Underwater vocal complexity of Arctic seal *Erignathus barbatus* in Kongsfjorden (Svalbard)

Ignazio Parisi,[1,a)] Giovanni de Vincenzi,[1,a)] Marco Torri,[2] Elena Papale,[1]
Salvatore Mazzola,[1] Angelo Bonanno,[2] and Giuseppa Buscaino[1,b)]
[1]*National Research Council, Institute for Coastal Marine Environment, Bioacousticslab Capo Granitola, Via del Mare, Campobello di Mazara (TP), 6-91021 Torretta Granitola, Italy*
[2]*National Research Council, Institute for Coastal Marine Environment, Via del Mare, Campobello di Mazara (TP), 6-91021 Torretta Granitola, Italy*

(Received 16 June 2017; revised 5 October 2017; accepted 26 October 2017; published online 20 November 2017)

In this study the description of underwater vocal repertoire of bearded seal in Svalbard (Norway) was extended. Two autonomous passive acoustic recorders were deployed for one year (August 2014–July 2015) in the inner and outer parts of the Kongsfjorden, and 1728 h were recorded and 17 220 vocalizations were found. Nine different vocalization classes were identified and characterized using ten acoustic parameters. The calls showed heterogeneous spectral features, but share the descending trend of frequency modulation. The different classes emerged were discriminated primarily by bandwidth and duration, and then by minimum frequency, central frequency, and maximum frequency in this order. This study represents a step forward to improve the understanding of the acoustic behaviour and the social function of these calls, and identified long passive acoustic monitoring as an effective method to assess vocal complexity and the ecology of marine species producing sounds. © *2017 Acoustical Society of America.* https://doi.org/10.1121/1.5010887

[KL] Pages: 3104–3115

I. INTRODUCTION

Bearded seals are circumpolar and boreal-Arctic pinnipeds that maintain a close association with sea-ice for critical life history activities such as reproduction, pupping, and moulting (Moore and Huntington, 2008; Nelson *et al.*, 1984). The role of the bearded seal in the Arctic trophic chain is of great importance, considering especially that, together with ringed seal (*Pusa hispida*), bearded seals are the primary food resource of polar bear (*Ursus maritimus*) (Derocher, 2005; Derocher *et al.*, 2002). During the reproduction period in spring time, bearded seal males produce elaborate calls for the purpose of advertising breeding condition (Cleator, 1996; Cleator *et al.*, 1989; Cleator and Stirling, 1990; Risch *et al.*, 2007), and/or to maintain territories (Van Parijs *et al.*, 2003; Van Parijs and Clark, 2006). Van Parijs *et al.* (2001) monitored the spatial and temporal patterns of vocalizing bearded seal males in Kongsfjorden (Svalbard Islands, Norway), attesting that the vocal complexity of this species is limited to four classes: trill, sweep, moan, and flat tone. However, in other Arctic regions like Alaska (Cleator *et al.*, 1989; Jones *et al.*, 2014; Risch *et al.*, 2007), Canadian high Arctic (Cleator *et al.*, 1989; Risch *et al.*, 2007), and western Canadian Arctic (Risch *et al.*, 2007), *Erignathus barbatus* is known to have a wider vocal repertoire. These vocalizations share common features among different sites, like the trilling trend, but other characteristics, such as the presence of ascent/plume or the sweep trend, are peculiarities belonging only to some Arctic populations (Risch *et al.*, 2007).

Information about the complexity of species-specific vocal repertoire, in term of its acoustical parameters, is a useful method to improve knowledge on animal behaviour, and the sum of acoustical and behavioural data can improve knowledge on the ecological role of communication. The production of a wide vocalization repertoire, characterized by vocalizations with internal spectral variability, may underlie the transmission of different communicative significance (Brudzynski, 2010). Further, the study of differentiated vocalizations of a species is also crucial to identify stocks or populations on the basis of regional dialects, providing information on movement and association patterns of species (e.g., Charrier *et al.*, 2013; McDonald *et al.*, 2006; Noad *et al.*, 2000; Papale *et al.*, 2017; Sciacca *et al.*, 2015).

In the Arctic region, because of its extreme weather and darkness in winter months, boat-based surveys are limited and the ability to carry out aerial surveys is highly restricted (Cleator and Stirling, 1990; Van Opzeeland, 2010). In these extreme conditions the possibility of using passive acoustic monitoring (PAM) has already permitted us to investigate various biological and ecological factors, such as the abundance estimate or the spatial and temporal distribution of marine mammals (e.g., MacIntyre *et al.*, 2013; MacIntyre *et al.*, 2015; Stirling *et al.*, 1983), as well as their acoustic behaviour (Simon *et al.*, 2010). In contrast to visual observation, acoustic recorders can be operated autonomously and under conditions where visual observation is not possible (Azzolin *et al.*, 2014; Sousa-Lima *et al.*, 2013; Van Opzeeland, 2010), allowing high-resolution sampling for long periods (up to one year) with a spatial coverage

---

[a)] Also at Association eConscience, Via Provinciale, 610–90046 Monreale (PA), Italy.

