

## Fin whale acoustic presence and song characteristics in seas to the southwest of Portugal

Andreia Pereira,<sup>1,a)</sup> Danielle Harris,<sup>2</sup> Peter Tyack,<sup>3</sup> and Luis Matias<sup>1</sup>

<sup>1</sup>Instituto Dom Luiz, Faculty of Sciences of University of Lisbon, Campo Grande Edifício C1, Piso 1, Lisbon, 1749-016, Portugal

<sup>2</sup>Centre for Research into Ecological and Environmental Modelling, The Observatory, Buchanan Gardens, University of Saint Andrews, Saint Andrews, Fife KY16 9LZ, Scotland, United Kingdom

<sup>3</sup>Sea Mammal Research Unit, Scottish Oceans Institute, University of Saint Andrews, Saint Andrews Fife KY16 8LB, United Kingdom

### ABSTRACT:

Fin whales were once abundant in the seas to the southwest of Portugal, but whaling activities decreased their numbers considerably. Acoustic data from ocean bottom seismometers provide an opportunity to detect fin whales from their notes, data that would otherwise be logistically challenging and expensive to obtain. Based on inter-note interval and frequency bandwidth, two acoustic patterns produced by fin whales were detected in the study area: pattern 1, described from fin whales in the Mediterranean Sea, and pattern 2, associated with fin whales from the north-east North Atlantic Ocean (NENA). NENA fin whales travel into the western Mediterranean Sea, but the Mediterranean population has not been documented to travel regularly into the NENA. In this study, 11 months of acoustic data recorded southwest of Portugal in the NENA were used to characterize 20-Hz fin whale notes into these patterns. Pattern 2 was the most common and occurred mostly in November–January. Pattern 1 occurred less frequently and mostly in September–December, February and April, which suggested a limited excursion of whales from the Mediterranean Sea. There were also occasions when the two patterns were recorded simultaneously. Results suggest that fin whales from the NENA and Mediterranean Sea might mix in the area during part of the year.

© 2020 Acoustical Society of America. <https://doi.org/10.1121/10.0001066>

(Received 5 August 2019; revised 20 March 2020; accepted 21 March 2020; published online 15 April 2020)

[Editor: Arthur N. Popper]

Pages: 2235–2249

### I. INTRODUCTION

The migration of baleen whales is generally assumed to be a seasonal latitudinal movement with animals feeding during the summer in high-latitude productive areas, also known as “feeding” or “summer” grounds, and spending the winter in tropical or subtropical areas for reproduction, also referred to as “winter” or “breeding” grounds (e.g., Kellogg, 1929). Researchers have suggested several selective pressures to explain baleen whale migration with some pressures related to energetics (Clapham, 2001) and others related to predator avoidance and calf survival (Corkeron and Connor, 1999; Connor and Corkeron, 2001) but the drivers of migration are still a matter of debate. Although the seasonal latitudinal movement theory is very appealing for researchers because it explains baleen whale movement in a simple manner, it does not explain all the types of movement strategies observed for this group of species. Records of humpback whales (e.g., Van Opzeeland *et al.*, 2013), gray whales (e.g., Calambokidis *et al.*, 2002), blue whales (e.g., Branch *et al.*, 2007), and fin whales (e.g., Geijer *et al.*, 2016) show that not all individuals undertake seasonal movement. For example, fin whales have year-round residency in low-latitude areas, such as in the Gulf of California (Urban *et al.*, 2005), the Mediterranean Sea (Notarbartolo di Sciara *et al.*,

2003) and the East China Sea (Mizroch *et al.*, 2009). Fin whales from other populations have also remained year-round around high-latitude areas, such as in the Gulf of Maine, Gulf of St. Lawrence, and Nova Scotia (Delarue *et al.*, 2009). Fin whales show a continuum of migratory strategies that seem to reflect local adaptations of different age-sex classes of whale (Geijer *et al.*, 2016). The movements of fin whales are usually very challenging to investigate because their distribution range extends to offshore areas and because of the difficult logistics needed to study them offshore. Whaling data provided the first descriptions of the long-distance movements of several large cetacean species, including fin whales (e.g., Clapham and Baker, 2002; Smith *et al.*, 2012). Currently, several types of data can be used to investigate movements of large whales such as remote sensing (Read, 2002), stable isotopes (e.g., Witteveen *et al.*, 2009; deHart and Picco, 2015), genetic data (e.g., Kershaw *et al.*, 2016), and photo-identification matches (e.g., Bannister *et al.*, 1999). In particular, acoustic data have been increasingly used to study the presence and movements of calling animals (e.g., Charif and Clark, 2000; Oleson *et al.*, 2006; Širović *et al.*, 2015). Although acoustic datasets can only provide information about acoustically active animals, they have several advantages over the other data types. Acoustic data provide a nonintrusive approach to study movements and behaviour of animals (Mellinger *et al.*, 2007). Instruments can record for extended periods of

<sup>a)</sup>Electronic mail: [afpereira@fc.ul.pt](mailto:afpereira@fc.ul.pt), ORCID: 0000-0002-5368-5707.

time during all hours of the day and through all seasons and weather conditions.

Fin whales produce fairly simple acoustic signals with fundamental frequencies between 15 and 142 Hz, but the most common and the most studied fin whale sound is the “20-Hz note,” also known as the “regular” note (Watkins, 1981; Hatch and Clark, 2004). This signal is a  $\sim 1$  s tone that sweeps downward from 30 to 15 Hz (Watkins *et al.*, 1987). Usually fin whales produce series of 20-Hz notes, called sequences, which are separated by two types of silent periods: rests, lasting between 1 and 20 min, and longer gaps, lasting between 20 min and 2 h (Watkins *et al.*, 1987; Delarue *et al.*, 2009; Soule and Wilcock, 2012). Within a sequence, individual notes are separated by stereotyped intervals (called the inter-note interval, INI), measured from the beginning of a signal to the beginning of the next signal (Watkins *et al.*, 1987). Sequences form bouts (also known as “songs”) that are separated by periods of silence greater than 2 h. Fin whales also produce other low frequency signals, such as the backbeat, which is relatively constant in frequency (between 18 and 20 Hz) and lasts  $\sim 0.8$  s (Clark *et al.*, 2002). Backbeats usually are produced in a repeated series or alternating with 20-Hz notes. The production of the numerous types of fin whale signals and their characteristics vary across geographical areas (e.g., Watkins *et al.*, 1987; Hatch and Clark, 2004; Castellote *et al.*, 2012; Oleson *et al.*, 2014). The 20-Hz note is recorded worldwide, and its temporal and spectral features and bout structure show geographic variation (Hatch and Clark, 2004). Hatch and Clark (2004) found differences in the median frequency of the 20-Hz note between the Gulf of California and other areas of the Pacific and North Atlantic Oceans. Backbeats and higher frequency upsweeps are less frequent and seem to be produced more in the Atlantic Ocean (Hatch and Clark, 2004). In the North Atlantic Ocean, 20-Hz note bouts have been differentiated by their INI (e.g., Delarue *et al.*, 2009), frequency bandwidth (e.g., Thompson *et al.*, 1992), and bout composition, i.e., variability in the types of signal comprising bouts (e.g., Clark and Gagnon, 2004). Of all 20-Hz note features, the INI is the most variable feature in fin whale bioacoustics (Hatch and Clark, 2004; Delarue *et al.*, 2009; Castellote *et al.*, 2012; Oleson *et al.*, 2014). INI patterns have been shown to vary throughout the year in some areas (Morano *et al.*, 2012; Oleson *et al.*, 2014), and in other areas INI patterns that were stable through several years also shifted over longer time periods (Širović *et al.*, 2017). It is not known whether those differences represent individuals that produce a specific INI pattern and move around areas or changes in the acoustic repertoire of individual whales within a population, i.e., individuals that change their INI pattern.

Whaling records of the 20th century show that fin whales were once abundant in the seas to the southwest of Portugal (Sanpera and Aguilar, 1992; Clapham and Hatch, 2000; Clapham *et al.*, 2008). Fin whales were found in dense concentrations, close to shore, and throughout the year, which suggested the presence of a local, nonmigratory

subpopulation (Sanpera and Aguilar 1992; Clapham and Hatch, 2000). For over two centuries, fin whales in this area were heavily hunted (Clapham *et al.*, 2008). Hunting activities ceased in 1960 due to the low profitability (Clapham *et al.*, 2008), and in 1986 the International Whaling Commission (IWC) Moratorium that regulates whaling activities went into effect, and hunting of these whales was officially ceased (Clapham and Baker, 2002). Recent sightings and strandings suggest that fin whale numbers in this area are small compared to the past abundance, and there is no evidence of the presence of the suggested resident subpopulation (Clapham *et al.*, 2008). Since the small number of recent sightings of fin whales have been made offshore, these whales might not be from the suggested resident subpopulation that was heavily exploited during the whaling period.

Recent research after the cessation of whaling shows that there are two groups of fin whales in the northeast North Atlantic Ocean (NENA) and Mediterranean waters (Hatch and Clark, 2004; Panigada and Notarbartolo di Sciara, 2012). According to genetic (Bérubé *et al.*, 1998; Palsbøll *et al.*, 2008), toxicological (Aguilar *et al.*, 2002), and stable isotope (Giménez *et al.*, 2013) data, fin whales from the Mediterranean Sea show a degree of divergence from the NENA fin whales. Historical records of the 19th century show the NENA fin whales entered the Mediterranean Sea (Richiardi, 1874), and throughout the 20th century there was further evidence of this movement (Jonsgård, 1966; Viale, 1985). However, scientists have debated whether Mediterranean fin whales are currently moving into the NENA or not (Castellote *et al.*, 2012; Giménez *et al.*, 2013; Castellote *et al.*, 2014). Mediterranean fin whales have been assumed to be resident (Notarbartolo di Sciara *et al.*, 2003), but the evidence showing that some individuals from the Mediterranean Sea travel toward the NENA is growing (Geijer *et al.*, 2016; Gauffier *et al.*, 2018). However, the frequency and the range of these excursions to the NENA are not clear (Notarbartolo di Sciara *et al.*, 2003). Passive acoustic data collected in the NENA and the Mediterranean Sea suggest the presence of two different acoustic groups of fin whales in these two areas (Clark *et al.*, 2002; Hatch and Clark, 2004; Castellote *et al.*, 2012). Castellote *et al.* (2012) found differences in the INI between 20-Hz notes recorded in the Mediterranean Sea and those recorded in the NENA. There were also differences in the frequency bandwidth of the 20-Hz notes that further suggested differences in the notes. The INI type 1 recorded in the Mediterranean Sea was close to 15 s, and the notes used to calculate this INI showed a frequency bandwidth of 5 Hz (Castellote *et al.*, 2012). The INI type 2 were 12–13 s, and the notes used to calculate this INI showed a frequency bandwidth of 6 Hz. Castellote *et al.* (2012) recorded the INI and note type 2 primarily in the NENA but also in several locations in the western Mediterranean Sea, while INI and note type 1 were only recorded in an area of the Mediterranean Sea further to the east. These results suggest the existence of two acoustic groups that the authors

associated with the NENA and Mediterranean fin whale populations.

