, K. Melis<sup>11,13</sup>, T. Michael<sup>11</sup>, P. Migliozzi<sup>36</sup>, A. Moussa<sup>21</sup>, C. Mueller<sup>22</sup>, E. Nezri<sup>9</sup>, G. E. Pávlaş<sup>16</sup>, C. Pellegrino<sup>17,23</sup>, C. Perrina<sup>14,15</sup>, P. Piattelli<sup>10</sup>, V. Popa<sup>16</sup>, T. Pradier<sup>37,42</sup>, C. Racca<sup>3</sup>, G. Riccobene<sup>10</sup>, K. Roensch<sup>5</sup>, M. Saldaña<sup>2</sup>, D. F. E. Samtleben<sup>11,12</sup>, M. Sanguineti<sup>4,28</sup>, P. Sapienza<sup>10</sup>, J. Schnabel<sup>5</sup>, F. Schüssler<sup>33</sup>, T. Seitz<sup>5</sup>, C. Sieger<sup>5</sup>, M. Spurio<sup>17,23</sup>, Th. Stolarczyk<sup>33</sup>, A. Sánchez-Losa<sup>10,18</sup>, M. Taiuti<sup>4,28</sup>, A. Trovato<sup>10</sup>, M. Tselengidou<sup>5</sup>, D. Turpin<sup>6</sup>, C. Tönnis<sup>8</sup>, B. Vallage<sup>1,33</sup>, C. Vallée<sup>6</sup>, V. Van Elewyck<sup>7</sup>, D. Vivolo<sup>36,38</sup>, S. Wagner<sup>5</sup>, J. Wilms<sup>29</sup>, J. D. Zornoza<sup>8</sup> & J. Zuñiga<sup>8</sup>

<sup>1</sup>Technical University of Catalonia, BarcelonaTech (UPC), Laboratory of Applied Bioacoustics (LAB), Rambla Exposició, 24, 08800 Vilanova i la Geltru, Barcelona, Spain. <sup>2</sup>Institut d'Investigació per a la Gestió Integrada de les Zones Costaneres (IGIC) - Universitat Politècnica de Valencia, C/Paranimf 1, 46730 Gandia, Spain. <sup>3</sup>GRPHE - Université de Haute Alsace - Institut universitaire de technologie de Colmar, 34 rue du Grillenbreit BP 50568 – 68008, Colmar, France. <sup>4</sup>INFN - Sezione di Genova, Via Dodecaneso 33, 16146 Genova, Italy. <sup>5</sup>Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen Centre for Astroparticle Physics, Erwin-Rommel-Str. 1, 91058 Erlangen, Germany. <sup>6</sup>Aix-Marseille Université, CNRS/IN2P3, CPPM UMR 7346, 13288 Marseille, France. <sup>7</sup>APC, Université Paris Diderot, CNRS/IN2P3, CEA/IRFU, Observatoire de Paris, Sorbonne Paris Cité, 75205 Paris, France. <sup>8</sup>IFIC - Institut de Física Corpuscular (CSIC – Universitat de València) c/Catedrático José Beltrán, 2 E-46980 Paterna, Valencia, Spain. <sup>9</sup>LAM - Laboratoire d'Astrophysique de Marseille, Pôle de l'Etoile Site de Château-Gombert, rue Frédéric Joliot-Curie 38, 13388 Marseille Cédex 13, France. <sup>10</sup>INFN - Laboratori Nazionali del Sud (LNS), Via S. So a 62, 95123 Catania, Italy. <sup>11</sup>Nikhef, Science Park, Amsterdam, The Netherlands. <sup>12</sup>Huygens-Kamerlingh Onnes Laboratorium, Universiteit Leiden, The Netherlands. <sup>13</sup>Universiteit van Amsterdam, Instituut voor Hoge-Energie Fysica, Science Park 105, 1098 XG Amsterdam, The Netherlands. <sup>14</sup>INFN - Sezione di Roma, P.le Aldo Moro 2, 00185 Roma, Italy. <sup>15</sup>Dipartimento di Fisica dell'Università La Sapienza, P.le Aldo Moro 2, 00185 Roma, Italy. <sup>16</sup>Institute for Space Science, RO-077125 Bucharest, Magurele, Romania. <sup>17</sup>INFN - Sezione di Bologna, Viale Berti-Pichat 6/2, 40127 Bologna, Italy. <sup>18</sup>INFN - Sezione di Bari, Via E. Orabona 4, 70126 Bari, Italy. <sup>19</sup>Geoazur, UCA, CNRS, IRD, Observatoire de la Côte d'Azur, Sophia Antipolis, France. <sup>20</sup>Univ. Paris-Sud, 91405 Orsay Cedex, France. <sup>21</sup>University Mohammed I, Laboratory of Physics of Matter and Radiations, B.P.717, Oujda 6000, Morocco. <sup>22</sup>Institut für Theoretische Physik und Astrophysik, Universität Würzburg, Emil-Fischer Str. 31, 97074 Würzburg, Germany. <sup>23</sup>Dipartimento di Fisica e Astronomia

dell'Università, Viale Berti Pichat 6/2, 40127 Bologna, Italy. <sup>24</sup>Laboratoire de Physique Corpusculaire, Clermont Université, Université Blaise Pascal, CNRS/IN2P3, BP 10448, F-63000 Clermont-Ferrand, France. <sup>25</sup>INFN - Sezione di Catania, Viale Andrea Doria 6, 95125 Catania, Italy. <sup>26</sup>LSIS, Aix Marseille Université CNRS ENSAM LSIS UMR 7296 13397 Marseille, France. <sup>27</sup>Royal Netherlands Institute for Sea Research (NIOZ), Landsdiep 4, 1797 SZ 't Horntje (Texel), The Netherlands. <sup>28</sup>Dipartimento di Fisica dell'Università, Via Dodecaneso 33, 16146 Genova, Italy. <sup>29</sup>Dr. Remeis-Sternwarte and ECAP, Universität Erlangen-Nürnberg, Sternwartstr. 7, 96049 Bamberg, Germany. <sup>30</sup>Moscow State University, Skobel'syn Institute of Nuclear Physics, Leninskie gory, 119991 Moscow, Russia. <sup>31</sup>Mediterranean Institute of Oceanography (MIO), Aix-Marseille University, 13288, Marseille, Cédex 9, France. <sup>32</sup>Dipartimento di Fisica ed Astronomia dell'Università, Viale Andrea Doria 6, 95125 Catania, Italy. <sup>33</sup>Direction des Sciences de la Matière - Institut de recherche sur les lois fondamentales de l'Univers - Service de Physique des Particules, CEA Saclay, 91191 Gif-sur-Yvette Cédex, France. <sup>34</sup>INFN - Sezione di Pisa, Largo B. Pontecorvo 3, 56127 Pisa, Italy. <sup>35</sup>Dipartimento di Fisica dell'Università, Largo B. Pontecorvo 3, 56127 Pisa, Italy. <sup>36</sup>INFN - Sezione di Napoli, Via Cintia 80126 Napoli, Italy. <sup>37</sup>Université de Strasbourg, IPHC, 23 rue du Loess 67037 Strasbourg, France. <sup>38</sup>Dipartimento di Fisica dell'Università Federico II di Napoli, Via Cintia 80126, Napoli, Italy. <sup>39</sup>Université de Toulon CNRS LSIS UMR 7296 83957 La Garde, France. <sup>40</sup>Institut universitaire de France, 75005 Paris, France. <sup>41</sup>Université du Sud Toulon-Var, 83957 CNRS-INSU/IRD UM 110, La Garde Cédex, France. <sup>42</sup>CNRS, UMR7178, 67037 Strasbourg, France. Correspondence and requests for materials should be addressed to M.A. (email: michel.andre@upc.edu)

**Despite dedicated research has been carried out to adequately map the distribution of the sperm whale in the Mediterranean Sea, unlike other regions of the world, the species population status is still presently uncertain. The analysis of two years of continuous acoustic data provided by the ANTARES neutrino telescope revealed the year-round presence of sperm whales in the Ligurian Sea, probably associated with the availability of cephalopods in the region. The presence of the Ligurian Sea sperm whales was demonstrated through the real-time analysis of audio data streamed from a cabled-to-shore deep-sea observatory that allowed the hourly tracking of their long-range echolocation behaviour on the Internet. Interestingly, the same acoustic analysis indicated that the occurrence of surface shipping noise would apparently not condition the foraging behaviour of the sperm whale in the area, since shipping noise was almost always present when sperm whales were acoustically detected. The continuous presence of the sperm whale in the region confirms the ecological value of the Ligurian sea and the importance of ANTARES to help monitoring its ecosystems.**

The last decade has seen tremendous progress in the construction of undersea neutrino observatories. These telescopes rely on the detection of Cherenkov light emitted along the path of charged particles produced in neutrino interactions<sup>1,2</sup>. When located under the sea, neutrino telescopes are also multidisciplinary observatories that have the potential to provide a new level of understanding of the marine ecosystems at great depths<sup>3</sup>.