[b)] Electronic mail: giuseppa.buscaino@cnr.it depending on the recording system and the monitored sound source level (for the bearded seal precedent study of Cleator *et al.*, 1989, assessed a mean spatial coverage of 5 km with a peak up to 20 km).

In this study, it long PAM was used to improve our knowledge on the acoustic behaviour of bearded seals in Svalbard, extending the recording period as compared to that of previous studies. A quantitative acoustic analysis of bearded seals vocalizations was performed and recorded during one-year monitoring in two sites of the Kongsfjorden (North West Svalbard Islands). Here, we provide a calls classification with detailed descriptions of each call type.

II. METHODS

A. Study area and acoustic data collection

The study area is located at 78°–79° north and 11°–12° east, within Kongsfjorden (Western Svalbard Islands, Norway). The hydrographic conditions of the fjord are dominated by the mixing of cold Arctic waters, warmer Atlantic water masses, and melting ice (Cottier *et al.*, 2005). Since tidewater glacial fronts act as upwelling areas for zooplankton (Svendsen *et al.*, 2002), they attract a number of top predators (Lydersen *et al.*, 2014). Data were collected all year-round from August 2014 to July 2015 through autonomous passive acoustic recorders (SM2, Wildlife Acoustics, Concord, MA), deployed in two sites (Fig. 1). The first (hereinafter called glacier-site) was deployed in the Svalbard Archipelago, in front of the Kronebreen glacier (one of the larger fast-flowing and actively calving marine-terminating glacier streams in Svalbard Islands) at a distance of about 4 km from its front (Fig. 1, coordinates 78°54.74′ N–12°24.31′ E); the other (hereinafter open-site) was deployed 27 km from the Kronebreen glacier at the mouth of the fjord (79°03.22′ N–11°32.83′E).

The hydrophone had a recording bandwidth of 25 Hz–150 kHz with a sensitivity of -170 ± 5 dB re 1 V/μPa. Recorders were located at 75 m depth, at 8 m from the bottom, anchored with a 35 kg ballast, and kept vertical by a small buoy. In order to avoid any noise due to the moving parts, the connections between the recorders, the acoustic release and the ballast, as well as with the buoy were made up of non-metallic ropes. Sampling frequency was set at 48 kHz with a resolution of 16 bits and no pre-amplification or filtering were applied (except for antialiasing filter, applied automatically by the recorder) during the recordings. Data were collected by setting a 50% duty cycle: the first 30 min of every hour were recorded for each day (24 h). The recorders were recovered for maintenance every 4 months to change batteries and storage memory. A subsample of 6 days per month was randomly selected for the analysis. A total of 1728 h of acoustic data were examined (864 h for each site).

B. Bearded seal vocal repertoire analysis

Each acoustic data file of 30 min was visually examined through the spectrograms view [fast Fourier transform (FFT) size 2048 point, Hann Window, sampling frequency 48 000, frequency band 0–24 kHz, resolution 23 Hz and 42 ms] in RX5 (iZotope, Cambridge, MA), audio editor software, to detect the presence of bearded seal vocalizations. For the acoustic analysis, only acoustic events with well-defined frequency contours, allowing unambiguous measurements of the acoustic parameters, were chosen. The analysis was started by adapting the four call types previously described by Van Parijs *et al.* (2001). These four classes (trill, sweep, moan, and flat tone) were distinguished basing on the variability in their duration, bandwidth, and spectrogram contour. Because in our data a highest variability of signals occurred, other vocal classes were added based on differences in duration (D_t), bandwidth (BW), and frequency modulation (FM). The use of these parameters permitted to underline nine differentiable vocalization classes.

To corroborate the new classification, the random forest (RF) ensemble method was performed using the following

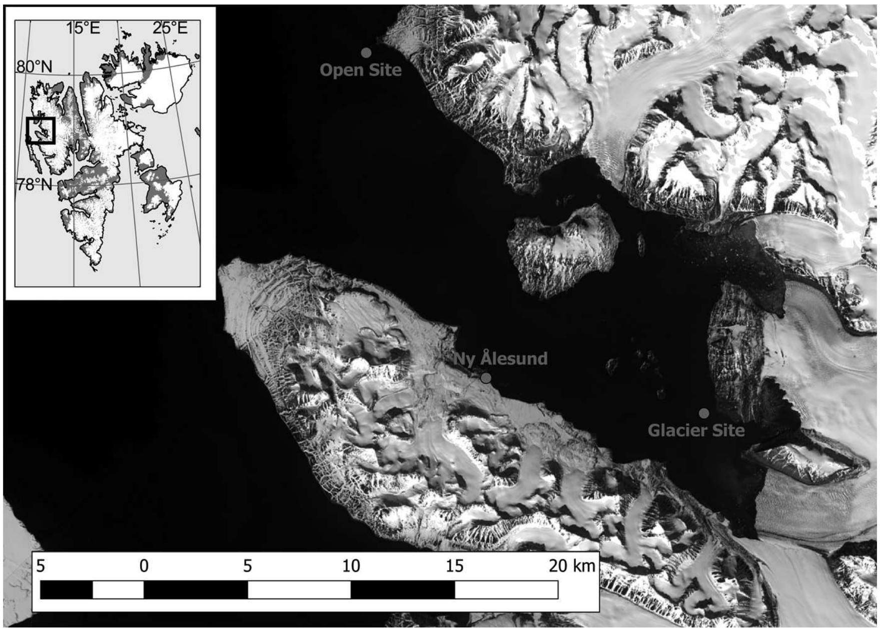


FIG. 1. (Color online) Study area: In the inset the Svalbard Islands are represented, while in the principal image Kongsfjorden, Western Svalbard, with the position of the recorders sites (red dots) and the Ny-Ålesund harbour (yellow dot) is shown. (Credit: U.S. Geological Survey, Department of the Interior/USGS.)