If it is assumed that each note type is associated with a subpopulation of fin whales, as suggested by other studies (e.g., [Hatch and Clark, 2004](#); [Delarue et al., 2009](#)), and individual whales do not change from one type to another, then acoustic data can reveal some insights about the movements of these two fin whale populations between the Mediterranean Sea and the NENA. This study provided an opportunity to monitor potential excursions of the subpopulation of fin whales in the Mediterranean Sea to the NENA. In this study, we used 11 months of recordings that were available from an opportunistic source off southwest Portugal to measure the spectral and temporal features of a subsample of recorded 20-Hz note bouts. The spectral and temporal measurements of the 20-Hz notes were used in a hierarchical cluster analysis to categorize different acoustic groups. Possible acoustic groups identified in this analysis were then further validated by comparing their characteristics to the note groups established by [Castellote et al. \(2012\)](#) for fin whales in the NENA and Mediterranean Sea. In a second stage of analysis, the temporal pattern of identified INI(s) was also assessed using the whole dataset of detections in order to obtain an occurrence pattern.

## II. METHODS

### A. Data collection

The data for this paper stem from ocean-bottom seismometers (OBS) deployed as part of the NEAREST (integrated observations from near shore sources of tsunamis: Towards an early warning system) seismic monitoring project, which aimed to investigate the local seismicity and Earth structure in the source region of the Lisbon earthquake and tsunami of 1755 ([Silva, 2017](#)). Each OBS included three channels from the seismometer (two channels for the horizontal components,  $X$  and  $Y$ , and one channel for the vertical component,  $Z$ ) and one channel from a hydrophone ( $H$ ), all with a sampling rate of 100 Hz ([Carrara et al., 2008](#)). [A more detailed description of the deployment and recording instruments is found in Harris et al. \(2013\)](#) and will not be repeated here. The recording period was not equal across all OBS because some components of the seismometers did not work properly at all times, and deployment lengths differed between instruments (Table I). Five OBS were not suitable for the automatic detection process because the  $Z$ -channel, which was the main channel used in the detection process, was not working in two OBS (OBS02 and OBS11), and the other three (OBS21, OBS22, and OBS23) had no detections, including false positives, possibly due to some undetected instrument issue. In addition, local air gun experiments were undertaken during several days in early September ([Somoza et al., 2007](#)). These days with air gun experiments were removed from the dataset because there were long periods of time with high levels of ambient noise, which masked fin whale notes.

### B. Signal detection

The  $Z$ -channel of the seismometer was chosen as the main acoustic recorder for data analysis because the  $Z$ -channel signals generally had a higher signal-to-noise ratio (SNR) than those from the  $H$ -channel, as also observed by [McDonald et al. \(1995\)](#) and [Brodie and Dunn \(2004\)](#). The target signal of the detection process was the 20-Hz fin whale note as described above. The 20-Hz fin whale note was detected using a modified normalized cross correlation of the recordings and a template of the signal waveform. A full description of the detection algorithm is found in [Harris et al. \(2013\)](#) and [Matias and Harris \(2015\)](#). Briefly, in the detection process, the normalized cross correlation between a master template of a single note and sequential 3-s samples of the waveform from the recording are calculated through the time-series. A detection is recorded when the normalized cross correlation between these two time-series reaches a threshold set by the researcher. The template of the 20-Hz note used was obtained from the waveform of a 20-Hz note with high SNR that was observed during a manual inspection of the dataset (recorded in OBS10 on 7 January 2008 at 18:12:17). The detection process was applied to all of the  $Z$ -channel recordings of the OBS using a low correlation threshold (0.2) and a buffer time of 3 s. These parameters were chosen in order to obtain as many notes as possible (the low threshold) but to eliminate the possibility of overlapping notes by using the buffer between detections (the lowest INI observed between two 20-Hz notes was 3 s, therefore a 3-s buffer was adequate to separate detections).

Once a signal was detected using the process from previous studies ([Harris et al., 2013](#); [Matias and Harris, 2015](#)), the spectrogram cross correlation was also calculated following the methodology of [Mellinger and Clark \(2000\)](#) to obtain spectral and temporal measurements of detections. Spectrogram cross correlation requires a spectrogram template of the target signal that is compared with the spectrogram of the recordings. The kernel spectrogram of the 20-Hz note was obtained by averaging the time and frequency characteristics of 39 notes with large amplitude spectral values (between 18 and 25 Hz) and high correlation ( $>0.9$ ) given by the waveform matched filter. When a detection was registered, the algorithm limited a time window starting 1 s before and extending 2 s after the detection time. The spectrograms of detections were computed for successive time windows with a 2.56-s length (256 samples) that were shifted in time by 0.1 s. The spectrum for each time window was computed by fast Fourier transform (FFT) after application of a modified Hanning window. The modified Hanning window left the central 10% of the time-series untouched. This meant that the Hanning window was applied only to 90% of the signal with a 45% taper at each end. At a 100-Hz sampling rate and after the computation of the FFT, the frequency interval in the frequency domain was 0.39 Hz. The spectral power (the square of the FFT amplitude) and spectral cumulative power were computed for each 3 s time window between 12 and 31 Hz in frequency.

TABLE I. Summary of acoustic data available for signal detection and measurement. Darkest gray OBS represent instruments with unavailable Z-channels. Lightest OBS with gridded pattern had the Z-channel available but there was no detection. Asterisks represent the presence of the bouts used in this study.

YEAR	2007				2008							
	S	O	N	D	J	F	M	A	M	J	J	A
OBS 01		*	*				*					
OBS02												
OBS 03	*	*										
OBS 04			*	*	*							
OBS 05		*		*		*	*					
OBS 06			*									
OBS 08												
OBS 09					*							
OBS 10												
OBS11												
OBS 12		*			*	*	*	*				
OBS 13			*			*						
OBS 14												
OBS 15												
OBS 16				*	*							
OBS 17								*				
OBS 18												
OBS 19			*	*			*					
OBS 20							*	*				
OBS 21				*								
OBS 22												
OBS 23												
OBS 24				*								
OBS 25												
MONTH	S	O	N	D	J	F	M	A	M	J	J	A

**C. Signal sampling**

Time constraints precluded verification that all detections were whale notes, and instead a series of rules were established to obtain a subsample small enough to allow manual analysis. The aim of the manual analysis was to ensure accurate spectral and temporal measurements by (1) identifying high quality bouts based on correlation and SNR, (2) checking for the presence of false positives and possible acoustic interference in the notes, and (3) determining whether any notes in the bout were missed by the detector. A cross-correlation value of 0.5 from both the detector and the spectrogram cross correlations and a SNR of 2.0 were used as thresholds to obtain a list of days with potentially high quality bouts. Then, only days with a minimum of 250 automatic detections (with correlation values from 0.2 to close to 1) and at least 20 days with high correlation

values ( $\geq 0.7$ ) were considered. From this list, a random sample of five days for each month was selected. Bouts were separated by 24 h, so no bouts came from the same day.

All hours in the selected days were then manually analyzed using spectrograms calculated with TRITON (Wiggins *et al.*, 2010), a software package written in MATLAB version R2014b. These spectrograms had a window length of 256 samples, a Hanning window taper, and 95% overlap. A detection was considered to be a true positive, i.e., a target 20-Hz fin whale note, when it showed acoustic characteristics in time and frequency similar to those previously published and described in Sec. I. False positives included backbeats and other signals recorded by the OBS not related to fin whale vocal activity such as distant air guns, levelling of the OBS, and other whale species. In order to have the complete structure of the bout, notes missed by the

automatic detection process, probably because of their low SNR, were manually added by the analyst. Notes were only added when INI and SNR were consistent with the ones detected in the automatic detection process for that particular bout. In general, bouts were of high quality and clearly identified in the spectrogram by their consistency of INI and SNR characteristics. Therefore, it was mostly clear when the notes missed by the detector belonged or did not belong to the main bout. There were rare occasions when days had more than one bout. In those cases, only the bout containing the 20-Hz notes with highest SNR (assessed visually) was considered for further analysis. This was done to reduce the risk of detecting the same whale multiple times on the same day and therefore increase the contribution of that individual to the overall note characteristic measurements. However, it was not possible to determine whether the same animal produced more than one bout across days. The 20-Hz note is a fairly simple signal that does not show obvious individual spectral differences, and the INI is also very stable across bouts. We could be sure that more than one whale was calling on some occasions when overlapping bouts were observed. Because the INI did not show large intra-bout variability, it was often possible to follow a specific bout by predicting the times when notes would be produced. However, when two simultaneously occurring bouts showed gaps between notes that were much longer than the INI it became too difficult to separate them. In those cases, the strongest bout was inspected and included until it was not possible to identify the “target” whale.

#### D. Spectral and temporal measurements

Several temporal and spectral features of the detections were extracted based on [Hatch and Clark \(2004\)](#): arrival time, starting frequency, median frequency, note duration, and frequency bandwidth. The cumulative power was calculated for each detection window. A section of the cumulative power window between 5% and 95% was chosen to represent the signal in order to mitigate the influence of noise inside the spectral window. Only the notes with the highest SNR ( $\geq 3.0$ ), high correlation values ( $\geq 0.75$ ), and no overlapping notes from other whales from each bout were used for measurements of the features mentioned above.

The INI was calculated as the difference between detection times obtained in the waveform cross-correlation detection process of each note. There are other time cues, obtained from the spectrogram computation, that are used to calculate the INI such as the time of the median or the maximum energy inside a time window. Most studies use the time of the median energy to calculate the INI (e.g., [Hatch and Clark, 2004](#); [Castellote et al., 2011](#)). There were occasions when 20-Hz notes showed changes in their spectral characteristics as a result of acoustic interference that produced the Lloyd’s mirror effect ([Pereira et al., 2016](#)). These notes were not used in this analysis, but the changes observed in the spectrogram could impact the calculation of the energy and thus impact the INI. A comparative analysis

of the different time cues used to calculate the INI was undertaken (see supplemental material document 2),<sup>1</sup> and since there were no large differences in the INI, the time cue obtained from the initial waveform cross-correlation detection method was maintained. Only INI measurements between 20-Hz notes were considered in the analysis, following the INI definition in [Castellote et al. \(2012\)](#).

Once INI measurements were taken from the selected subsamples, it was possible to search for specific values of the INI in the automatic detections dataset. Therefore, the distribution of INI measurements was calculated across the automated detection dataset with cross correlation  $\geq 0.5$  and SNR  $\geq 7.0$ . A medium cross-correlation threshold (0.5) was used to minimize the occurrence of false positives without constraining the possibility of detecting 20-Hz notes with slightly different acoustic features than our cross-correlation template. The high SNR value ensured that detections were high quality (close to the recording instrument and with low levels of background noise). Here the main aim was to obtain a temporal pattern of occurrence of each INI across the automatic dataset. Because the INI is fairly consistent within bouts, the variability was expected to be low.