The ANTARES (Astronomy with a Neutrino Telescope and Abyss environmental REsearch) experiment is the largest neutrino telescope presently operating in the Northern Hemisphere<sup>1</sup>. It is located in the deep Mediterranean Sea at a depth of 2,475 m. In addition to light sensors, the ANTARES infrastructure also incorporates the AMADEUS (ANTARES Modules for Acoustic DETection Under the Sea<sup>2</sup>), system. This system comprises acoustic sensors (hydrophones) aiming at evaluating the potential for detection of ultra-high-energy neutrinos through the pressure waves they produce when interacting in water. The system of ANTARES comprises many acoustic hydrophones, designed to perform measurements of acoustic backgrounds in the deep sea.

Most oceanographic instruments moored on the seafloor are not connected with the surface: they have to run on batteries and store data locally, implying that scientists only have access to their data after the recovery of the instrument<sup>3</sup>. Cabled observatories such as ANTARES (France), JAMSTEC (Japan), NEPTUNE (Canada) and OBSEA (Spain), remove these restrictions by providing not only electrical power and real-time data acquisition, but also the possibility for scientists to interact directly with their sensors. They can control them and monitor the environment in reaction to the observation of peculiar events. New real-time opportunities are therefore given to marine sciences and geosciences: listening to marine organisms, observing bioluminescent creatures<sup>4</sup>, studying global change<sup>5</sup> and its consequence on marine circulation and biodiversity.

Understanding the link between natural and anthropogenic processes is indeed essential for predicting the magnitude and impact of future changes of the natural balance of the oceans. Amongst this wide variety of changes, the next decades will see increasing levels of offshore industrial development, and this will almost certainly lead to increased amounts of noise pollution in the oceans.

Noise pollution in the marine environment is an emerging but serious concern to science and society<sup>6</sup>. Its implications are less well understood than other global threats and, until recently, largely undetectable to everyone but the expert. In addition, the assessment of the acoustic impact of artificial sounds in the sea is not a trivial task, certainly because there is a lack of information on how marine organisms process and analyse sounds and how relevant these sounds are for the balance and development of their populations<sup>7</sup>. Organisms that are exposed to sound can be adversely affected both on a short timescale (acute effect) and on a long timescale (permanent or chronic effects). These adverse effects can be widespread and the European Commission decided in September 2010, under the Marine Strategy Framework Directive (MSFD) that two indicators for underwater noise be used in describing ocean Good Environmental Status<sup>8</sup>.

As an essential step reaching Good Environmental Status, European Member States were asked to establish monitoring programmes for assessment, enabling the state of the marine waters concerned to be evaluated

on a regular basis. The MSFD comes with criteria and methodological standards on GES of marine waters (Commission Decision 2010/477/EU), including two indicators on noise (Descriptor 11, Noise/Energy): Low and mid frequency impulsive sounds as well as continuous low frequency sound (ambient noise)<sup>6</sup>.

The need of including noise indicators in the MSFD is supported by recent findings, which patently showed that an exposure to low frequency sounds would trigger massive acoustic trauma in non-hearing specialists like cephalopods<sup>8–10</sup>. If we add that the negative consequences of a short or long term exposure to artificial sounds may not be immediately detected, one can understand how challenging it is to obtain objective data allowing an efficient control of the effects of anthropogenic sound in the sea.

Definitive studies on the response of marine mammals to anthropogenic sound are typically hampered by the short time spent at the surface and the deep-diving lifestyle of many vocalising species. Implemented within the framework of the European Sea-Floor Observatory Network of Excellence (ESONET) in 2007, the “Listen to the Deep Ocean Environment (LIDO, <http://listentothedeep.com>)” programme has developed and applied techniques for continuous noise measurement and passive acoustic monitoring (PAM) to world-wide cabled deep sea platforms and moored stations, representing the first and only PAM system available online (<http://listentothedeep.com>, ref. 11). The combination of this real-time data management and passive acoustic monitoring with the use of the latest technological developments in underwater neutrino acoustic detection has provided the scientific community with a hitherto non-available technology to reveal the daily behaviour of deep-sea marine organisms, therefore opening for the first time Internet access to deep ocean sound information.

The sperm whale (*Physeter macrocephalus*) is a successful deep-sea predator with a worldwide distribution. Sperm whales consume a wide diversity of prey: although squids appear to be their primary resource, fishes also are important in certain locations. It is largely accepted that sperm whales forage during deep dives that routinely exceed a depth of 400 m, sometimes reaching more than 2,000 m, and 30–40 min duration<sup>12</sup>. Their need of individually ingesting one ton of squid per day on average leads them to rarely remain in the same region for long, but to migrate year-round over large distances in search of food<sup>12</sup>.

While foraging, sperm whales produce short duration sound signals named clicks. *Usual clicks* are produced in a 10 Hz–30 kHz frequency band, at a rate defining an inter-click interval (ICI) that varies from 0.5 to 2.0 seconds during descent from the surface until the whale begins its ascent to the surface. There is evidence that usual clicks produced during foraging dives are directional, with an intense, forward-directed beam, presenting levels as high as 236 dB re 1  $\mu$ Pa at 1 m<sup>13</sup>. Their off-axis low frequency components can be detected up to a distance of 15 km in sea state 3. Recent data on sperm whale sounds suggested that usual clicks are appropriate for echolocation on these low-reflecting targets, thus providing support that sperm whales typically spend more than 80% of their time immersed in foraging processes<sup>13–15</sup>. In addition to usual clicks, sperm whales produce *creak clicks* during foraging that have properties equivalent to signals in buzzes, the terminating pulse trains known from echolocating bats during prey capture<sup>15</sup>. The frequency content of an on-axis creak click resembles that of a P1 pulse of on-axis usual clicks, which validates the same speculations for a match between the frequency content and backscatter of prey items in the meso- and bathypelagic ecosystem. Hence, creak clicks are just as suited for echolocation as are usual clicks, but the reduced output and high repetition rate suggest a shorter sonar range. The directional nature and the reduced ASL on-axis, compared with usual clicks, of 185–205 dB/1  $\mu$ Pa rms also make creak clicks hard to detect on a sufficient number of hydrophones to allow for localization of the source and subsequent derivation of source parameters<sup>15</sup>.

In this paper, we present the analysis of 2 years of uninterrupted acoustic data collected by the ANTARES telescope and automatically processed in real-time by the LIDO software package.: one year of data (2011) was used for training which results were then used for classification in the second year (2012). This allowed the training of classifiers that revealed and displayed directly on the Internet the daily long-range echolocation behaviour of the sperm whale in the deep Ligurian Sea, and a first year-round study of the presence and potential effects of shipping noise on this particular species.

## Results

Data was processed in consecutive data segments of 16.8 seconds length. Acoustic features were identified and classified for each segment, with 1323397 analyzed segments in total for the second year. Presence of a class was then evaluated per segment. In the following, SW, UC, and IS stand for the sperm whale, ultrasonic cetacean and shipping classes. When they are followed by a C (SWC, UCC, ISC) they describe the output of the GMM classifier as explained in the Methods section. When they are followed by an I (SWI, UCI, ISI) they represent the smoothed number of classified impulses in an analysed data segment. When they are followed by a P (SWP, ISP) they represent the output of the GLM presence model.