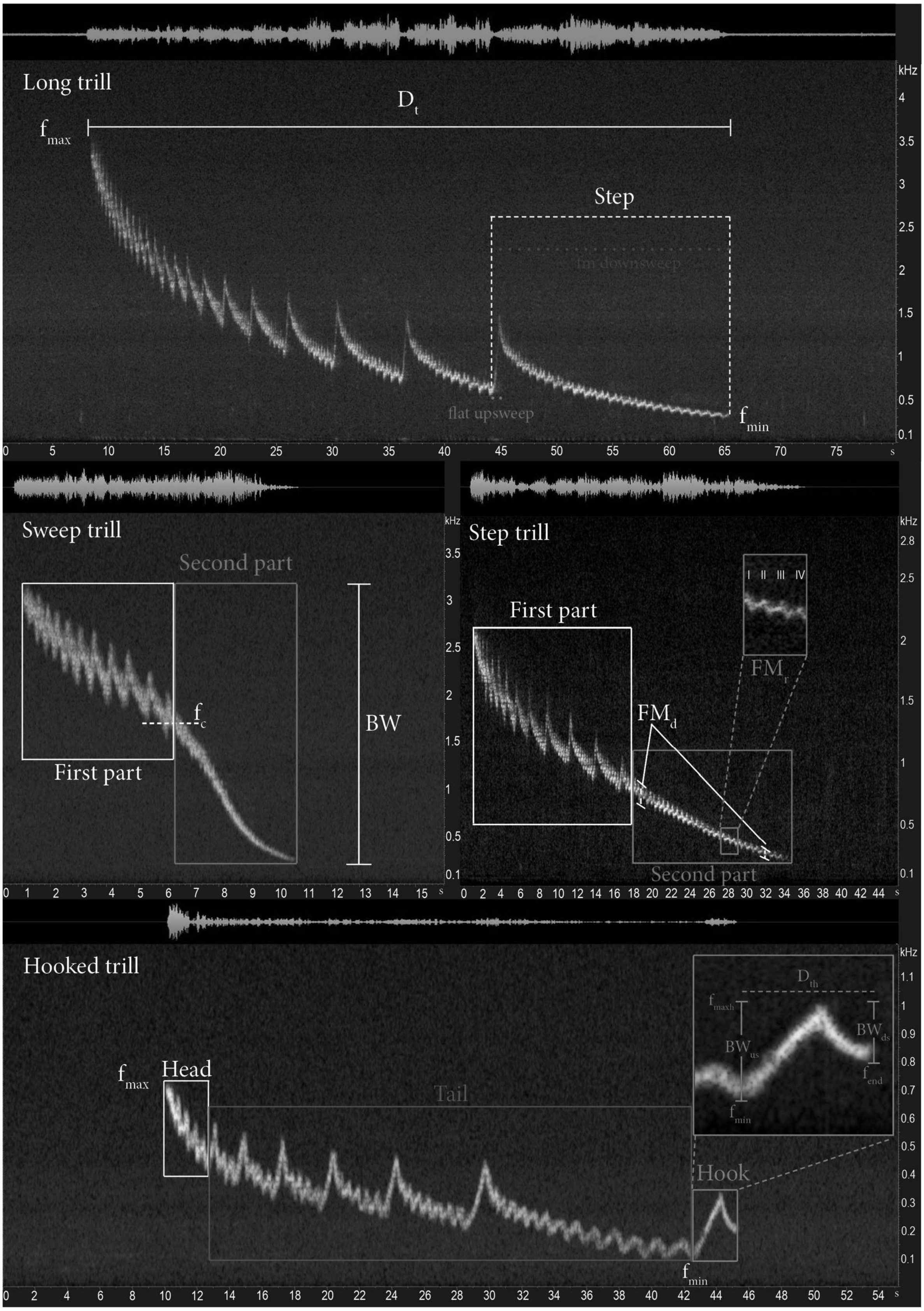


FIG. 2. (Color online) Spectrograms (FFT size 2048 point; Hann Window) and oscillogram (blue line) of the vocalizations of *E. barbatus* recorded in Kongsfjorden. (a) Long trill, sweep trill, step trill, and hooked trill; (b) moan, linear trill, strange trill, deep trill, slope trill, and linear sweep; (c, top) a group of moan with strange trill events; (c, bottom) sequence of five short vocalizations resembling the groans (red ellipse) previously described by Cleator *et al.*, 1989. D_t, total duration; BW, bandwidth, $f_{\min}$, minimum frequency; $f_{\max}$, maximum frequency; f_c, central frequency; FM_r, frequency modulation rate; FM_d, frequency modulation range; f_{mid}, frequency at $D_t/2$; Harm, harmonics; Head, first portion of vocalization; Tail, portion of vocalization after head; Step, up flat-sweep followed by FM downsweep.