The upper and lower limits of the INI pattern(s) were defined as 1.5 standard deviation (SD) around the mean INI(s). The SD limit for the INI(s) was chosen based on the percentage of 20-Hz notes from the subsample with manual verification that fell under this limit (which was 83.13%). The manual inspection of the selected subsample showed that the most common number of consecutive 20-Hz notes was two. Therefore, the pattern for the INI(s) was obtained by the counts of sequences of two detections that showed the expected INI(s). For months where the number of sequences was less than or equal to 34, all detections were inspected in order to check if detections were from 20-Hz notes. In the cases of months with the number of sequences higher than 34, a systematic random subsample of 30 detections was inspected instead. Whenever a detection was uncertain, it was considered a false positive. The false positive rate from this systematic random subsample was then used to adjust the number of sequences for each month. The total number of detections for each month was multiplied by the true positive rate. Since the aim was to obtain a temporal pattern of the INI(s) and not an abundance estimate, we did not consider the percentage of missed notes.

#### E. Data analysis

Cluster analysis was used to investigate whether the spectral and temporal measurements differed between the subsamples of bouts. Cluster analysis has been used in cetacean bioacoustics to identify and classify different types of sounds ([Chabot, 1988](#); [Karlsen et al., 2002](#); [Nemiroff and Whitehead, 2009](#)), assign sounds to social units ([Rendell and Whitehead, 2003](#)), and assess geographical variation in bout structure (e.g., [Delarue et al., 2009](#)). In this study, no pre-classification of the 20-Hz notes was made, i.e., there was no class assignment of area or subpopulation.

Therefore, a hierarchical cluster analysis was undertaken in order to identify different clusters using the *stats* package version 3.4.0 in *R* (R Core Team, 2015). This cluster analysis made two assumptions about the data: (1) the sample was assumed to be representative for the population, and (2) variables were assumed not to be correlated (Hair *et al.*, 2010). Therefore, highly correlated variables needed to be identified in order to assess a possible variable reduction. The linear correlation between spectral and temporal note features was evaluated by calculating the correlation coefficient  $r$  (Upton and Cook, 2014) using all high quality notes from each bout (notes with  $\text{SNR} \geq 3.0$  and correlation value  $\geq 0.75$ ). This ensured a large enough sample size to assess relationships between variables. If two variables were highly correlated ( $r \geq 0.75$ ), the variable that was present in the fewest references about acoustic characterization was removed from the cluster analysis (as suggested by Dawson and Trapp, 2011).

For the cluster analysis, each bout was characterized by the median of the spectral and temporal features. Bouts were considered points in a coordinate space where the observed similarities of the points corresponded to metric distances between them. In some cases, when the variables had large differences in their units, rescaling of the data was performed to give equal importance to all variables. This normalization is usually done by subtracting the mean of a particular variable from all measurements and dividing by the SD (Upton and Cook, 2014). In this study, the data were scaled to give equal importance to all note features in the clustering process. The distance matrix of the bouts was calculated using the Euclidean distance. The average linkage method was chosen as the amalgamation rule, in which the distance between two clusters is defined as the average distance between each point in one cluster to every point in the other cluster. To assess the clustering procedure, the silhouette width (SW) was computed, which reflects the strength of the clusters (Rousseeuw, 1987). Kaufman and Rousseeuw (1990) give some guidance as to the desirable size of the SW. A SW above 0.5 is considered a reasonable classification, and a SW below 0.2 reflects a lack of substantial cluster structure. The optimal number of clusters was evaluated by calculating the average silhouette width (ASW; Rousseeuw, 1987) using the *R* package *factoextra* version 1.0.4 (Kassambara and Mundt, 2017). The ASW assesses the optimal ratio where the intra-cluster distance is minimized and the inter-cluster distance is maximized.

The visual results obtained by this traditional approach were compared with the results from the *R* package *pvclust* version 2.0.0 (Suzuki and Shimodaira, 2015). This package was developed to assess statistical significance in the hierarchical setting (Suzuki and Shimodaira, 2015) by calculating an approximately unbiased  $p$ -value, based on a multistep bootstrap resampling (Shimodaira, 2004). The null hypothesis associated with the  $p$ -value of each cluster states that the cluster does not exist (Suzuki and Shimodaira, 2017). The  $p$ -value is converted to percentage value and represents the proportion of occurrence of a given cluster with the same

members when a number of bootstrap replications (in this case 10 000) are performed. Based on a significance level of  $\alpha = 0.05$ , a  $p$ -value  $> 95\%$  suggests the rejection of the null hypothesis. The same distance matrix and amalgamation rule were used for 10 000 bootstrap replications of the hierarchical clustering.

Nonparametric statistical tests, such as the Kruskal-Wallis test and the *post hoc* pairwise Wilcoxon-Mann-Whitney rank sum test with  $p$ -values from each test adjusted with a Bonferroni correction ( $\alpha = 0.05$ ), were used to assess statistical differences of spectral and temporal features between bouts. The Kruskal-Wallis test was used to perform a comparison among all bouts. The assumption of the Kruskal-Wallis test is that the observations in each bout come from populations with the same distribution. The null hypothesis of the Kruskal-Wallis test is that the mean ranks for each variable are identical between bouts (McDonald, 2008). Wilcoxon-Mann-Whitney rank sum tests were used to compare two bouts at a time if the Kruskal-Wallis test revealed statistical differences among bouts. The sample of each bout consisted of all notes suitable for reliable spectral and temporal measurements as discussed in Sec. IID.

In the hierarchical analysis, two acoustic clusters were defined. The automatic detection process was undertaken using a single template from cluster 2, which was the most common in our dataset. The performance of detecting notes from cluster 1 (bouts manually confirmed and produced in September 2007) was assessed by undertaking an exploratory performance analysis of the detection process with four days when the bouts of cluster 1 were produced and four other random days of the subsample with bouts of cluster 2. In this analysis, the strongest 20-Hz note (high amplitude and high SNR) from cluster 1 was used as a second master template for the waveform correlation processing. Because the spectrograms of those eight days were manually scanned, it was possible to know the total number of 20-Hz notes in each cluster. The detection performances of the cluster 2 template used initially and the second master template from cluster 1 were assessed by calculating the true positive rate, i.e., the number of correct detections of the 20-Hz note divided by the total number of the 20-Hz notes from that cluster.

### III. RESULTS

Between September 2007 and August 2008, a network of 24 OBS was deployed off southwest Portugal (Fig. 1). During visual scans of the recording's spectrograms, several types of acoustic signals were identified: two types of signals from fin whales, 20-Hz notes and backbeats [Fig. 2(A)], notes produced by blue whales [Fig. 2(B)], and non-biological signals such as air guns [Fig. 2(C)] and high intensity signals related with the OBS levelling [Fig. 2(D)].

Between 10 September 2007 and 11 August 2008, a total of 2 918 312 automatic detections of 20-Hz notes were made in recordings of 19 of the 24 OBS. A subsample of these data was used in the following analysis based on the

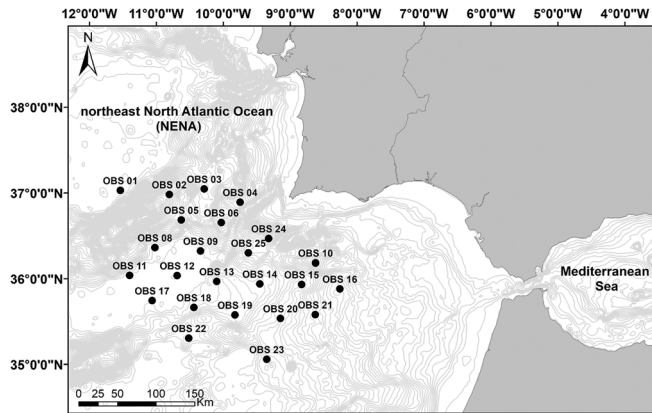


FIG. 1. Location of the seismic monitoring array off southwest Portugal.

days with detections and a set of rules in order to obtain high quality bouts and a good representation of the characteristics of the 20-Hz note. A total of 888 h of sound recordings (37 days) was visually analyzed and resulted in a total of 30 009 20-Hz notes identified and manually confirmed.<sup>1</sup> This included automatic detections that were positively confirmed in the spectrogram and notes that were added to complete each bout.

In almost every day of the recording period, 20-Hz notes were detected, but days with high quality bouts showed a seasonal occurrence. The 37 bouts selected using the criteria mentioned in Sec. II (Methods) were produced between September 2007 and April 2008, which matches the period of highest fin whale vocal activity described in

the literature for the Northern Hemisphere (Watkins *et al.*, 1987; Watkins *et al.*, 2000; Stafford *et al.*, 2007). It also matches the peak for the southwest Mediterranean Sea and Strait of Gibraltar, as described by Castellote *et al.* (2012), and the peak presence in the Strait of Gibraltar as described by Gauffier *et al.* (2018). Between May and August 2008, there were no days with high quality bouts, and automatic detections were reduced considerably. Of the 37 bouts considered, only 32 had enough notes that met the criteria for the measurement analysis, and these were used in the analysis. From the 32 bouts, a total of 4170 20-Hz notes were of high enough quality ( $SNR \geq 3.0$  and correlation values  $\geq 0.75$ ) to be considered for acoustic characterization. The median of the spectral and temporal features of the 20-Hz notes within a bout<sup>1</sup> was used to characterize each bout in the hierarchical cluster analysis. The INI between 20-Hz notes was calculated with the detection times obtained in the waveform cross-correlation detection process of each note. One of the assumptions for cluster analysis is that variables are not correlated (Hair *et al.*, 2010). No correlations between spectral and temporal note features above 0.75 were detected, so all 20-Hz spectral and temporal note features were maintained in the hierarchical cluster analysis. Figure 3 shows the resulting dendrogram from the traditional hierarchical cluster analysis with scaled data. The height of the dendrogram represents the distance between clusters in which larger distances are more conservative. The highest split of the dendrogram, 4.14, divided the bouts into two main clusters. The ASW of 0.41 corroborated these