**Sperm whale presence model.** A generalized linear model was fitted with response variable sperm whale presence and as independent variables the log transformed SWI (SWI') and ISI (ISI'). The interaction between these two variables was also included. The noise level SPL and the log transformed UCI were initially tried as covariates but were not found to significantly model animal presence. The explained variance was 53%. As shown in Table 1, both SWI' and the interaction SWI' \* ISI' were significant while ISI' by itself did not have a significant influence. The 95% confidence interval of the number of segments with sperm whale presence (SWP) was [115.14, 142.9] corresponding to 12% and 15% of the total number of segments analysed.

$$SWP = \frac{1}{1 - \exp(-X\beta)} = \frac{1}{1 - \exp(-(-3.956 + 3.645 SWI' - 0.073 ISI' - 0.379 SWI' * ISI'))} \quad (1)$$

| Coefficient | Estimation | Std. Error | Z value | Pr (> z ) |
|-------------|------------|------------|---------|-----------|
| Intercept   | -3.956     | 0.303      | -13.057 | <2e-16    |
| SWI'        | 3.645      | 0.355      | 10.286  | <2e-16    |
| ISI'        | -0.073     | 0.153      | -0.477  | 0.6332    |
| SWI' * ISI' | -0.379     | 0.116      | -3.264  | 0.001     |

**Table 1. Sperm whale acoustic presence model: estimated coefficients, standard errors on the coefficients, corresponding Z value and two-sides probability.** Intercept, SWI' and the interaction SWI'\*ISI' were significant. ISI' by itself was not significant.

| Coefficient | Estimation | Std. Error | Z value | Pr (> z ) |
|-------------|------------|------------|---------|-----------|
| Intercept   | -1.949     | 0.199      | -9.774  | <2e-16    |
| ISI'        | 2.636      | 0.174      | 15.121  | <2e-16    |
| UCI'        | -0.444     | 0.062      | -7.191  | 2e-13     |
| SWI'        | 0.028      | 0.233      | 0.121   | 0.903     |
| SWI' * ISI' | -0.627     | 0.095      | -6.586  | 4e-11     |

**Table 2. Impulsive ship noise presence model coefficients: estimated coefficients, standard errors on the coefficients, corresponding Z value and two-sides probability.** Intercept, ISI', UCI' and interaction SWI' \* ISI' were significant. SWI' by itself was not significant.

| Month | Jan    | Feb    | Mar    | Apr    | May    | Jun    | Jul    | Aug    | Sep    | Oct    | Nov    | Dec    |
|-------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| SWP   | 36763  | 14207  | 27447  | 10452  | 12210  | 12969  | 41661  | 28112  | 27700  | 19363  | 22672  | 18269  |
| ISP   | 42294  | 34262  | 42445  | 36166  | 47073  | 48394  | 52335  | 57071  | 48622  | 40194  | 48833  | 33399  |
| Total | 159428 | 149142 | 159428 | 154285 | 159428 | 154285 | 159428 | 159428 | 154285 | 159428 | 154285 | 159428 |

**Table 3. Source presence estimated values during 2012: total number of segments with sperm whale and shipping impulse presence by month.** The bottom row shows the number of 16.8 second segments per month.

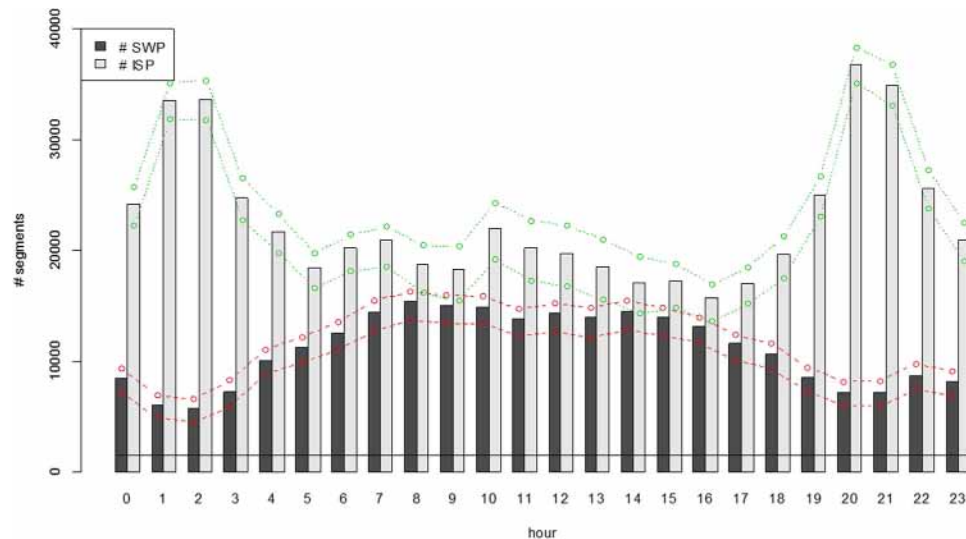
**Impulse ship presence model.** Impulsive shipping noise presence (ISP) was modelled by a GLM using the information in the log transformed ISI, SWI and UCI. As for the sperm whale model, SPL did not have a significant influence. The explained deviation of the model was 55%. As shown in Table 2, the interaction SWI' \* ISI' was included in the model. The 95% confidence interval of the number of segments with ISP was [367.5, 403.7], corresponding to between 38% and 42% of the total number of segments analysed.

**Missing data.** Several gaps in time were found, mostly caused by shipping noise presence: the noise produced by e.g., a ferry that passed by twice a day occasionally required shutting down the acoustic data acquisition due to system saturation. Therefore, the number of missing segments increased between 01:00–02:00 and 20:00–21:00 UTC with the highest percentage of missing data at 02:00 UTC (28% of the segments missing). Nonetheless, all daily hours were well represented in the recorded data. In order to allow comparison between hours of the day, the estimates of the hourly SWP and ISP were normalized with the percentage of missing data in each hour. Impulses were detected in more than half of the total number of segments, with maximum rates between 19:00 and 03:00 UTC.

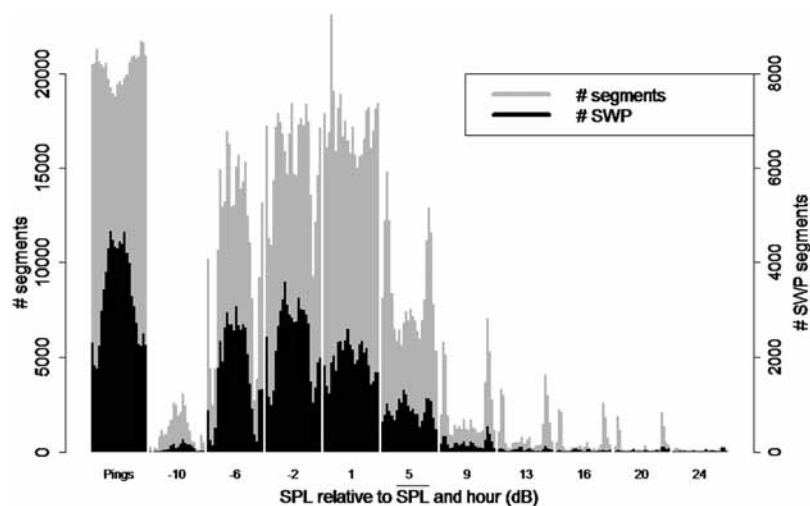
**Sperm whale and shipping presence recorded by the acoustic system of ANTARES during 2012.** The number of segments with sperm whales and shipping impulses was estimated for each month of 2012 (Table 3). All the segments processed in each month were evaluated by the SWP model and the ISP model. The resultant predictions of each class were normalised to the proportion of missing data. Shipping presence was more common than sperm whale presence: 29% versus 15% of the time respectively.

Most importantly, Table 3 shows the continuous presence of sperm whales in the vicinity of ANTARES and an increased presence in summer.