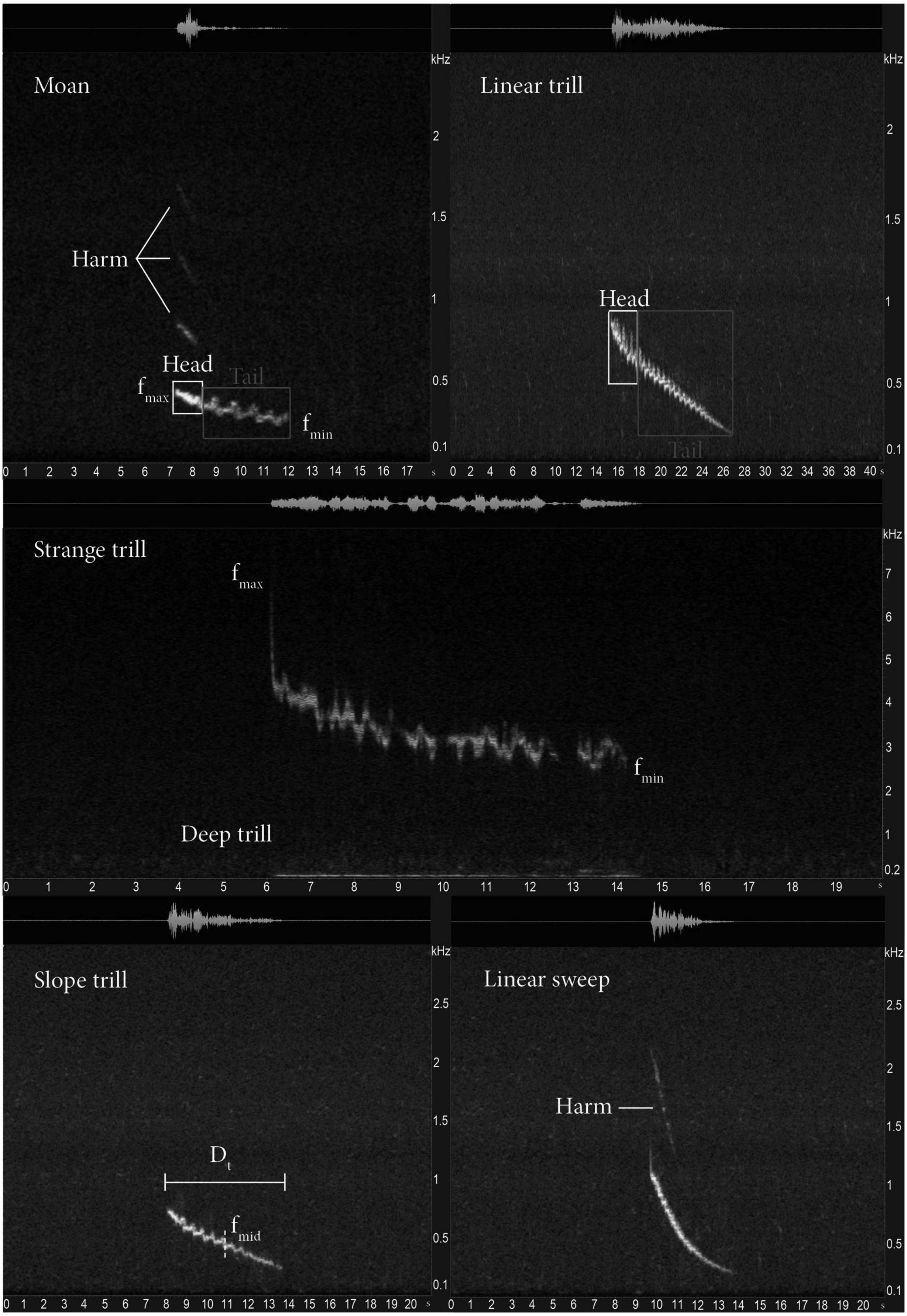
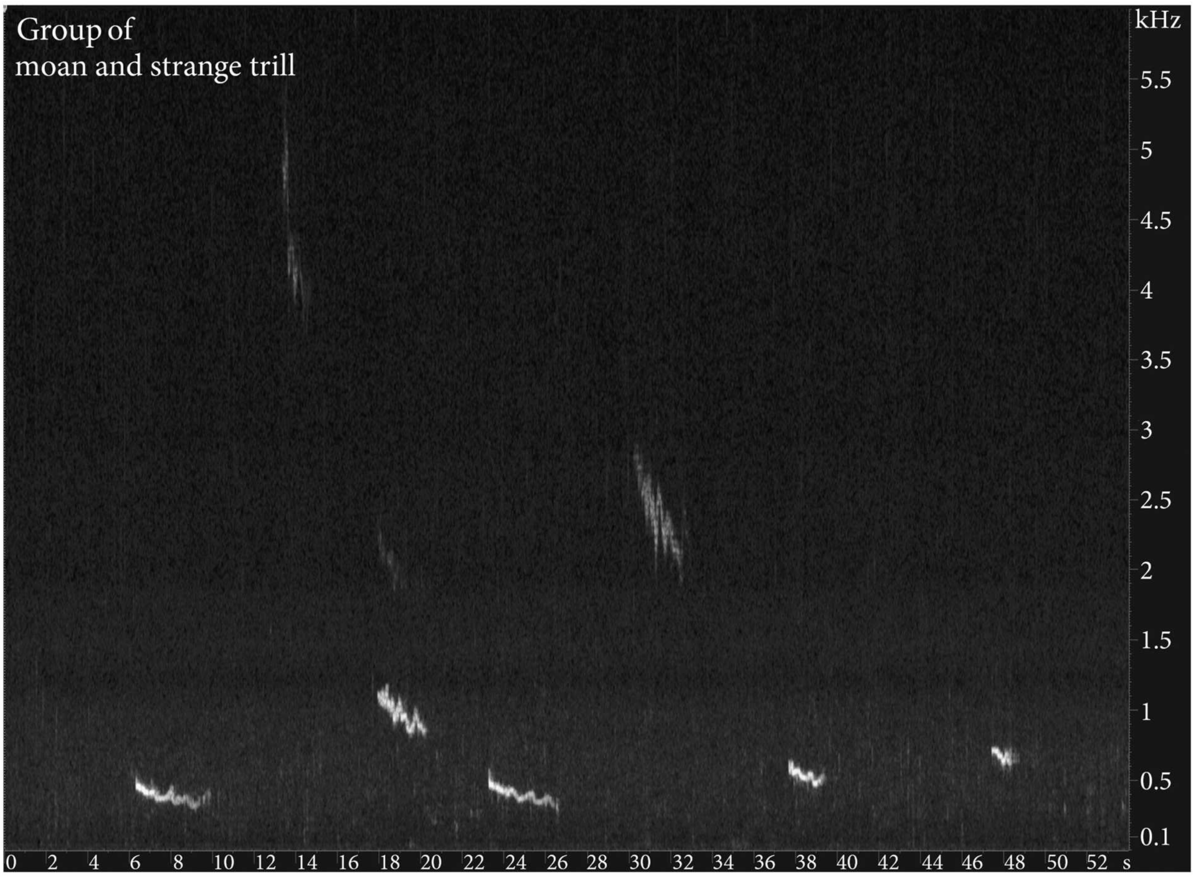