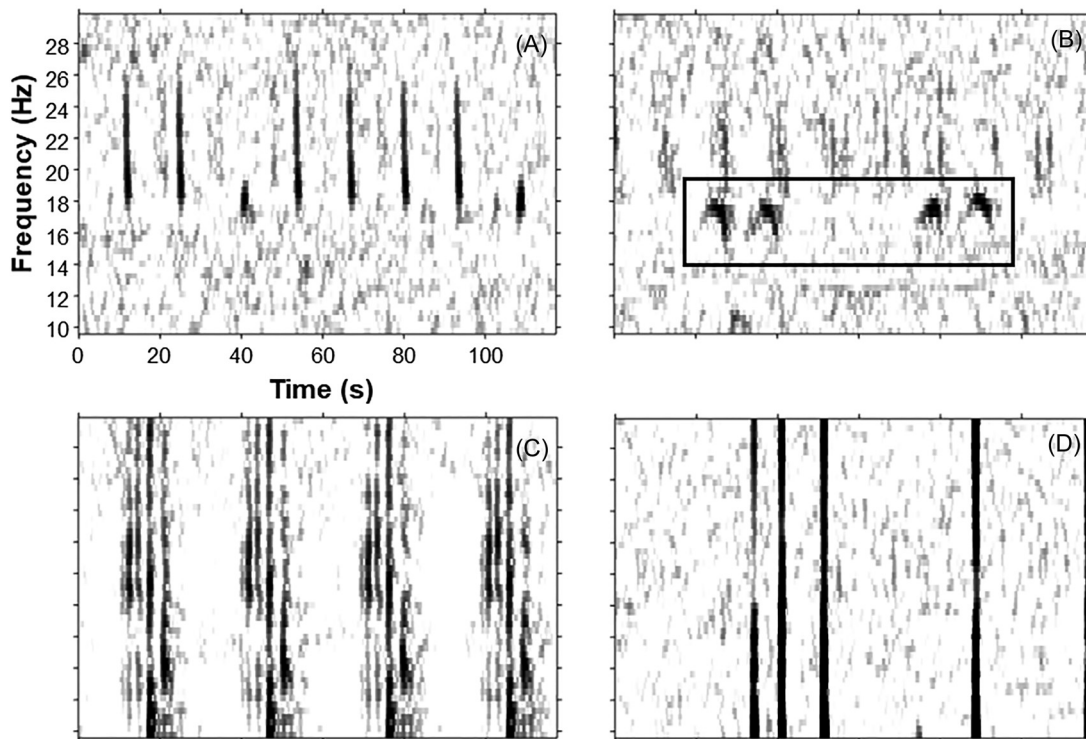


FIG. 2. Examples of spectrograms (120 s) of several types of signals identified during visual scanning of the OBS recordings. (A) Fin-whale signals, backbeats (third and last signals), and 20-Hz notes; (B) blue whale calls (black square); (C) air gun shots; (D) high intensity signals that are related with the OBS levelling. Frame size, 256 samples, 95% overlap, Hanning window, equalized.

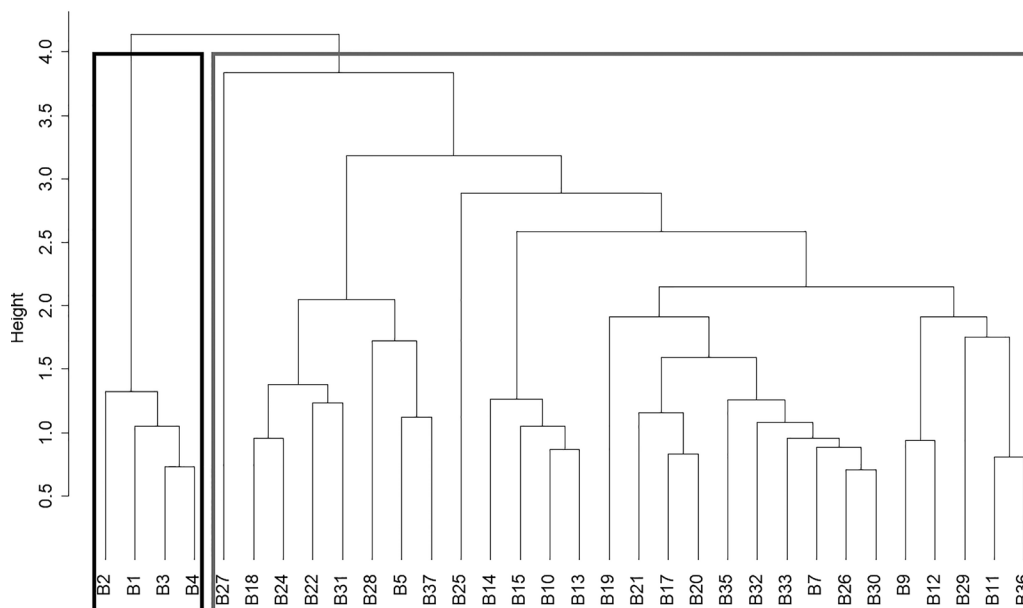


FIG. 3. Dendrogram of fin whale bouts calculated using Euclidean distance and the average linkage method. Bold black and gray lines represent the two clusters validated by the average silhouette width (ASW).

two main clusters. At smaller heights of the dendrogram, bouts could also be grouped into three (height of 3.84) or four (height of 3.18) clusters. However, the ASWs of 0.32 and 0.33 for a structure with 3 or 4 clusters, respectively, were smaller than the structure with two clusters. Therefore, the division of the bouts into two clusters was preferred.

The dendrogram that resulted from bootstrap sampling showed the same two main clusters (Fig. 4). The  $p$ -value ( $p > 95%$ ) for node 18 (Fig. 4, first left-hand branch) indicated the existence of the same cluster of four bouts on the

left side of Fig. 3. The third cluster, indicated by node 29 (Fig. 4, second right-hand branch), in which all of the remaining bouts except one were included, also showed high  $p$ -values ( $p > 95%$ ).

Cluster analysis does not reveal which spectral/temporal measurements contribute to the clustering; therefore, further analysis was required. An exploratory data analysis of the spectral and temporal features showed that the INI [Fig. 5(a)] and frequency bandwidth [Fig. 5(b)] could discriminate between bouts. Kruskal-Wallis nonparametric tests

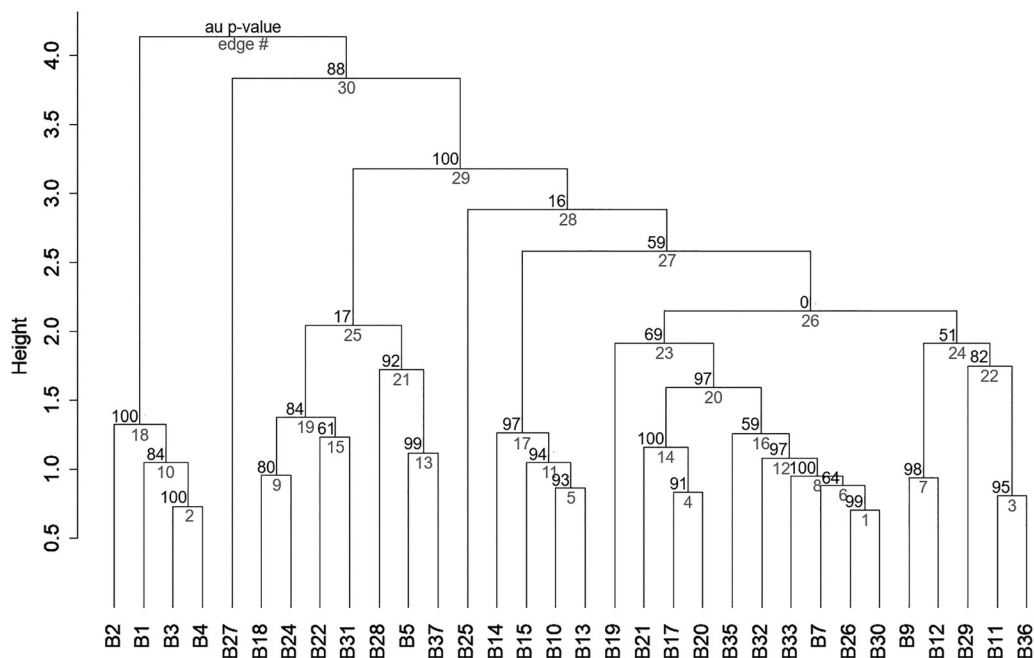


FIG. 4. Dendrogram of fin whale bouts calculated using Euclidean distance and the average linkage method for 10000 bootstrap replications. Top values represent the approximately unbiased  $p$ -value. The lightest gray numbers represent the node number.



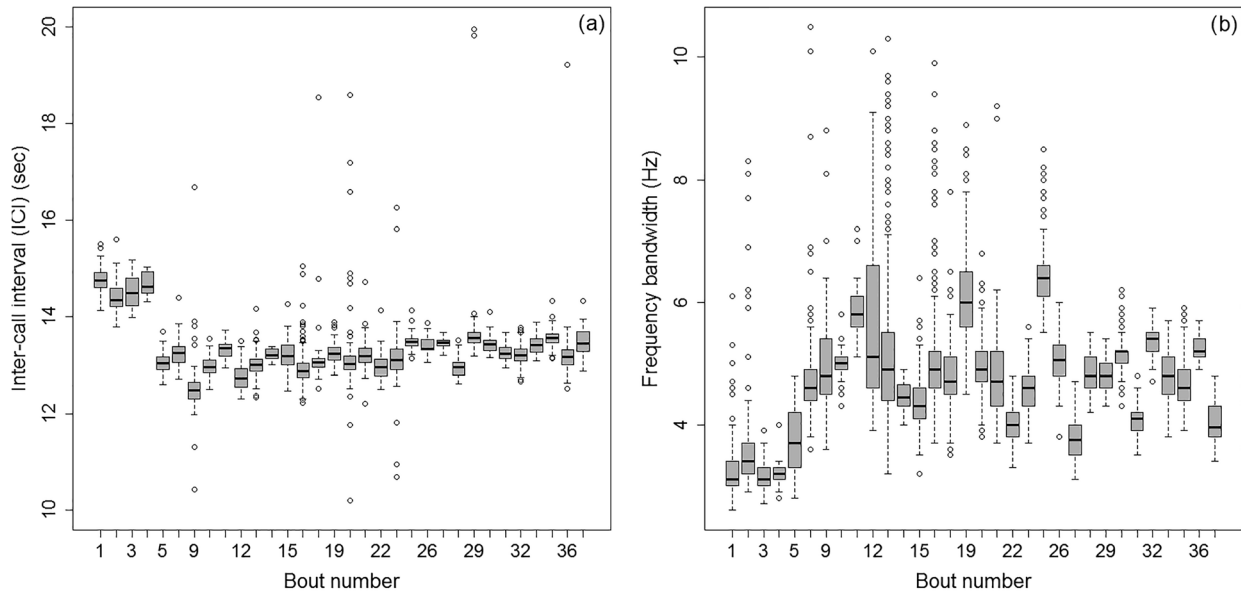


FIG. 5. Box-plots of the (a) INI and (b) frequency bandwidth between bouts. Bouts B6, B8, B16, B23, and B34 were not included due to lack of measurable calls.

suggested differences in the INI ( $H = 1725.5$ ,  $df = 31$ ,  $p$ -value  $< 2.2e-16$ ) and frequency bandwidths ( $H = 2407.4$ , degrees of freedom ( $df$ ) = 31,  $p$ -value  $< 2.2e-16$ ) between bouts. Regarding the INIs, the Wilcoxon-Mann-Whitney tests suggested differences between bouts 1–4 and the rest of the bouts (adjusted  $p$ -values  $< 0.05$ ). The results further suggested differences between bout B1 and B2 ( $W = 2963.5$ , adjusted  $p$ -value =  $2.1e-08$ ) and similarities between B1 and B3 ( $W = 1283.5$ , adjusted  $p$ -value = 0.06) and B4 ( $W = 408$ , adjusted  $p$ -value = 1.00). The test results suggested similarities among bouts B2, B3, and B4 (adjusted  $p$ -values  $> 0.05$ ). Concerning the frequency bandwidth, test results suggested differences between bouts B1, B3, and B4 and all the remaining bouts, even B2. The test results suggested that B2 only showed similarities between B2 and B5 ( $W = 8000.5$ , adjusted  $p$ -value = 0.39).