Shipping and sperm whale acoustic presence had contrasting behaviour during the day (Fig. 1). Although both classes contained many segments with presence over all 24 hours of the day, ISP had higher estimated values at night and SWP was larger during daylight hours. In terms of ISP, there were two apparent peaks at 01:00 and 20:00 UTC. Both corresponded to the regular passage of a ferry during those hours. The SWP predictions presented significant differences (Wilcoxon paired test: p-value < 0.001) over daylight hours (from 06:00 to 18:00 UTC) and night (from 18:00 to 06:00 UTC). In all twelve months of 2012 the SWP rate observed in daylight hours was larger than the rate at night. In Fig. 1, a 95% confidence interval is shown for the SWP predictions (red line) and ISP predictions (green line). The apparent diel shape for sperm whale presence anticorrelated with the shipping presence visible in Fig. 1 will be analyzed in detail at the discussion section below.



**Figure 1.** Number of segments with sperm whales and shipping impulse presence per hour in 2012. Sperm whale acoustic presence was higher during the daylight hours, whereas at night shipping impulses were more prominent. The red and green lines show the confidence interval at 95% for the SWP and ISP predictions, respectively, and the black line the sum of SWP predictions if the sperm whale model was evaluated with SWI = 0.



**Figure 2.** Number of recorded segments (grey, left y-axis) and segments with sperm whale acoustic presence (black, right y-axis) as function of hour and noise level (relative to SPL and hour). SPL were divided and rounded in ten bins. Each SPL features 24 bars corresponding to 24 hours.

**Sperm whale presence as function of noise level.** The sperm whale hourly acoustic presence for different levels of noise is presented in Fig. 2. At each SPL, 24 bars corresponding to each hour of the day were drawn, showing the total recorded segments and the segments with sperm whale presence. As explained in the methods section, the SPL was always high in segments that contained pings from the line localisation echo sounders. To avoid mixing these segments with segments that had increased SPL due to a continuous noise source, the segments with pings are shown separately in the first group of bars.

The grey bars (Fig. 2) show differences according to the SPL. At lower levels, many more segments are processed in daylight hours than in the night. As the noise level increases, the tendency is inverted, with a larger quantity of segments being processed at night (with two local maxima at the beginning and at the end of the day). The distribution of the total number of segments also exhibits a long right tail with many detected impulses at high SPL. These are likely produced by the same ships that also caused the increase in the noise.

Sperm whales were acoustically detected during all hours of the day regardless of noise levels. The hourly SWP distribution, with a peak around midday, is present at SPL between  $-6$  dB and  $5$  dB, and in the large dataset with the pings. When the relative SPL were over  $9$  dB very few sperm whales (or impulses in general) were detected and the hourly pattern of SWP (Fig. 2) is not visible anymore.

| Statistic | $\lambda$ | AUC              | FPR              | FNR              | NPV              | PPV              |
|-----------|-----------|------------------|------------------|------------------|------------------|------------------|
| SWC       | 3         | 0.84 (0.78–0.90) | 0.03 (0.02–0.05) | 0.42 (0.30–0.54) | 0.93 (0.92–0.95) | 0.75 (0.64–0.85) |
| ISC       | 5         | 0.93 (0.90–0.96) | 0.09 (0.06–0.13) | 0.14 (0.08–0.20) | 0.93 (0.90–0.95) | 0.82 (0.77–0.87) |
| UCC       | 9         | 0.93 (0.89–0.96) | 0.14 (0.08–0.21) | 0.07 (0.04–0.10) | 0.83 (0.77–0.88) | 0.94 (0.92–0.96) |

**Table 4. Training data accuracy: threshold ( $\lambda$ ), area under the curve (AUC), false positive rates (FPR), false negative rates (FNR), negative predictive value (NPV) and positive predictive value (PPV) from sperm whales, shipping and dolphins presence.** Confidence intervals are at 95%, with the AUC, FPR and FNR calculations using 2,000 stratified bootstrap replicates and the NPV and PPV for the asymptotic limits.

## Discussion

**Presence model.** The training data used to fit the model contained some segments with false negative observations when there were actually sperm whale impulses. These impulses were either (1) too faint to be picked up at the detection stage or (2) the classifier did not recognize these impulses as being SWC. This is reflected in Table 4, where even when no sperm whales are detected ( $SWI = 0$ ), the non-zero intercept results in a prediction of presence of about 0.02. The assumption that sperm whales can be present at any time (i.e. hour) at ANTARES throughout the year has been confirmed through human observation over the last three years.

Occasionally, when a high number of shipping impulses were produced, the classifier may have overestimated SWI by generating false positives. Shipping can produce hundreds of impulses in a single segment and even a small false positive rate may produce a significant SWI. This is reflected in the negative coefficient  $-0.379$  for the  $SWI'$  and  $ISI'$  product, which lowers the influence caused by false positives during ISI detections.

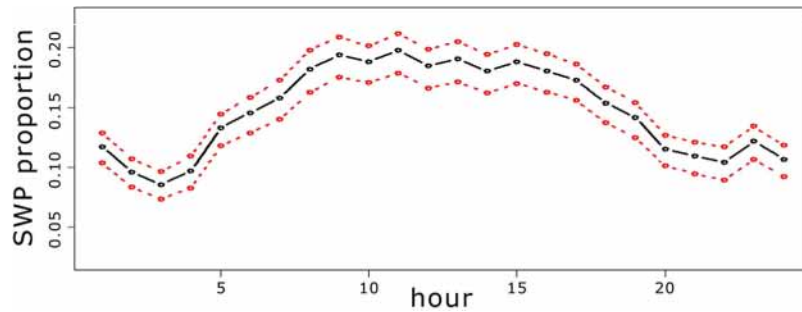
One problem that the model in eq. (2) cannot take into account is false negatives in the detection and classification phases due to masking of sperm whale sonar. While some information from the SPL measurements is included in the model through ISI (a high number of shipping impulses indicates a nearby ship, which usually produces high noise levels), SPL itself had no significance. When all sperm whale echolocation signals are masked to the point that it is not even possible to manually establish acoustic presence, then no training set can be created that takes into account this kind of false negative.

**Reasons for divergence between shipping and sperm whale diel behaviour.** The contrasting behaviour between the presence of ISI and sperm whales does not necessarily mean that sperm whales were avoiding high levels of noise caused by ships. Four different hypotheses could be posed: (1) sperm whales are more often present at daylight hours independently of shipping appearance; (2) sperm whales are present at least as often at night as at daylight hours; due to shipping occurring more at night; (3) sperm whale echolocation signals are misclassified as shipping at night but not at daylight hours; (4) sperm whale echolocation signals are masked by shipping impulses at night. Note that only hypothesis (1) and (2) are mutually exclusive.

By taking a look at Fig. 1, we could presume that hypothesis (1) is the most probable one: there is a significant difference of sperm whale acoustic presence between night and day hours. However, this assumption cannot be proven with the results presented above. It is not clear if the daily presence is due to what is indicated in hypothesis (3) or in hypothesis (4), which could then lead to hypothesis (2) as the real situation. To have a better idea of which hypotheses are more probable and how these could affect the results, two additional arguments are introduced:

- The number of data segments that contained a high relative SPL (more than 7 dB over  $\overline{SPL}$ ) but did not contain an ANTARES localisation signal was less than 4% of the total number of segments. Thus, the masking of impulses did not occur often enough to drastically change the obtained results. When SPL was low enough to avoid masking, about 13.25% of the analysed segments contained sperm whales (the average ratio between black and grey bars in Fig. 1). If we suppose a similar ratio when the SPL was high (as per hypothesis 4) then the error in sperm whale presence estimation due to masking would not be larger than 600 segments for each hour or less than 8% of the segments with sperm whale presence in Fig. 1. In that case, even if hypothesis (4) is true, then it would not significantly affect the results.
- An hourly behaviour analysis was reproduced with only the noise levels where no masking was found. For instance, relative SPL between  $-10$  to 1 dB were rarely affected by high noise level events (Fig. 2). Taking into account only segments in this SPL range, the ratio between the sum of SWP predictions and the total number of segments processed during each hour of the day was computed (Fig. 3).