FIG. 2. (*Continued*).



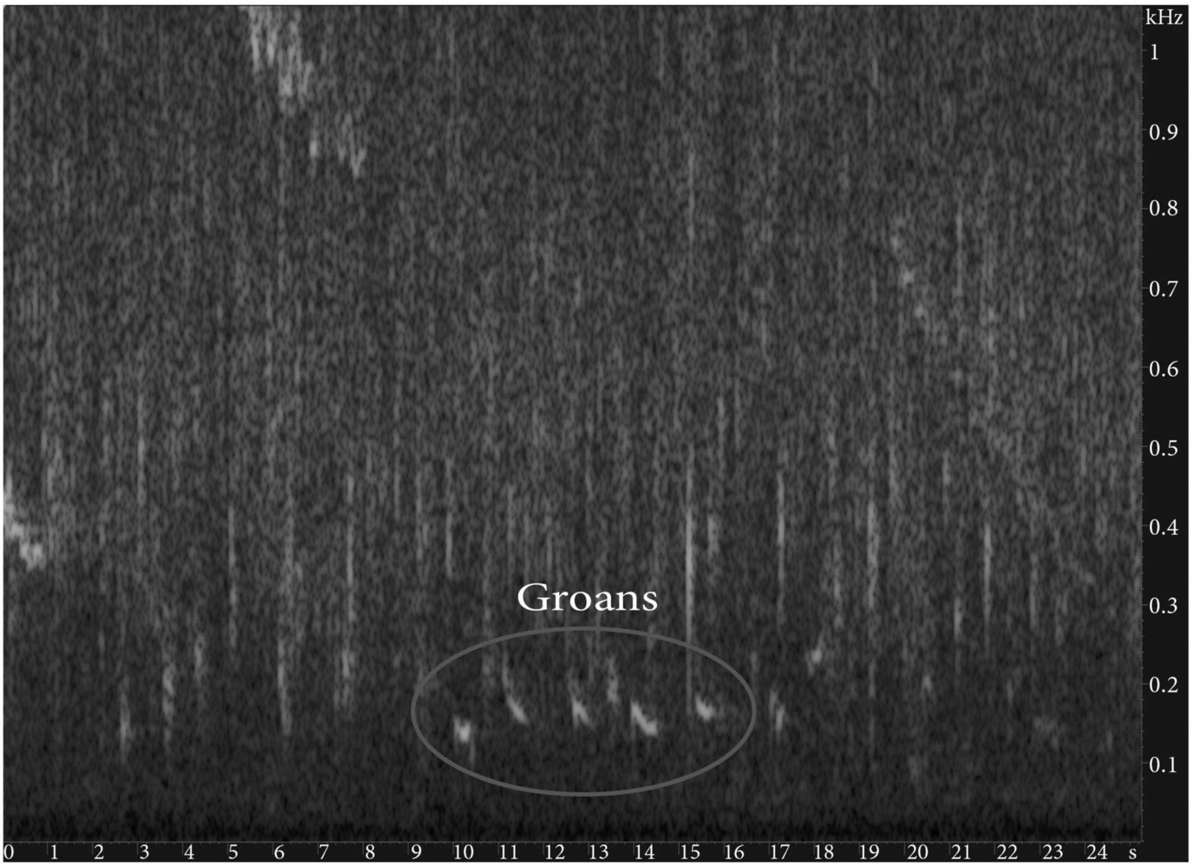


FIG. 2. (*Continued*).

acoustic parameters measured from the spectrogram in all vocalization classes [Figs. 2(a) and 2(b)]:

- Duration (D_t, s): total duration of the vocalization;
- Minimum frequency ($f_{\min}$, Hz): minimum frequency of the vocalization measured on the spectrogram within the total vocalization;
- Maximum frequency ($f_{\max}$, Hz): maximum frequency of the vocalization measured on the spectrogram within the total vocalization;
- Bandwidth (BW, Hz): range included between $f_{\max}$ and $f_{\min}$;
- Central frequency (f_c, Hz): frequency in the half of the bandwidth.

Moreover, to better characterize the different vocalizations, other acoustic parameters were measured in the spectrogram only in long trill, sweep trill, step trill, and hooked trill [Figs. 2(a) and 2(b)]

- The number of steps: number of frequency modulated downsweep preceded by a flat upsweep;
- The presence/absence of harmonics: defined as the presence or absence of harmonics over the fundamental frequency;
- The frequency revealed at the half of the D_t (f_{mid});
- The frequency modulation rate (FM_r): the number of cycles per second of FM;
- The frequency modulation range (FM_d): FM elongation.

Finally, to characterize the final part of the hooked trill class [hook, Fig. 2(a)], the following parameters were extracted from the spectrogram:

- Hook duration (D_{th}, s): duration included between $f_{\min}$ and f_{end} points;
- Hook maximum frequency (f_{maxh}, Hz): measured on the spectrogram within the hook;
- Hook ending frequency (f_{end}, Hz): measured on the spectrogram within the total vocalization;
- Bandwidth upsweep (BW_{up}, Hz): bandwidth included between f_{maxh} and $f_{\min}$;
- Bandwidth downsweep (BW_{ds}, Hz): bandwidth included between $f_{\max}$ and f_{end}.

An operator manually identified bearded seal calls to count vocal events for each class.

C. Statistical analysis

The statistical analyses were performed using package "randomForest" (Liaw and Wiener, 2002) implemented in R software (R Core Team, 2016).