The hierarchical cluster analysis divided the bouts into two clusters of INIs. When the INI of the 20-Hz note bouts were pooled by cluster, the median INI for cluster 1 was 14.6 s (SD  $\pm 0.3$  s) and the median INI for cluster 2 was 13.1 s ( $\pm 0.5$ ; see Table I of the supplemental material for details about bouts).<sup>1</sup> The two types of INIs were searched across the dataset of detections: Sequences of two detections with the INI between 14.08 and 15.07 s were attributed to cluster 1, and detections with the INI between 12.42 and 13.82 s were attributed to cluster 2.

The search results revealed different occurrence patterns for each INI (Fig. 6). Series of 20-Hz notes with the longer INI occurred less frequently and showed a peak during September 2007 [Fig. 6(a)]. Series of notes with the shorter INI occurred mostly between November 2007 and January 2008 [Fig. 6(b)].

It is important to note that there was one occasion during the validation of the bouts used for the song characterization, recorded on OBS05, one of the most western OBS,

starting at 01:00:00 on 21 October 2007, when two bouts were recorded simultaneously (Fig. 7). One bout showed a longer INI and shorter frequency bandwidth, which was similar to the characteristics of cluster 1, and the other bout showed a shorter INI and longer frequency bandwidth, which were associated with cluster 2.

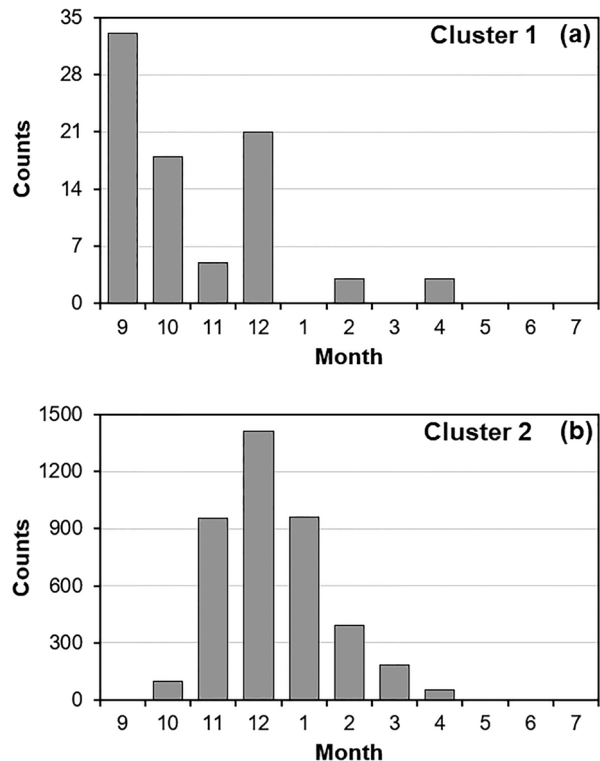


FIG. 6. Distribution of the occurrence of sequences of two detections with (a) intervals between 14.08 and 15.07 s, observed in cluster 1, and (b) intervals between 12.42 and 13.82 s, associated with cluster 2. The number of sequences was obtained from detections of the automatic detection dataset with correlation values  $\geq 0.5$  and SNR  $\geq 7.0$ .

#### IV. DISCUSSION

Two clusters of fin whale note bouts with different sets of acoustic characteristics were identified in a subsample of the OBS recordings. They were distinguished by the median INI and median frequency bandwidth of their bouts: Cluster 1 had a longer median INI of 14.6 s and a lower median frequency bandwidth of 3.2 Hz, and cluster 2 had a shorter median INI of 13.1 s and a higher frequency bandwidth of 4.9 Hz. Cluster 1 was formed from manually verified bouts recorded in September 2007, and cluster 2 was formed from manually verified bouts that were recorded between October 2007 and February 2008. The results of the automatic search of the two INI patterns showed that the INI of cluster 1 also occurred in other periods, and the occurrence of the INI of cluster 2 followed the same seasonal pattern as the one observed with the manually verified days.

##### A. Occurrence of spectral and temporal note characteristics

The manually verified bouts grouped in cluster 1 were recorded during a 12-day period from only one OBS (OBS03) and could have been produced by a single animal that stayed close to the recording instrument during that time. In that case, cluster 1 could be formed by the bout characteristics of a single animal and might not be representative of several individuals. However, since the same INI pattern was recorded at different times even after the 12-day period when cluster 1 bouts were manually measured, it is likely that more than one animal produced this INI. Nonetheless, characteristics of cluster 1 seem to be detected less often than the characteristics of cluster 2.

The detectability of notes with the two different acoustic characteristics was assumed to be the same. Different detection probabilities could impact the results of the seasonality analysis but not the identification of the two acoustic groups. The automatic detection, based on the waveform and spectrogram of the recordings, was performed using a master template obtained from a 20-Hz note of cluster 2. This cluster 2 template might not detect notes from cluster 1 as well as from cluster 2. Indeed, the true positive rate of 20-Hz notes from cluster 1, calculated for the exploratory

performance analysis, was higher when using a template from that cluster in the detection process. Cross-correlation values for the correct detections of the 20-Hz notes of cluster 1 using a template from the same cluster increased, on average, 0.3 (SD  $\pm$  0.08). Although these results suggest that bouts from cluster 1 were possibly underrepresented in the initial subsample of days with potentially good bouts, the main aim of this study, which was the characterization of 20-Hz fin whale notes, was still achieved. Further, the decision to subsample the detections based on SNR and correlation threshold was done assuming that the note characteristics did not vary as a function of these criteria. If the detection of 20-Hz notes from one cluster resulted in mostly low correlation thresholds, then most of them would have not been considered for the subsample. Future work on the two acoustic groups will be undertaken with different sets of templates in order to minimize the bias in the detection process.

A bout with characteristics attributed to cluster 1 was also recorded and manually verified on 21 October 2007 while a bout from cluster 2 was simultaneously produced by another whale in the same area (Fig. 7). Only the cluster 2 bout was formally analyzed because of the established set of rules for sampling and signal measurement. This was the only bout for which a secondary bout had different spectral and temporal characteristics from the primary one. After the identification of the characteristics of the two clusters, it was possible to identify the secondary as having characteristics of cluster 1. The visual confirmation of both clusters occurring simultaneously on the same OBS on the same day shows that these clusters were not always temporally and geographically separated.

The INI and frequency bandwidth measurements obtained in this study showed the same trends found in [Castellote et al. \(2012\)](#) for the two clusters, but there are some differences in the measurements. [Castellote et al. \(2012\)](#) retrieved a large amount of acoustic data from recording instruments deployed in three locations inside the Mediterranean Sea but also from the Strait of Gibraltar and the Azores archipelago. Their study showed a good temporal coverage in strategic areas where the two types of INIs and frequency bandwidths could potentially occur. Some of the data used in [Castellote et al. \(2012\)](#) were recorded during

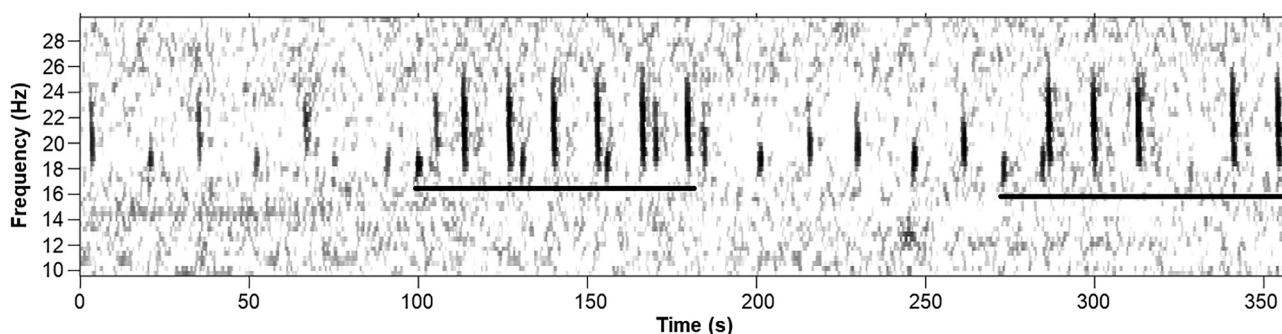


FIG. 7. Spectrogram showing several minutes when two whales with distinct note features attributed to each cluster were calling at the same time. Black lines indicate the whale associated with cluster 2. The whales were recorded on 21/10/2007 around 01:00:00 on OBS05. Frame size, 256 samples, 95% overlap, Hanning window, equalized.

the same time period as this study. More specifically, the recordings in [Castellote \*et al.\* \(2012\)](#) from the Azores and Alboran basin were obtained between November 2007 and May 2008. This study used data recorded between September 2007 and August 2008. Since both studies recorded data during the same time, it was expected that the recorded 20-Hz note characteristics would be similar, thus the observed differences need further consideration.

The different values found for the NENA in both studies could be related to unreported details in the signal processing such as differences in the window size used to measure the features. Some studies indicate intra-annual ([Oleson \*et al.\*, 2014](#)) and interannual ([Morano \*et al.\*, 2012](#)) differences in the INI. This could complicate the use of the INI as the main discriminatory feature of acoustic groups and its comparison between areas. Differences in INIs and frequency bandwidths could also be related to differences in geographical locations of the recordings. The acoustic characteristics of fin whales recorded in the Mediterranean Sea and NENA have previously been assessed by [Clark \*et al.\* \(2002\)](#), [Hatch and Clark \(2004\)](#), and [Castellote \*et al.\* \(2012\)](#). This study further identified the INI and frequency bandwidth as significant acoustic features to define acoustic groups among fin whales ([Hatch and Clark 2004](#); [Castellote \*et al.\*, 2012](#)). Comparing the results of this study with the current literature, the INI value of cluster 1 matched the INI recorded inside the Mediterranean Sea ([Clark \*et al.\*, 2002](#); [Hatch and Clark 2004](#); [Castellote \*et al.\*, 2012](#)), and the INI value of cluster 2 matched the INI recorded in the NENA ([Hatch and Clark, 2004](#); [Castellote \*et al.\*, 2012](#)). In addition, cluster 1 showed a larger INI and a shorter frequency bandwidth than cluster 2. This same trend, a shorter frequency bandwidth associated with a larger INI and a larger frequency bandwidth associated with a shorter INI, was also found for fin whales recorded in the Mediterranean Sea and NENA ([Castellote \*et al.\*, 2012](#)). The INIs and frequency bandwidths presented in [Castellote \*et al.\* \(2012\)](#) for the Azores matched the ones obtained in the Alboran basin, 12.7 s and around 6.8 Hz, respectively. However, the INI and frequency bandwidth from the Strait of Gibraltar were slightly different, 13.1 s and 5.8 Hz, respectively ([Castellote \*et al.\*, 2012](#)), and were more similar to the characteristics of cluster 2 obtained in this study. Despite these differences, [Castellote \*et al.\* \(2012\)](#) grouped the characteristics from these three areas (Azores, Alboran Basin, and the Strait of Gibraltar) together in the NENA cluster. Further analysis needs to be undertaken in order to assess if these smaller scale differences are representative of potential separate acoustic characteristics between the NENA and the Mediterranean Sea that could be associated with a separate subpopulation present in the Strait of Gibraltar during the whaling period as suggested by [Clapham \*et al.\* \(2008\)](#).