At least two conclusions can be drawn from Fig. 3: (I) when no masking is produced the pattern seen in Fig. 3 seems consistent with the result obtained in Fig. 1. Hence, even if hypothesis (4) were true, when segments containing masking are discarded, providing more confidence in the reliability of the model, the conclusions seem to be in agreement with what was presented for the whole data set. (II) If we rely on the validity of the model, accurately estimating false positives in SWP and ISP from the training data, and we suppose that masking is due to high levels of noise instead of impulse ship presence, then the plot in Fig. 3 that includes all segments without masking should give the most reliable result. Although the hourly pattern could change slightly, it seems strong enough to indicate significant differences in the sperm whale acoustic presence during the different hours of the day.



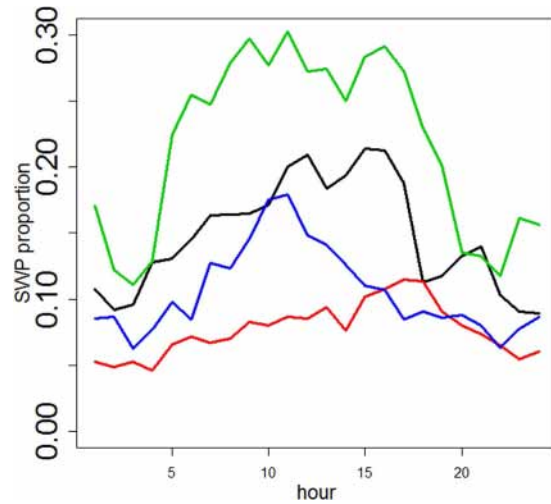
**Figure 3.** Proportion of sperm whale acoustic presence predictions considering only the segments with relative SPL lower than 1. A total of 917,233 segments were processed confirming a maximum at midday. The red line shows the confidence interval at 95%.

**Discussion and conclusions in the framework of this study.** The sperm whale is one of eight common cetacean species in the western Mediterranean Sea, and one of only two great whales encountered in the Mediterranean Sea<sup>16–22</sup>. Social groups have occasionally been reported in the recent past<sup>17,18</sup>, although schools of more than 15 individual animals were observed in the 1950s<sup>20</sup>. Genetic and observational evidence support the notion that the Mediterranean sperm whales constitute an isolated population, separated from the Atlantic stock. Robust information is lacking on their population status but their number would be in the hundreds rather than thousands and current densities appear to be much lower than those reported in the 1950s suggesting that a special attention should be devoted to the conservation status of this population<sup>22</sup>. By-catches linked to driftnet<sup>23</sup>, although illegal in the Mediterranean since 2002, have still been reported in recent times, while ship strikes and disturbance from maritime traffic represent a constant threat to this population<sup>22</sup>. This makes it vitally important to carefully monitor these potential sources of anthropogenic impact and to gather information on the current sperm whale distribution and behaviour to take conservation actions.

The analysis of two years of continuous acoustic data provided by the ANTARES neutrino telescope revealed the year-round presence of sperm whales in the Ligurian Sea, probably associated to the availability of their main prey in the region. Although creak clicks could not be detected at ANTARES probably due to distance, directionality and masking potential from other sources, the continuous presence of long-range echolocation sounds over time suggests that the sperm whales are foraging in the area. Several authors have previously noticed a link between sperm whale density and primary production<sup>24</sup>. While most of the Mediterranean is considered an oligotrophic sea, the western Ligurian Sea is characterised by high levels of primary productivity, caused by the interplay of climatic, oceanographic and physiographic factors<sup>25,26</sup>. A dominant cyclonic (counterclockwise) current, flowing north along Corsica and Tuscany and thence hugging the coast of Liguria and mainland France in a westerly direction, creates a permanent frontal system, which separates coastal and offshore waters<sup>27</sup>. Intense biological activity is generated along this water mass boundary by the enhanced productivity and retention associated with this frontal system<sup>28</sup>. Indeed, sperm whale vocalisations recorded at ANTARES are compatible with foraging behaviour, thus presumably indicating the availability of cephalopods – the sperm whale favourite diet – in the Ligurian Sea region, not only in summer but all year round.

The diel variation shown here in the feeding activity of the sperm whales in this region is probably also of relevance. Note that no seasonal variation was found in the detection of long-range echolocation sounds of the sperm whales (see Fig. 4). Diel patterns in behaviour have been documented for many marine mammals. For some whales, diel patterns are generally related to the timing of feeding, and hence should be appropriately related to environmental conditions, and food availability. Sperm whales feed mainly on mesopelagic and bathypelagic squid, such as squid from the families Gonatid, Histioteuthid and Onychoteuthid<sup>29</sup>. Some of these squid species, such as several species of *Gonatus*, conduct diel vertical migrations<sup>30</sup>, which may affect the foraging behaviour of sperm whales. However, the present results of increased sperm whale acoustic detection during daytime are in contrast with the findings by Matsushita<sup>31</sup> who proposed that the sperm whales in the Antarctic feed mainly at night when the squid rise to the sea surface. This conclusion was made after examination of stomach contents of whales caught early in the morning and at night, whose size was greater than that of those caught during the day.

Although the results presented here suggest foraging diel behaviour, it must be taken into account that the detection range of sperm whale signals is limited at around 15 km from the ANTARES telescope, not discarding that sperm whale preys may be moving away from the area (and sperm whales may consequently feed outside Antares acoustic range), nor that sperm whales can remain quiet for some time when gathered at the surface. Nevertheless, the hitherto unknown continuous presence of the Ligurian Sea sperm whales was demonstrated for the first time through the real-time analysis of audio data streamed from a cabled-to-shore deep-sea observatory. The unprecedented use of this technique and infrastructures allowed the hourly tracking of their long-range echolocation sounds on the internet, revealing a probable diel presence pattern: the sperm whales would mostly feed on cephalopods during daylight off Toulon coastlines. Whether this behaviour can be related to their prey vertical migration or simply to the absence of prey at certain times of the day, is still under study. Most interestingly, the same acoustic analysis revealed that the influence of surface shipping noise doesn't seem to condition the long-range echolocation behaviour of sperm whales in the area, since shipping noise was almost always present when sperm whales were detected. And although ship noise challenged the classification performance, sperm whales sounds were always detected at any time of the day. This continuous presence in the region confirms the



**Figure 4.** Proportion of sperm whale presence predictions per season (black line: winter, red line: spring, green line: summer and blue line: autumn) considering only the segments with relative SPL lower than 1.

ecological value of the Ligurian sea and the importance of ANTARES to help monitor its ecosystems. However, the high level of noise that marine fauna is continuously exposed to might represent in the medium term a threat to its conservation and efforts should be conducted to mitigate its effects, especially on a species like the sperm whale whose populations may be declining in the Mediterranean Sea.

## Methods

**The Acoustic Setup.** The AMADEUS acoustic detection equipment is designed to perform background measurements over a period of several years in an arrangement of acoustic sensors allowing studies at different temporal scales. ANTARES comprises 12 vertical structures, the detection lines, each one holding up to 25 storeys that are arranged at equal distances of 14.5 m along the lines<sup>2</sup>. Standard storeys contain sensors for the detection of Cherenkov light. Each line is anchored to the sea floor and held taut by a submerged buoy. A 13th line, called the Instrumentation Line (IL), is equipped with instruments for monitoring the environment. It holds six storeys. The acoustic sensor clusters of AMADEUS are distributed over the ANTARES detector with distances of 200 m horizontally and ranging from 10 m to 100 m vertically (Fig. 5). In total, 2 (out of 13) ANTARES lines are equipped with acoustic detection devices to form 6 so-called *acoustic storeys* representing the clusters of sensors, allowing position reconstruction of acoustic point sources for the frequency range from about 1 to 125 kHz<sup>2</sup>. Channel 15 in storey 24 (middle red dot hydrophone, L12, Fig. 5) was chosen for the analysis described in this paper because of its very low self-noise. Sampling rate was 250 kHz, hydrophone sensitivity  $-170$  dB re 1 V/ $\mu$ Pa, gain 10 dB, quantization +2 V.