The RF ensemble method (Breiman, 2001) was used in order to corroborate the vocal classes and discriminate vocalizations on the basis of the acoustic parameters of D_t, $f_{\min}, f_{\max}$, BW, f_c. Classification trees were constructed using a bootstrap aggregating algorithm (Breiman, 1996) that allows a reduced variance of predicted values and decreased risk of overfitting. Consequentially, each tree was built on a random sub-sampled training dataset while the subsequent predictions were carried out considering the remaining data (called out-of-bag, OOB) and allowing an unbiased estimate of the classification error. Prediction performances of the model are additionally improved, introducing a further source of diversity by a random restriction of the predictor variables used in each split (Breiman. 2001). Optimal model parameters were identified setting up a grid of tuning parameters in order to maximize correct predictions, using the OOB estimate of misclassification rates as a measure of model performance. As a consequence, a number of 1500 trees and 2 random variables at each split were considered. Variable importance was used in order to identify the predictor's contribution to the fitted model. Therefore, for each predictor, the mean decrease in accuracy (MDA) was defined as the normalized difference of the classification accuracy between two models: one considering the original predictor and one considering a randomly permuted predictor (Liaw and Wiener, 2002). Hence, variables importance was evaluated in dependence of their impact on the model predictions in terms of MDA. Partial dependence analysis (Friedman, 2001) was then performed with the aim to identify the most important predictor's values that characterize each vocal class. It performs an estimation of the partial relationship between the most important predictor identified (i.e., highest MDA) and the model outcome. Following this approach, for each variable of interest (X_j), it was fixed an equally spaced grid of values over the range of X_j and an averaged prediction function over all the combinations of observed values of the other predictors in the dataset were considered (Liaw and Wiener, 2002). The response function for each vocal class (*K*th) was estimated as given by

$$fk(X) = \log pk(X) - \Sigma j \log pj(X)/K$$

(Liaw and Wiener, 2002), where $pk(X)$ was the probability of membership in the *K*th class given the predictors.