## **B. Comparison with other studies that characterize notes of fin whales in the NENA and Mediterranean Sea**

The acoustic characteristics associated with cluster 1 had only been recorded inside the Mediterranean Sea until

now. This study provides the first evidence of acoustic characteristics associated with Mediterranean fin whales recorded in the NENA. The acoustic characteristics associated with cluster 2 have been recorded in the NENA and the western Mediterranean Sea ([Castellote \*et al.\*, 2012](#)) but not in the eastern Mediterranean. This study shows an extension of the geographical overlap to the NENA of the two acoustic clusters of fin whales that were previously observed in [Castellote \*et al.\* \(2012\)](#) in the western Mediterranean Sea. This study also showed a temporal overlap of the two acoustic clusters in the manual inspection of a high quality subsample and in the automatic search of the two INI patterns. A temporal overlap of singers from the NENA and Mediterranean Sea was previously recorded in the Balearic Basin, inside the Mediterranean Sea ([Castellote \*et al.\*, 2011](#); [Castellote \*et al.\*, 2014](#)). These results further suggest that fin whales from the two subpopulations might travel to the same areas and are both present in the study area during a part of the year. However, the acoustic characteristics associated with NENA fin whales were recorded more often than the acoustic characteristics associated with Mediterranean fin whales in the study area.

Cluster 1 was recorded during several months of the study—between September and December 2007 and between February and April 2008. [Gauffier \*et al.\* \(2018\)](#) performed an extensive study of the movements of fin whales between the Mediterranean Sea and NENA combining data from direct observations, vessel-, and land-based surveys with photo-identification over a period of 15 years. Between May and October, they found a directional movement toward the NENA of all of the whales observed during that period. They also observed 69% of the sighted fin whales moving from the NENA toward the Mediterranean Sea between November and April. Considering the directional movements found in [Gauffier \*et al.\* \(2018\)](#) and the seasonal presence of the acoustic characteristics found in this study, the results suggest that fin whales from the NENA may travel more frequently between the NENA and the Mediterranean Sea than Mediterranean fin whales may swim to the NENA. From the 37 bouts that were manually analyzed in this study, only four showed acoustic characteristics associated with Mediterranean fin whales. The distribution of the INI associated with Mediterranean fin whales also showed low occurrence, although this could also be related to the detection process as mentioned above.

Based on [Gauffier \*et al.\* \(2018\)](#), some individuals of the Mediterranean subpopulation may exit the Mediterranean basin between May and October, and they may stay in the NENA until April. The limited April records of the INI of cluster 1, obtained from the automatic search, could correspond to these few Mediterranean fin whales returning from the NENA to the Mediterranean basin. The lack of acoustic detections during the summer suggests that both the NENA and Mediterranean fin whales either stop producing 20-Hz notes or leave the study area. Multiple sources of visual data from sightings ([Husum, 2011](#); [Verborgh, 2012](#); [Boisseau, 2014](#)) and strandings ([Sousa, 2012](#)) found in the OBIS-

SEAMAP dataset (Halpin *et al.*, 2009) show the presence of fin whales off southwest Portugal during the time of recordings with no acoustic presence. This suggests that they may not produce 20-Hz notes there during the summer.

Croll *et al.* (2002) suggest that note bouts are only produced by male fin whales. Since the production of fin whale bouts is seasonal and peaks during the fall and winter months (Thompson and Friedl, 1982; Watkins *et al.*, 2000; Stafford *et al.*, 2007), which coincide with the breeding season, it is suggested that these signals are used as advertisement to attract females (Payne and Webb, 1971; Croll *et al.*, 2002; Clark and Gagnon, 2004). Based on this information and the seasonal analysis of the INI patterns, our study area might be used for breeding purposes since we recorded acoustic characteristics associated with the NENA and few Mediterranean fin whales in the study area with a peak during the fall for Mediterranean characteristics and a peak during the winter for NENA characteristics. The breeding areas of fin whales in the Mediterranean Sea are still unknown, but several authors suggest the North African continental shelf as a breeding site (Marini *et al.*, 1996; Notarbartolo di Sciara *et al.*, 2003). Although there are no suggested breeding areas for NENA fin whales around the study area, the whaling records of a year-round presence of fin whales suggest they might potentially use this area for breeding purposes as well. The visual data mentioned above suggest that non-calling fin whales also use the study area during the summer, but the identity of the population of fin whales present is undetermined. Fin whales have shown feeding activities in middle latitude areas and for extended periods of time that can last into the summer (e.g., Silva *et al.*, 2013). The Gorringer Bank is a volcanic seamount that is located in the southwest of the study area and has been defined as a Site of Community Interest in accordance with European Union legislation. The topographic features of seamounts affect the movement of water and, in some occasions, can promote local productivity (Genin, 2004). In the case of baleen whales, seamounts can provide increased feeding opportunities (Silva, 2016). The results of a hydrodynamic model applied to circulation and primary production around the area of the Gorringer Bank pointed to the development of a large anticyclonic water circulation with associated upwelling of deep nutrient rich water (Coelho and Santos, 2002). Field observations made during the summer were consistent with the modelling results (Coelho and Santos, 2002). These results suggest that non-calling fin whales that could be present during the summer in the study area and may use the area to forage.

Future management actions should account for the possible repopulation of the pre-whaling fin whale habitat and the growing evidence of overlap between NENA and Mediterranean fin whales.

### C. Conclusions

Acoustic data provided evidence of the presence of at least two acoustic groups of fin whales from September to

April in the seas to the southwest of Portugal. The acoustic cluster associated with the NENA was more commonly present in the study area with a peak between November and January. Few records of the acoustic cluster initially recorded in the Mediterranean Sea by other authors were also detected, mainly during September, which could suggest an extension of the distribution of the Mediterranean fin whale subpopulation. Although the occurrence peaks of the two acoustic clusters were different, there were occasions with temporal overlap of the two clusters. The results presented in this study suggest that the movements between the NENA and Mediterranean Sea may be undertaken mostly by NENA fin whales visiting the Mediterranean basin from spring to fall in accordance with previously proposed movement patterns (Castellote *et al.*, 2012). Only a small fraction of whales from the Mediterranean Sea may travel to the NENA during the summer. Further studies need to be developed in order to document the extent of these movements. A multidisciplinary approach that includes genetic, acoustic, and movement analysis would make a considerable contribution to the study of fin whales between the Mediterranean Sea and North Atlantic Ocean. Acoustic recorders could be deployed between September and November and in April off southwest Portugal to record calling fin whales in the area that could be in transit. Additional acoustic records could also be collected in other areas central to the NENA population such as in the Azores. The time periods selected here match the timing of the movements observed in Gauffier *et al.* (2018). After the deployment and during the time of the acoustic recording, biopsies could be taken in order to identify the subpopulation of a sample of calling individuals. Acoustic recordings could also enable the tracking of calling animals, and the directionality of the movement of the sampled individuals could be assessed. However, tags could also be deployed to a sample of animals in the case of non-vocalizing animals. The existing OBS dataset presented in this study is a valuable initial resource to assess the movement of calling fin whales because the instruments can also estimate tracks of calling animals (McDonald *et al.*, 1995; Wilcock, 2012; Harris *et al.*, 2013). A future task will be to identify tracks using the detections from the existing OBS dataset and calculate different movement parameters, such as directionality, for each of the two song types identified in this study. The assessment of genetic differences between NENA and Mediterranean fin whales (Bérubé *et al.*, 1998; Palsbøll *et al.*, 2008) associated with acoustic recordings could provide crucial information to confirm the proposed relationship of song type and population origin.

### ACKNOWLEDGMENTS

The NEAREST project was conducted on behalf of the European Union (EU) Specific Programme “Integrating and Strengthening the European Research Area,” Sub-Priority 1.1.6.3, “Global Change and Ecosystems,” Contract No. 037110. The authors would like to thank Tiago Marques for his helpful comments on the analysis of this paper and

Francisco Martinho for assistance in the manual verification of the spectrograms and data processing. The data analysis was carried out as part of the Ph.D. thesis by A.P. The thesis was funded by Fundação para a Ciência e a Tecnologia (FCT) through a Ph.D. scholarship SFRH/BD/52554/2014 from the Doctoral Programme—PD/143/2012—Lisbon Doctoral School on Earth System Science, Instituto Dom Luiz (IDL). A.P. was also funded by the Project AWARENESS (Project No. PTDC/BIA-BMA/30514/201) and FCT through Project No. UIDB/50019/2020—IDL. D.H. was funded by the Office of Naval Research (ONR; Award No. N00014-14-1-0394). P.T. acknowledges funding from ONR Award No. N00014-15-1-2553 and the MASTS pooling initiative (The Marine Alliance for Science and Technology for Scotland), and their support is gratefully acknowledged. MASTS is funded by the Scottish Funding Council (Grant Reference HR09011) and contributing institutions. The authors would like to thank two anonymous reviewers for their valuable comments and suggestions that improved the manuscript.

<sup>1</sup>See supplementary material at <https://doi.org/10.1121/10.0001066> for the summary of spectral and temporal characteristics of each bout and the comparative analysis of several time cues to calculate the INI.