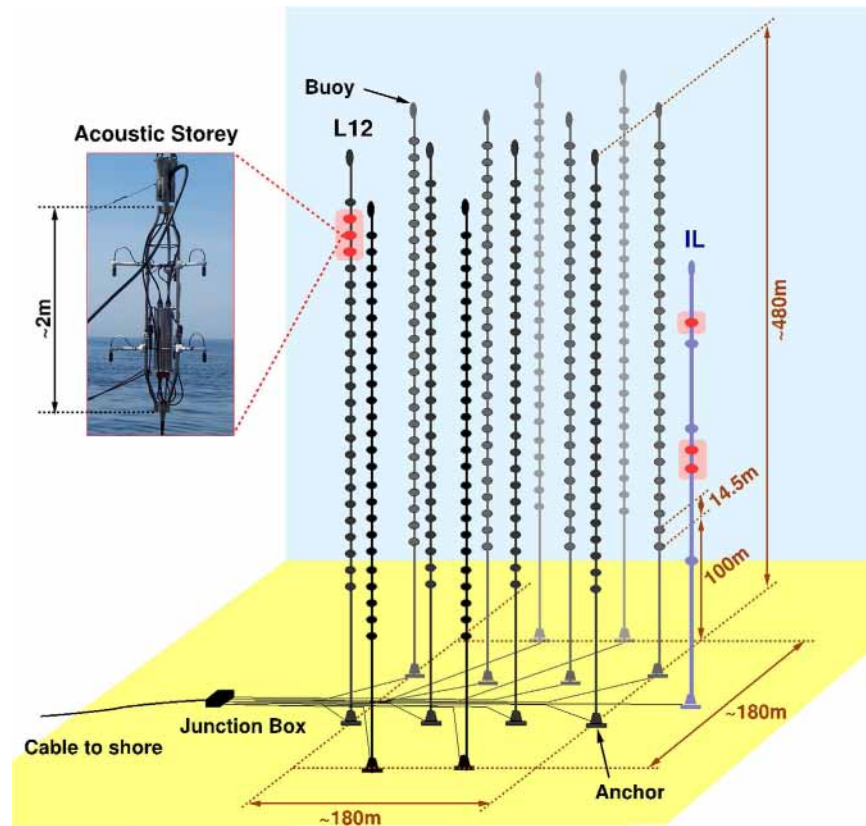
**The LIDO classification process.** The automated detection and classification of impulsive sounds in the marine environment is challenging. The intensity and the spectrum of the background noise is generally variable over periods of hours or even minutes due to changes in sea state or local anthropogenic activity, particularly shipping, which reduces the ability to detect weak impulses with a reasonably small false detection rate. Impulses are short and often only cover a fraction of the data stream. They must be automatically identified with sufficient precision by a detection stage. The acoustic features, which are extracted from each impulse, must give consistent values even under variations of the intensity and the spectrum of the background noise and the occurrence of interfering sounds.

The classification modules used in this article were designed to operate under such conditions; they are part of a larger system<sup>11</sup>, which includes impulse and short tonal sound detectors and classifiers applied to several frequency bands covering the whole bandwidth of the data. The modules work on segments ( $\sim 16.8$  seconds with a sampling rate of 250 kHz) of the audio stream that are automatically displayed in real-time on the <http://listen-tothedeeep.com> interface (Fig. 6).

In this paper, we concentrate on the sperm whale impulse classification<sup>32</sup> (SWC) and its performance in the presence of other impulsive sounds that were classified like ultrasonic cetacean impulses (UCC) and impulsive shipping noise (ISC). Ultrasonic cetacean impulses refer to other cetacean species vocalisations, in particular Delphinid and Zifid echolocation sounds whose lower frequencies overlap with the sperm whale frequency range. Other impulsive sources overlapping with the sperm whale frequency range were found to be very rare at the ANTARES platform.

Each detected impulse was systematically processed with a feature extraction step<sup>32</sup> and assigned a class with a classifier based on Gaussian Mixture Models (GMM). After classification, each detected impulse was tagged as SWC, ISC, or UCC. Impulses that were not recognized by the classifier received no tag and their effect on subsequent processing was neutral.





**Figure 5.** A sketch of the ANTARES detector. The six AMADEUS acoustic storeys are highlighted in pink-red (see text for details).

False positive rates - proportion of positives incorrectly classified - and false negative rates - proportion of positives not detected or detected but incorrectly classified - were estimated for different thresholds in the three classes (see Table 4).

The number of impulses assigned to each class per segment (the *detection rate*) was recorded; it is abbreviated SWI for the sperm whale class, ISI for the impulsive shipping class, and UCI for the ultrasonic cetacean class. The sources of these classes tended to produce long impulse trains covering multiple segments and providing reasonably continuous detection time intervals. For sperm whales this situation was different when only a single animal was present or when the animals were far away from the hydrophone. The number of detected impulses in a segment could differ considerably between consecutive segments. In some cases, a segment had zero detections while neighbouring segments did show sperm whale impulses. To obtain a smoother detection rate for sperm whale impulses (SWI) a running weighted average was computed over the number of impulses in a segment that were classified as sperm whale (SWC), as demonstrated in the paper by Zaugg *et al.*<sup>32</sup>. Thus SWI for a segment  $i$  was given by (1):

$$SWI_i = \sum_{t=i-5}^{i+5} \frac{SWC_t}{w_t}, \quad (2)$$

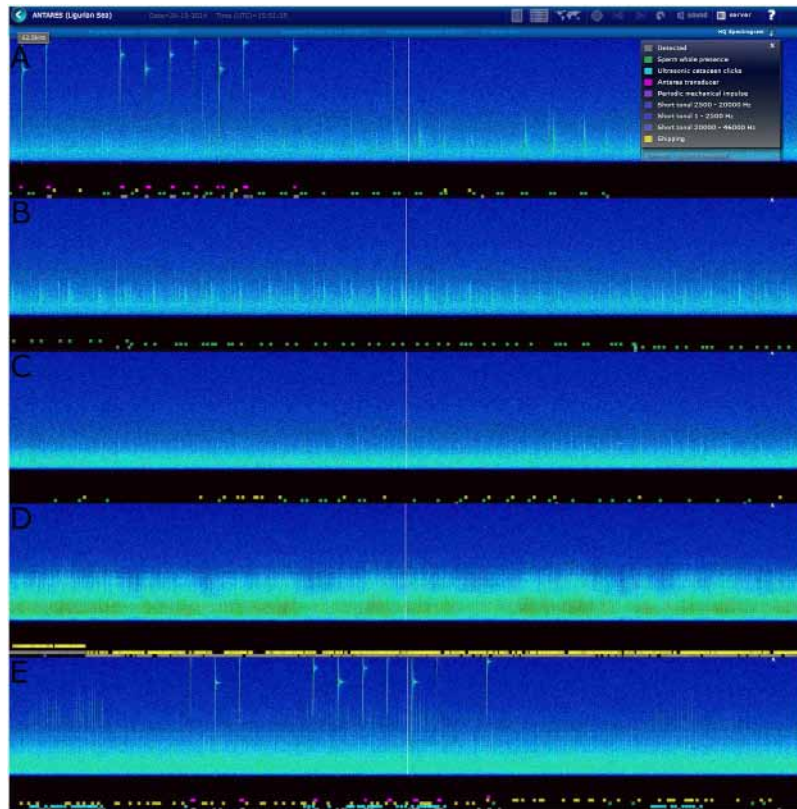
where  $w_t = 5$  if  $t = i$  and 1 otherwise, effectively evaluating the presence of sperm whale impulses in a 3 minute time interval<sup>32</sup>. The same was done for ultrasonic and shipping impulses.

In addition to the detections, noise levels were measured as well, calculating the sound pressure level (SPL) in each segment (2):

$$SPL_i = 10 \log \frac{1}{N_i} \sum_{k=1}^{N_i} \frac{p_{k,i}^2}{1 \mu Pa^2}, \quad (3)$$

where  $p_{k,i}$  is the measured pressure in sample  $k$  and  $N_i$  is the number of samples in a segment  $i$ .

The computed level does not represent the true sound pressure level at the site because the hydrophones of the AMADEUS system operate a high pass filter at 2 kHz. Since there is no interest here in absolute sound levels, but only in knowing when high levels of sound were present that might mask sperm whale signals or otherwise affect the classification system, the SPL (Sound Pressure Levels) reported in this paper are relative to the median SPL calculated over all available segments in 2012 (SPL).