Finally, higher values of the response function allowed delimiting optimal range of each predictor for which the prediction of the *K*th class is maximized and hence identifying the values of the acoustic parameters that mostly characterized the vocal classes.

III. RESULTS

A. Bearded seals vocal repertoire

In total, 17 220 bearded seal vocalizations were detected: 10 387 at the glacier site and 6833 at the open site. Specifically, in the glacier site the vocalizations were recorded among February to June, while in the open site only between May and June. Further, in the period between May and June, all the classes defined have been recorded simultaneously both in glacier and open site (see Fig. 3).

Nine classes of vocalizations, some of those have been defined previously by other authors (see Table I), were identified and described:

(1) Vocalization with marked FM
  (a) Trill with steps
    - long trill,
    - sweep trill,
    - step trill,
    - hooked trill;

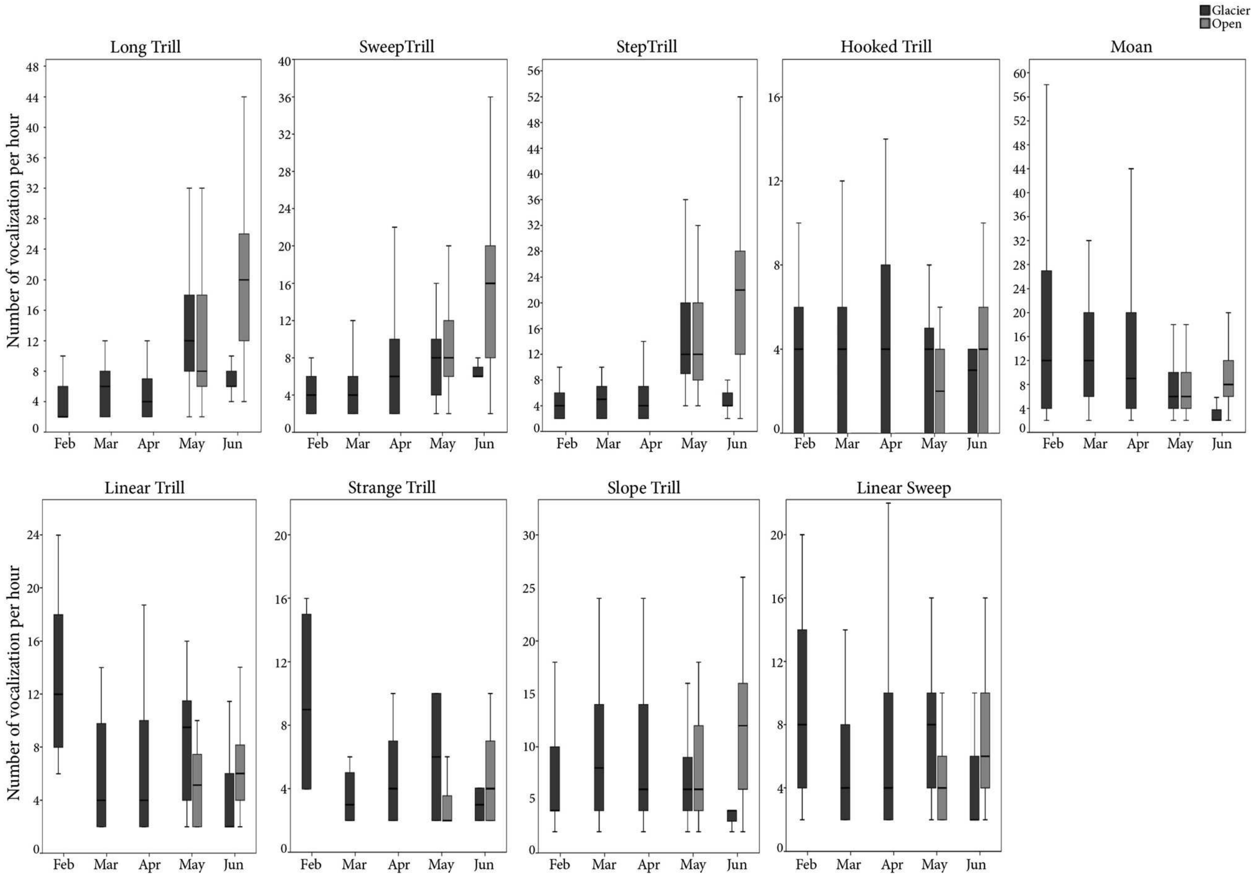


FIG. 3. (Color online) Number of vocalizations per hour in glacier site (blue plots) and open site (green plots) during the months of vocal activity (from February to June). (Median ± 25th–75th percentiles; whiskers ± 1st–99th percentiles.)

(b) Trill without steps
  - moan,
  - linear trill,
  - strange trill,
  - slope trill;

(2) Vocalization without FM

- linear sweep

The presence of steps, structural units composed of a FM downsweep preceded by a short upsweep flat, has been detected in different vocalizations classes [Fig. 2(a)]. These steps are characterized by an exponential downward trend with progressively increasing in duration and bandwidth. Another common feature of the steps is that at the descending phase modulated in frequency corresponds to an increase in the modulation depth, while a reduction of intensity [see the blue oscillogram in the Fig. 2(a)] in the upsweep phase and first portion of downsweep is observed [Fig. 2(a)].