- Aguilar, A., Borrell, A., and Reijnders, P. J. H. (2002). "Geographical and temporal variation in levels of organochlorine contaminants in marine mammals," *Mar. Environ. Res.* **53**, 425–452.
- Bannister, J. L., Pastene, L. A., and Burnell, S. R. (1999). "First record of movement of a southern right whale (*Eubalaena australis*) between warm water breeding grounds and the Antarctic Ocean, south of 60°S," *Mar. Mammal Sci.* **15**, 1337–1342.
- Bérubé, M., Aguilar, A., Dendanto, D., Larsen, F., Notarbartolo di Sciara, G., Sears, R., Sigurjónsson, J., Urban-Ramirez, J., and Palsbøll, P. J. (1998). "Population genetic structure of North Atlantic, Mediterranean Sea and Sea of Cortez fin whales, *Balaenoptera physalus* (Linnaeus, 1758); analysis of mitochondrial and nuclear loci," *Mol. Ecol.* **7**, 585–599.
- Boisseau, O. (2014). "Visual sightings from Song of the Whale 1993–2013, data downloaded from OBIS-SEAMAP," available at <http://seamap.env.duke.edu/dataset/1158> (Last viewed 22 November 2019).
- Branch, T. A., Stafford, K. M., Palacios, D. M., Allison, C., Bannister, J. L., Burton, C. L. K., Cabrera, E., Carlson, C. A., Galletti Vernazzani, B., Gill, P. C., Hucce-Gaete, R., Jenner, K. C. S., Jenner, M. N. M., Matsuoka, K., Mikhalev, Y. A., Miyashita, T., Morrice, M. G., Nishiwaki, S., Sturrock, V. J., Tormosov, D., Anderson, R. C., Baker, A. N., Best, P. B., Borsari, P., Brownell, R. L., Jr., Childerhouse, S., Findlay, K. P., Gerrodette, T., Ilangakoon, A. D., Joergensen, M., Kahn, B., Ljungblad, D. K., Maughan, B., Mccauley, R. D., McKay, S., Norris, T. F., Oman Whale and Dolphin Research Group, Warneke, R. M., Rankin, S., Samaran, S., Thiele, D., and Van Waerebeek, K. (2007). "Past and present distribution, densities and movements of blue whales *Balaenoptera musculus* in the southern hemisphere and northern Indian Ocean," *Mammal Rev.* **37**, 116–175.
- Brodie, D. C., and Dunn, R. A. (2014). "Low frequency baleen whale calls detected on ocean-bottom seismometers in the Lau basin, southwest Pacific Ocean," *J. Acoust. Soc. Am.* **137**, 53–62.
- Calambokidis, J., Darling, J. D., Deecke, V., Gearin, P., Gosho, M., Megill, W., Tombach, C. M., Goley, D., Toropova, C., and Gisborne, B. (2002). "Abundance, range and movements of a feeding aggregation of gray whales (*Eschrichtius robustus*) from California to southeastern Alaska in 1998," *J. Cetacean Res. Manage.* **4**, 267–276.
- Carrara, G., Matias, L., Geissler, W., D'Oriano, F., Lagalante, M., Cianchini, G., Chierini, F., Cuffaro, M., Diaconov, A., Doormann, U., Favali, P., Feld, C., Geissler, W., Gerber, H., Gossler, J., Hansen, M., Innocenzi, L., Labahn, E., Langner, W., Lo Bue, N., Riminucci, F., Romsdorf, M., Salocchi, A., Unglert, K., Veneruso, M., Wolter, R. J., and Zitellini, N. (2008). "NEAREST 2008 cruise preliminary report R/V Urania," available at <http://nearest.bo.ismar.cnr.it/documentation/nearest-2008-preliminary-report> (Last viewed 28 February 2019).
- Castellote, M., Álvarez, C., Laran, S., La Manna, G., Gazo, M., Manghi, M., and Lammers, M. (2011). "Further acoustic insights into the Mediterranean fin whale migration and population identity," in *25th Annual Conference of the European Cetacean Society*, March, Cadiz, Spain.
- Castellote, M., Clark, C. W., and Lammers, M. O. (2012). "Fin whale (*Balaenoptera physalus*) population identity in the western Mediterranean Sea," *Mar. Mammal Sci.* **28**, 325–344.
- Castellote, M., Clark, C. W., and Lammers, M. O. (2014). "Dispersal of North Atlantic fin whales (*Balaenoptera physalus*) into the Mediterranean Sea and exchange between populations: Response to Giménez *et al.*," *Rapid Commun. Mass Spectrom.* **27**, 1801–1806, [Rapid Commun. Mass Spectrom.](#) **28**, 665–667.
- Chabot, D. (1988). "A quantitative technique to compare and classify humpback whale (*Megaptera novaeangliae*) sounds," *Ethology* **77**, 89–102.
- Charif, R. A., and Clark, C. W. (2000). "Acoustic monitoring of large whales off north and west Britain and Ireland: A two year study, October 1996–September 1998," JNCC Report No. 313, available at <http://jncc.defra.gov.uk/pdf/jncc313web.pdf> (Last viewed 5 August 2019).
- Clapham, P. (2001). "Why do baleen whales migrate? A response to Corkeron and Connor," *Mar. Mammal Sci.* **17**, 432–436.
- Clapham, P. J., Aguilar, A., and Hatch, L. T. (2008). "Determining spatial and temporal scales for management: Lessons from whaling," *Mar. Mammal Sci.* **24**, 183–201.
- Clapham, P. J., and Baker, C. S. (2002). "Modern whaling," in *Encyclopedia of Marine Mammals*, edited by W. F. Perrin, B. Würsig, and J. G. M. Thewissen (Academic, San Diego, CA), pp. 1328–1332.
- Clapham, P. J., and L. T. Hatch (2000). "Determining spatial and temporal scales for population management units: Lessons from whaling," International Whaling Commission document SC/52/ID2, Adelaide, Australia.
- Clark, C. W., Borsari, F., and Notarbartolo di Sciara, G. (2002). "Vocal activity of fin whales, *Balaenoptera physalus*, in the Ligurian Sea," *Mar. Mammal Sci.* **18**, 281–285.
- Clark, C. W., and Gagnon, G. J. (2004). "Low-frequency vocal behaviors of baleen whales in the North Atlantic: Insights from integrated undersea surveillance system detections, locations, and tracking from 1992 to 1996," *U.S. Navy J. Underw. Acoust.* **52**, 609–640.
- Coelho, H., and Santos, R. (2002). "Enhanced primary production over sea-mounts: A numerical study," paper presented at the *4th Symposium on the Iberian Atlantic Margin*, July, Vigo, Spain.
- Connor, R. C., and Corkeron, P. J. (2001). "Predation past and present: Killer whales and baleen whale migration," *Mar. Mammal Sci.* **17**, 436–439.
- Corkeron, P. J., and Connor, R. C. (1999). "Why do baleen whales migrate?," *Mar. Mammal Sci.* **15**, 1228–1245.
- Croll, D., Clark, C. W., Acevedo, A., Tershy, B., Flores, S., Gedamke, J., and Urban, J. (2002). "Only male fin whales sing loud songs," *Nature* **417**, 809.
- Dawson, B., and Trapp, R. G. (2011). *Basic and Clinical Biostatistics* (McGraw-Hill Education, New York). 416 pp.
- deHart, P. A. P., and Picco, C. M. (2015). "Stable oxygen and hydrogen isotope analyses of bowhead whale baleen as biochemical recorders of migration and arctic environmental change," *Polar Sci.* **9**, 235–248.
- Delarue, J., Todd, S. K., VanParijs, S. M., and DiIorio, L. (2009). "Geographic variation in Northwest Atlantic fin whale (*Balaenoptera physalus*) song: Implications for stock structure assessment," *J. Acoust. Soc. Am.* **125**, 1774–1782.
- Gauffier, P., Verborgh, P., Giménez, J., Esteban, R., Sierra, J. M. S., and de Stephanis, R. (2018). "Contemporary migration of fin whales through the Strait of Gibraltar," *Mar. Ecol. Prog. Ser.* **588**, 215–228.
- Geijer, C. K. A., Notarbartolo di Sciara, G., and Panigada, S. (2016). "Mysticete migration revisited: Are Mediterranean fin whales an anomaly?," *Mammal Rev.* **46**, 284–296.
- Genin, A. (2004). "Bio-physical coupling in the formation of zooplankton and fish aggregations over abrupt topographies," *J. Mar. Syst.* **50**, 3–20.