**Figure 6.** Series of screenshots from <http://listentothedeep.com> interface where the output of the classification process is displayed. The legend in A shows coloured squares that are attributed to each class of detected signals displayed underneath the spectrogrammes: green for SWC (sperm whales), yellow for ISC (shipping), blue for UCC (cetacean sonar), pink for the ANTARES localisation mechanical pings. (A) Sperm whales are detected and classified as well as the ANTARES transducer pings. (B) Presence of a group of sperm whales. (C) Sperm whales in presence of incoming ship noise. (D) Ship noise is masking the low frequency part of the spectrogramme. (E) SWC, UCC, ISC and ANTARES transducer pings.

**Statistical analysis of the LIDO classifications.** In this study, we aim at presenting an analysis of sperm whale abundance at two temporal scales (hours and months), and how shipping can influence the sperm whale detection rates in both scales.

We define sperm whale acoustic presence (SWP) as a binary variable whose value is 1 if a human specialist is able to identify a sperm whale impulse registered with the AMADEUS system in a user-defined interval of time (in this case, in a segment of 16.8 s, see Zaugg *et al.*<sup>32</sup>). Its value is 0 otherwise. Shipping presence (ISP), expressed as ship clicks, is also defined with the same strategy. Presence is to be understood as presence within the detection range of the observatory. This assignment by a human specialist is a gold standard commonly used to calibrate automated methods. From this point of view, we modelled SWP and ISP as a function of the information provided by the LIDO classification process. This enabled us to automatically predict presence for new segments and correct for possible false positive and false negative impulse situations.

To fit this model we constructed a training data set with 963 segments spread over all hours of the day (at least 30 observations per hour). Initially, 400 segments of 16.8 s were randomly selected from 625 recordings of approximately 67 s (4 segments) each, which were stored every 7 hours from June 2011 to December 2011. Furthermore, 563 segments were added to the data set. These were randomly selected from 490 recordings of >5 min (20 segments) each, which were stored every 2 hours from January 2013 to March 2013. Experienced observers analysed every segment to decide whether sperm whale impulses and impulse shipping noise were present or not (i.e. they assigned SWP and ISP to those segments). Here, we assumed that the environment conditions found in this data were representative of what occurred in other segments, which were not in the training data regardless of the hour, day or month.

We considered a generalised linear model with a binomial response and a logit link function (GLM, Nelder 1972<sup>33</sup>) to predict sperm whale presence (SWP = 0: absence; SWP = 1: presence) as a function of SWI and ISI, including the interaction between the two. A logarithmic transformation was applied both to SWI and to ISI ( $X' = \log[X + 1]$ ). Above a certain threshold the exact value of SWI/ISI is not very important and taking the logarithm reduces the range of the parameters. Similarly, to predict the presence of shipping impulses in a segment (ISP = 0: absence; ISP = 1: presence), a generalised linear model in R was fitted taking into account the information in ISI, SWI and UCI. The same logarithmic transformation used for the first model was applied to the three independent variables. Both models were fitted in R version 2.15.2 (R Core Team 2012). The prediction of the

GLM model is a continuous value between 0 and 1, which can be interpreted as an index of sperm whale acoustic presence.

The main objective of this study was a reliable prediction of sperm whale (acoustic) presence during a particular hour of the day. To make an hourly prediction, SWP was summed over the segments that were analysed during that hour, i.e. a percentage of segments with SWP. In order to obtain a measure of how well the model predicts sperm whale presence, a bootstrap approach was applied to compute a confidence interval around SWP. A thousand bootstrap resamples (drawn with replacement, same size = 963 segments) were obtained from the training data. The model fitted to each of these data sets was then used to estimate SWP on the original training data (summing SWP over all segments).

Sperm whale presence and impulse ship presence predictions were aggregated by month and by hour. Confidence intervals were found using 500 bootstrap samples.

Moreover, SWP predictions in each month were separated in two groups of hours: daylight segments [6 to 18 h) and night segments [18 to 6 h). A Wilcoxon paired test of size 12 (the number of months) was used to determine whether significant differences between the two groups were observed.

**Sperm whale presence predictions as function of SPL (noise level).** The detector reliability was generally weaker when high SPL were found in the recordings. This effect can be caused by a sound source that completely covered the sonar detection band with continuous noise. This phenomenon is called masking. However, a high SPL does not always indicate that sonar was masked. Due to the nature of the SPL computation, a single high-level transient signal can produce a high SPL output computed over a segment. At ANTARES this happened quite often because the neutrino detection lines are calibrated using high-frequency high-level echo sounders that transmit a sequence of pings about every once 100 s (see Fig. 6A and E). On these occasions, there would be no masking of sonar if no other continuous source was present. Segments that contained such a ping were easily identified due to their particular acoustic signature.

In order to obtain more information on the sperm whale behaviour and to see how the SPL affected the daily distribution of SWP, an overview was made showing SWP during each hour of the day and for increasing SPL (shown relative to the yearly median  $\overline{\text{SPL}}$ ).

## References

1. Aguilar, J. A. *et al.* ANTARES: the first undersea neutrino telescope. *Nucl Inst and Met Phys Res A*. **656**, 11–38 (2011a).
2. Aguilar, J. A. *et al.* AMADEUS - The Acoustic Neutrino Detection Test System of the ANTARES Deep-Sea Neutrino Telescope -. *Nucl Inst and Met Phys Res A*. **626–627**, 128–143 (2011b).
3. Ruhl, H. *et al.* Societal need for improved understanding of climate change, anthropogenic impacts, and geo-hazard warning drive development of ocean observatories in European Seas. *Prog Oceanog*. **91**, 1–33 (2011).
4. Tamburini, C. *et al.* Deep-sea bioluminescence blooms after dense water formation at the ocean surface. *PLoS One*. **8**(7), e67523. doi: 10.1371/journal.pone.0067523 (2013).
5. Van Haren, H. *et al.* Acoustic and optical variations during rapid downward motion episodes in the deep North Western Mediterranean. *Deep Sea Res I*. **58**, 875–884 (2011).
6. Van der Graaf, A. J. *et al.* European Marine Strategy Framework Directive - Good Environmental Status (MSFD GES): Report of the Technical Subgroup on Underwater noise and other forms of energy (2012).
7. Hatch, L. T., Clark, C. W., Van Parijs, S. M., Frankel, A. S. & Ponirakis, D. W. Quantifying Loss of Acoustic Communication Space for Right Whales in and around a U.S. National Marine Sanctuary. *Conserv Biol*. **26**, 983–994 (2012).
8. André, M. *et al.* Low-frequency sounds induce acoustic trauma in cephalopods. *Front. Ecol. Environ*. **9**, 489–493 (2011).
9. Solé, M. *et al.* Does exposure to noise from human activities compromise sensory information from cephalopod statocysts? *Deep Sea Res. II*. **95**, 160–181 (2013).
10. Solé, M. *et al.* Ultrastructural damage of *Loligo vulgaris* and *Illex coindetii* statocysts after low frequency sound exposure. *PLoS One* **8**(10), e78825. doi: 10.1371/journal.pone.0078825 (2013).
11. André, M. *et al.* Listening to the Deep: Live monitoring of ocean noise and cetacean acoustic signals. *Mar Pollut Bull*. **63**, 18–26 (2011).
12. Whitehead, H. *Sperm whales: social evolution in the ocean* (The University of Chicago Press, Chicago, 2003).
13. Mohl, B., Wahlberg, M., Madsen, P. T., Heerfordt, A. & Lund, A. The monopulsed nature of sperm whale clicks. *J Acous Soc Am*. **114**, 1143–1154 (2003).
14. André, M., Johansson, T., Delory, E. & van der Schaar, M. Foraging on squid: the sperm whale mid-range sonar. *Jour Mar Biol Assoc*. **87**, 59–67 (2007).
15. Madsen, P., Wahlberg, M. & Möhl, B. Male sperm whale (*Physeter macrocephalus*) acoustics in a high-latitude habitat: implications for echolocation and communication. *Behav Ecol Sociobiol*. **53**, 31, doi: 10.1007/s00265-002-0548-1 (2002).
16. Gannier, A., Drouot, V. & Gould, J. C. Distribution and relative abundance of sperm whales in the Mediterranean Sea. *Mar Ecol Prog Ser*. **243**, 281–293 (2000).
17. Drouot, V., Gannier, A. & Gould, J. C. Summer social distribution of sperm whales in the Mediterranean Sea. *J Mar Biol Ass*. **84**, 675–680 (2004).
18. Pavan, G. *et al.* G. Short Term and Long Term Bioacoustic Monitoring of the Marine Environment. Results from NEMO ONDE Experiment and Way Ahead in *Computational bioacoustics for assessing biodiversity. Proceedings of the International Expert meeting on IT-based detection of bioacoustical patterns* (ed. Frommolt, K. H., Rolf Bardeli, R. & Clausen, M.) 7–14 (Federal Agency for Nature Conservation, Bonn, 2008).
19. Frantzis, A. *et al.* Sperm whale presence off South-West Crete, Eastern Mediterranean Sea in *Proc. 13th Ann. Conf. ECS*. 214–217 (Eur Res Cet, Valencia, 1999).
20. Notarbartolo-Di-Sciara, G. Sperm whales, *Physeter macrocephalus*, in the Mediterranean Sea: a summary of status, threats, and conservation recommendations. *Aquatic Conserv. Mar. Freshw. Ecosyst*. **24**, 4–10. doi: 10.1002/aqc.2409 (2014).
21. Pace, D. S., Mussi, B., Gordon, J. C. D. & Würtz, M. Ecology, Behaviour and Conservation of Sperm Whale in the Mediterranean Sea in *Aquatic Conserv. Mar. Freshw. Ecosyst*. 24 (ed. Wiley, J.) 1–118 (Wiley Online library, 2014).
22. Rendell, L. E. & Frantzis, A. Mediterranean sperm whales, *Physeter macrocephalus*: the precarious state of a lost tribe In *Medit. Mar. Mam. Ecol. Cons.* 75 (ed. Notarbartolo di Sciara, G., Podestà, M. P. & Curry, B. E.) 37–74, doi: 10.1016/bs.amb.2016.08.001 (Advances in Marine Biology, Academic Press/Elsevier, 2016).
23. Di Natale, A. & Notarbartolo di Sciara, G. A review of the passive fishing nets and trap fisheries in the Mediterranean Sea and of the cetacean bycatch In *Gillnets and cetaceans* (ed. Perrin, W. F., Donovan, G. P. & Barlow, J.) 189–202 (Rep Int Whal Comm, 1994).