- Giménez, J., Gómez-Campos, E., Borrell, A., Cardona, L., and Aguilar, A. (2013). "Isotopic evidence of limited exchange between Mediterranean and eastern North Atlantic fin whales," *Rapid. Commun. Mass Spectrom.* **27**, 1801–1806.
- Hair, J. F., Black, W. C., Babin, B. J., and Anderson, R. E. (2010). *Multivariate Data Analysis—A Global Perspective* (Pearson Education, Upper Saddle River, NJ), 816 pp.
- Halpin, P. N., Read, A. J., Fujioka, E., Best, B. D., Donnelly, B., Hazen, L. J., Kot, C., Urian, K., LaBrecque, E., Dimatteo, A., Cleary, J., Good, C., Crowder, L. B., and Hyrenbach, K. D. (2009). "OBIS-SEAMAP: The world data center for marine mammal, sea bird, and sea turtle distributions," *Oceanography* **22**, 104–115.
- Harris, D., Matias, L., Thomas, L., Harwood, J., and Geissler, W. H. (2013). "Applying distance sampling to fin whale calls recorded by single seismic instruments in the Northeast Atlantic," *J. Acoust. Soc. Am.* **134**, 3522–3535.
- Hatch, L. T., and Clark, C. W. (2004). "Acoustic differentiation between fin whales in both the North Atlantic and North Pacific Oceans, and integration with genetic estimates of divergence," Report of the International Whaling Commission SC/56/SD6, 37 pp.
- Husum, A. M. (2011). "HMAP—History of marine animal populations," available at <http://seamap.env.duke.edu/dataset/103150007> (Last viewed 22 November 2019).
- Jonggård, Å. (1966). "Biology of the north Atlantic fin whale *Balaenoptera physalus* (L.): Taxonomy, distribution, migration and food," *Hvalråd. Skr.* **49**, 1–62.
- Karlsen, J., Bisther, A., Lydersen, C., Haug, T., and Kovacs, K. (2002). "Summer vocalisations of adult male white whales (*Delphinapterus leucas*) in Svalbard, Norway," *Polar Sci.* **25**, 808–817.
- Kassambara, A., and Mundt, F. (2017). "factoextra: Extract and visualize the results of multivariate data analyses [software package]," available at <http://www.sthda.com/english/rpkgs/factoextra/index.html> (Last viewed 23 January 2019).
- Kaufman, L., and Rousseeuw, P. J. (1990). *Finding Groups in Data: An Introduction to Cluster Analysis* (Wiley, New York), 368 pp.
- Kellogg, R. (1929). "What is known of the migrations of some of the whalebone whales," *Smithsonian Inst. Ann. Rep.* **1929**, 467–494.
- Kershaw, F., Carvalho, I., Loo, J., Pomilla, C., Best, P. B., Findlay, K. P., Cerchio, S., Collins, T., Engel, M. H., Minton, G., Ersts, P., Barendse, J., Kotze, P. G. H., Razafindrakoto, Y., Nguouessono, S., Mejer, M., Thornton, M., and Rosenbaum, H. C. (2016). "Multiple processes drive genetic structure of humpback whale (*Megaptera novaeangliae*) populations across spatial scales," *Mol. Ecol.* **26**, 977–994.
- Marini, L., Consiglio, C., Angradi, A. M., Catalano, B., Sanna, A., Valentini, T., Fioino, M. G., and Villetti, G. (1996). "Distribution, abundance and seasonality of cetaceans sighted during scheduled ferry crossings in the central Tyrrhenian Sea: 1989–1992," *Ital. J. Zool.* **63**, 381–388.
- Matias, L., and Harris, D. (2015). "A single-station method for the detection, classification and location of fin whale calls using ocean-bottom seismic stations," *J. Acoust. Soc. Am.* **138**, 504–520.
- McDonald, J. H. (2008). *Handbook of Biological Statistics* (Sparky House, Baltimore, MD), 287 pp.
- McDonald, M. A., Hildebrand, J. A., and Webb, S. C. (1995). "Blue and fin whales observed on a seafloor array in the Northeast Pacific," *J. Acoust. Soc. Am.* **98**, 712–721.
- Mellinger, D. K., and Clark, C. W. (2000). "Recognizing transient low-frequency whale sounds by spectrogram correlation," *J. Acoust. Soc. Am.* **107**, 3518–3529.
- Mellinger, D. K., Stafford, K. M., Moore, S. E., Dziak, R. P., and Matsumoto, H. (2007). "An overview of fixed passive acoustic observation methods for cetaceans," *Oceanography* **20**, 36–45.
- Mizroch, S. A., Rice, D. W., Zwiefelhofer, D., Waite, J., and Perryman, W. L. (2009). "Distribution and movements of fin whales in the North Pacific Ocean," *Mammal Rev.* **39**, 193–227.
- Morano, J. L., Salisbury, D. P., Rice, A. N., Conklin, K. L., Falk, K. L., and Clark, C. (2012). "Seasonal and geographical patterns of fin whale song in the western North Atlantic Ocean," *J. Acoust. Soc. Am.* **132**, 1207–1212.
- Nemiroff, L., and Whitehead, H. (2009). "Structural characteristics of pulsed calls of long-finned pilot whales *Globicephala melas*," *Bioacoustics* **19**, 67–92.
- Notarbartolo di Sciara, G., Zanardelli, M., Jahoda, M., Panigada, S., and Airoldi, S. (2003). "The fin whale *Balaenoptera physalus* (L. 1758) in the Mediterranean Sea," *Mammal Rev.* **33**, 105–150.
- Oleson, E. M., Širović, A., Bayless, A. R., and Hildebrand, J. A. (2014). "Synchronous seasonal change in fin whale song in the North Pacific," *PLoS One* **9**, e115678.
- Oleson, E. M., Wiggins, S. M., and Hildebrand, J. A. (2006). "Temporal separation of blue whale call types on a southern California feeding ground," *Anim. Behav.* **74**, 881–894.
- Palsbøll, P. J., Bérubé, M., Aguilar, A., Notarbartolo di Sciara, G., and Nielsen, R. (2008). "Discerning between recurrent gene flow and recent divergence under a finite-site mutation model applied to north Atlantic and Mediterranean sea fin whale (*Balaenoptera physalus*) populations," *Evolution* **58**, 670–675.
- Panigada, S., and Notarbartolo di Sciara, G. (2012). "Balaenoptera physalus," in *IUCN 2017, IUCN Red List of Threatened Species*, version 2017.3, available at <http://www.iucnredlist.org> (Last viewed 31 January 2019).
- Payne, R. S., and Webb, D. (1971). "Orientation by means of long range acoustic signaling in baleen whales," *Ann. N. Y. Acad. Sci.* **188**, 110–141.
- Pereira, A., Harris, D., Tyack, P., and Matias, L. (2016). "Lloyd's mirror effect in fin whale calls and its use to infer the depth of vocalizing animals," *Proc. Meet. Acoust.* **27**, 070002.
- R Core Team (2015). "R: A language and environment for statistical computing [software package]," The R Foundation for Statistical Computing, Vienna, Austria, available at <http://www.R-project.org/> (Last viewed 31 January 2019).
- Read, A. J. (2002). "Telemetry," in *Encyclopedia of Marine Mammals*, edited by W. F. Perrin, B. Würsig, and J. G. M. Thewissen (Academic, San Diego, CA), pp. 1232–1235.
- Rendell, L., and Whitehead, H. (2003). "Vocal clans in sperm whales (*Physeter macrocephalus*)," *Proc. Biol. Sci.* **270**, 225–231.
- Richiardi, S. (1874). "On the individual variations of the *Balaenoptera musculus*" *Atti d. Soc. Toscana di Sc. Nat.* **1**, 246–259.
- Rousseeuw, P. J. (1987). "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," *J. Comput. Appl. Math.* **20**, 53–65.
- Sanpera, C., and Aguilar, A. (1992). "Modern whaling off the Iberian peninsula during the 20th century," *Rep. Int. Whaling Comm.* **42**, 723–730.
- Shimodaira, H. (2004). "Approximately unbiased tests of regions using multistep-multiscale bootstrap resampling," *Ann. Stat.* **32**, 2616–2641.
- Silva, M. (2016). "Role of oceanic island and seamount ecosystems to baleen whales," in *30th Annual Conference of the European Cetacean Society*, March, Funchal, Portugal.
- Silva, M. A., Prieto, R., Jonsen, I., Baumgartner, M. F., and Santos, R. S. (2013). "North Atlantic blue and fin whales suspend their spring migration to forage in middle latitudes: Building up energy reserves for the journey?," *Plos One* **8**, e76507.
- Silva, S. D. M. M. F. (2017). "Strain partitioning and the seismicity distribution within a transpressive plate boundary: SW Iberia-NW Nubia," Ph.D. thesis, University of Lisbon, Lisbon, Portugal, 399 pp.
- Širović, A., Oleson, E. M., Buccowich, J., Rice, A., and Bayless, A. R. (2017). "Fin whale song variability in southern California and the Gulf of California," *Sci. Rep.* **7**, 10126.
- Širović, A., Rice, A., Chou, E., Hildebrand, J. A., Wiggins, S. M., and Roch, M. A. (2015). "Seven years of blue and fin whale call abundance in the Southern California Bight," *Endanger. Species Res.* **28**, 61–76.
- Smith, T. D., Reeves, R. R., Josephson, E. A., and Lund, J. N. (2012). "Spatial and seasonal distribution of American whaling and whales in the age of sail," *PLoS One* **7**, e34905.
- Somoza, L., Anahnah, F., Bohoyo, F., González, J., Hernández, J., Iliev, I., León, R., Llave, E., Maduro, C., Martínez, S., Pérez, L. F., and Vázquez, T. (2007). "Informe científico-técnico, N/O L'ATALANTE" ("Scientific and technical report, R/V L'Atalante"), August 23 to September 9, 2007, project report, available at [http://tierra.rediris.es/moundforce/Moundforce\\_informe\\_final.pdf](http://tierra.rediris.es/moundforce/Moundforce_informe_final.pdf) (Last viewed 28 February 2019).
- Soule, D. C., and Wilcock, J. S. (2013). "Fin whale tracks recorded by a seismic network on the Juan de Fuca Ridge, Northeast Pacific Ocean," *J. Acoust. Soc. Am.* **133**, 1751–1761.
- Sousa, A. (2012). "Historical strandings of cetaceans on the Portuguese coast," available at <http://seamap.env.duke.edu/dataset/829> on 2016-08-04 (Last viewed 22 November 2019).
- Stafford, K. M., Mellinger, D. K., Moore, S. E., and Fox, C. G. (2007). "Seasonal variability and detection range modeling of baleen whale calls in the Gulf of Alaska, 1999–2002," *J. Acoust. Soc. Am.* **122**, 3378–3390.
- Suzuki, R., and Shimodaira, H. (2015). "pvclust: Hierarchical clustering with p-values via multiscale bootstrap resampling (version 2.0-0)

- [software package],” available at <http://CRAN.R-project.org/package=pvclust> (Last viewed 31 January 2019).
- Suzuki, R., and Shimodaira, H. (2017). “pvclust—An R package for hierarchical clustering with *p*-values,” available at <http://stat.sys.i.kyoto-u.ac.jp/prog/pvclust/> (Last viewed 1 August 2019).
- Thompson, P. O., and Friedl, W. A. (1982). “A long-term study of low frequency sound from several species of whales off Oahu, Hawaii,” *Cetology* **45**, 1–19.
- Thompson, P. O., Findley, L. T., and Vidal, O. (1992). “20-Hz pulses and other vocalizations of fin whales, *Balaenoptera physalus*, in the Gulf of Mexico,” *J. Acoust. Soc. Am.* **92**, 3051–3057.
- Upton, G., and Cook, I. (2014). *A Dictionary of Statistics* (Oxford University Press, Oxford, UK), 496 pp.
- Urban, R. J., Rojas-Bracho, L., Guerrero-Ruiz, M., Jaramillole Gorreta, A., and Findley, L. T. (2005). “Cetacean diversity and conservation in the Gulf of California,” in *Biodiversity, Ecosystems, and Conservation in Northern Mexico*, edited by J. E. Cartron, G. Ceballos, and R. S. Felger (Oxford University Press, New York), pp. 276–297.
- Van Opzeeland, I., Van Parijs, S., Kindermann, L., Burkhardt, E., and Boebel, O. (2013). “Calling in the cold: Pervasive acoustic presence of humpback whales (*Megaptera novaeangliae*) in Antarctic coastal waters,” *PLoS One* **8**, e73007.
- Verborgh, P. (2012). “CIRCE marine mammals off Spain 2001–2012,” available at <http://seamap.env.duke.edu/dataset/840> on 2016-08-04 (Last viewed 22 November 2019).
- Viale, D. (1985). “Cetaceans in the northwestern Mediterranean: Their place in the ecosystem,” *Oceanogr. Mar. Biol.* **23**, 491–571.
- Watkins, W. A. (1981). “Activities and underwater sounds of fin whales,” *Sci. Rep. Whales Res. Inst.* **93**, 83–117.
- Watkins, W. A., Daher, M. A., Reppucci, G. M., George, J. E., Martin, D. L., DiMarzio, N. A., and Gannon, D. P. (2000). “Seasonality and distribution of whale calls in the North Pacific,” *Oceanography* **13**, 62–67.
- Watkins, W. A., Tyack, P., and Moore, K. E. (1987). “The 20-Hz signals of finback whales (*Balaenoptera physalus*),” *J. Acoust. Soc. Am.* **82**, 1901–1912.
- Wiggins, S. M., Roch, M. A., and Hildebrand, J. A. (2010). “TRITON software package: Analyzing large passive acoustic monitoring data sets using MATLAB,” *J. Acoust. Soc. Am.* **128**, 2299.
- Wilcock, W. S. D. (2012). “Tracking fin whales in the northeast Pacific Ocean with a seafloor seismic network,” *J. Acoust. Soc. Am.* **132**, 2408–2419.
- Witteveen, B. H., Worthy, G. A. J., and Roth, J. D. (2009). “Tracing migratory movements of breeding North Pacific humpback whales using stable isotope analysis,” *Mar. Ecol. Prog. Ser.* **393**, 173–283.