24. Jaquet, N., Whitehead, H. & Lewis, M. Relationship between sperm whale distribution and primary productivity over large spatial scale in the Pacific ocean. *Eur Res Cet.* **9**, 188–192 (1995).
25. Millot, C. Circulation in the Western Mediterranean Sea. *Oceanol Acta.* **10**, 143–150 (1987).
26. Morel, A. & André, J. M. Pigment distribution and primary production in the Western Mediterranean as derived from coastal zone color scanner observations. *J Geophys Res.* **96**, 2685–12698 (1991).
27. Crépon, M., Wald, L. & Monget, J. M. Low-frequency waves in the Ligurian Sea during December 1977. *J Geophys Res.* **87**, 595–600 (1982).
28. Prieur, L. & Sathyendranath, S. An optical classification of coastal and oceanic waters based on the specific spectral absorption curves of phytoplankton pigments, dissolved organic matter, and other particulate materials. *Limnol Oceanogr.* **26**, 671–89 (1981).
29. Kawakami, T. A review of sperm whale food. *Sci Rep Whales Res Inst.* **32**, 199–218 (1980).
30. Roper, C. F. E. & Young, R. E. Vertical distribution of pelagic cephalopods. *Smithson Contrib Zool.* **209**, 1–51 (1975).
31. Matsushita, T. Daily rhythmic activity of the sperm whales in the Antarctic. *Bull Jpn Soc Sci Fish.* **20**, 770–73 (1955).
32. Zaugg, S. *et al.* Real-time acoustic classification of sperm whale clicks and shipping impulses from deep-sea observatories. *Appl Acoust.* **71**(11), 1011–1019 (2010).
33. Nelder, J. & Wedderburn, R. Generalized linear models. *J R Stat Soc.* **135**, 370–384 (1972).

## Acknowledgements

The authors acknowledge the financial support of the funding agencies: Centre National de la Recherche Scientifique (CNRS), Commissariat à l'énergie atomique et aux énergies alternatives (CEA), la Commission Européenne (FEDER fund and Marie Curie Program), Institut Universitaire de France (IUF), IdEx program and UnivEarthS Labex program at Sorbonne Paris Cité (ANR-10-LABX-0023 and ANR-11-IDEX-0005-02), Région Île-de-France (DIM-ACAV), Région Alsace (contrat CPER), Région Provence-Alpes-Côte d'Azur, Département du Var and Ville de La Seyne-sur-Mer, France; Bundesministerium für Bildung und Forschung (BMBF), Germany; Istituto Nazionale di Fisica Nucleare (INFN), Italy; Stichting voor Fundamenteel Onderzoek der Materie (FOM), Nederlandse organisatie voor Wetenschappelijk Onderzoek (NWO), the Netherlands; Council of the President of the Russian Federation for young scientists and leading scientific schools supporting grants, Russia; National Authority for Scientific Research (ANCS), Romania; Ministerio de Economía y Competitividad (MINECO), Prometeo and Grisolia programs of Generalitat Valenciana and MultiDark, Spain; Agence de l'Oriental and CNRST, Morocco. We also acknowledge the technical support of Ifremer, AIM and Foselev Marine for the sea operation and the CC-IN2P3 for the computing facilities.

## Author Contributions

M.A., M.V.D.S. planned the research and designed the study. M.A., A.C. M.V.D.V., A.S. and S.Z. conducted lab work and analysed the data with help from L.H., A.M.S., J.V.C., M.S., F.-V. and D.D., M.A. and A.C. wrote the article and all authors reviewed it. M.A. provided funding. M.A., A.C., M.V.D.V., A.S., L.H., S.Z., A.M.S., J.V.C., M.S., F.V., D.D., S.A., A.A., M.A.N., G.A., M.A.R., J.A., T.A., B.B., J.B., S.B., V.B., S.B.I., R.B., M.C.B., R.B.R.J.B.R., J.B.U., A.C.A., L.C., J.C., S.C., T.C.H., M.C., A.C.O., R.C., H.C., P.C., A.C.R., A.D., G.D.B., C.D., I.D.P., C.D.O., D.D.O., D.D.R., T.E., I.E.B., D.E., A.E., K.F., I.F., L.A.F., S.G., P.G., S.G.E., K.G., V.G., A.G., H.G., R.G., K.G.R., S.H., H.V.H., A.J.H., Y.H., J.J.H., J.H., J.H.O., C.H., G.I., C.W.J., M.D.J., M.J., M.K., O.K., U.K., D.K., A.K., M.K.R., I.K., V.K., C.L., R.L., D.L., E.L., S.L., M.M., A.M., A.M.A., J.A.M., A.M.A.T., K.M., T.M., P.M., A.M.O., C.M., E.N., G.E.P., C.P., C.P.E., P.P., V.P., T.P., C.R., G.R., K.R., M.S.A., D.F.E.S., M.S.A.N., P.S., J.S., F.S., T.S., C.S., M.S.P., T.H.S., A.S.A., M.T., A.T., M.T.S., D.T., C.T., B.V., C.V., V.V.E., D.V., S.W., J.W., J.D.Z., J.Z. reviewed and edited the manuscript.

## Additional Information

**Competing Interests:** The authors declare no competing financial interests.

**How to cite this article:** André, M. *et al.* Sperm whale long-range echolocation sounds revealed by ANTARES, a deep-sea neutrino telescope. *Sci. Rep.* **7**, 45517; doi: 10.1038/srep45517 (2017).

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017