The long trills [Fig. 2(a)] are the vocalizations with greater D_t and BW (Table I). This vocalization follows an exponential downward trend and is compounded by a series of steps.

The sweep trill [Fig. 2(a)] is characterized by the succession of two distinct phases [Fig. 2(a)]; the two parts might have the same duration (Van Parijs *et al.*, 2001; Risch *et al.*, 2007), although the first part is on average longer than the first one (their duration ratio can range from 0.75 to 1.79 with a mean [±standard deviation (SD)] of 1.29 ± 0.24, $N = 50$). The first part, following a descending logarithmic curve, is composed of steps, from a minimum of 8 up to 13 (Table I). The second part does not present a FM while tracing a sigmoid curve. Because the ratio between f_c and f_{mid} is unitary (Mean ± SD, 1.00 ± 0.06, $N = 50$), the f_c is reached to half of the total D_t.

The step trills [Fig. 2(a)], as the sweeps trills, are composed by two distinct parts. While the first part is composed of descending steps modulated in frequency, the second part proceeds without steps with a progressive reduction of the FM$_d$, a feature that was also found in linear trill [Figs. 2(a) and 2(b)]. The durations of the two parts are almost equal (their ratio is 1.01 ± 0.01, mean ± SD, $N = 46$), while the bandwidth of the first part occupies 2/3 of the total bandwidth (the ratio between the bandwidth of the first and second part is 1.99 ± 0.33, mean ± SD, $N = 46$). Moreover, the second part of the vocalization has a FM$_r$ lower than the first [Fig. 2(a)].

The hooked trills [Fig. 2(a)] acquired their name from the characteristic hook at the end of the vocalization. The head is characterized by rapid frequency decay and presents similarities with moan [Figs. 2(a) and 2(b)]. The central part (tail) is composed of 5–6 steps (Table I), although sometimes only 3–4 steps were observed. The last step of the tail has an inverse relationship between FM$_r$ and FM$_d$ [to a gradual decreasing of the FM$_r$ corresponds an increase of the FM$_d$;

TABLE I. Mean values (±standard deviation) of acoustic variables of the nine vocal classes of *E. barbatus* with percentage and number (no.) of occurrence refer to the total number of calls recorded across the two sites. *N.* is the number of samples used for the RF ensemble and for the measurements of the acoustic variables. In the last two columns similar classes previously described by others in and out Svalbard. D_t, total duration; D_r (min - max), minimum and maximum duration values of recorded events; $f_{\max}$, maximum frequency; $f_{\min}$, minimum frequency; f_c, central frequency; f_{mid}, frequency detected at $D_t/2$; BW, bandwidth; Harm, presence or absence of harmonics; Step, number of steps. SV, Svalbard vocalization. HCA, High Canadian Arctic vocalization. WCA, West Canadian Arctic vocalization.

		Acoustic variables										
Class	Percentage no.	D_t (s)	$D_{r(\mathrm{min\ -\ max})}$ (s)	$f_{\max}$ (Hz)	$f_{\min}$ (Hz)	f_c (Hz)	f_{mid} (Hz)	BW (Hz)	Harm (present/absent)	Step (no.)	Class previously described in Svalbard	Class previously described out of Svalbard
Long trill (*N*.50)	13.2% 2271	64.6 ± 10.1	47.6–84	5003.3 ± 1437.3	318.3 ± 76.6	2664.3 ± 731.7	—	4686.4 ± 1407.6	Present	10 ± 2.6	Van Parijs *et al.*, 2001; Trill Risch *et al.*, 2007; SV1 (T)	Risch *et al.*, 2007; HCA1 (T) Cleator *et al.*, 1989
Sweep trill (*N*.50)	9.1% 1561	9 ± 0.7	6.8–10.1	2659.6 ± 562.1	302 ± 64.4	1492.6 ± 298.8	1584 ± 323	2357.5 ± 559.4	Present	10.7 ± 1.1	Van Parijs *et al.*, 2001; Sweep Risch *et al.*, 2007; SV3 (S)	Risch *et al.*, 2007; SV3 (S)
Step trill (*N*.46)	14.9% 2570	28.6 ± 3.6	16.6–34.6	1637.6 ± 347.5	238.4 ± 57.8	935.6 ± 176.4	—	1354.8 ± 365	Present	10.8 ± 1.8		Cleator *et al.*, 1989
Hooked trill (*N*.56)	2.9% 505	34.1 ± 2	29.6–38.1	706.9 ± 36.5	106.3 ± 7.6	406.6 ± 19.8	—	600.5 ± 36.5	Present	5.4 ± 0.5		Cleator *et al.*, 1989
Moan (*N*.105)	37.3% 6418	3 ± 1.3	0.8–6.7	487.5 ± 213.2	322.7 ± 166.2	405.4 ± 187.9	—	165.3 ± 72.6	Present	0	Van Parijs *et al.*, 2001 Moan Risch *et al.*, 2007; SV4 (M)	Risch *et al.*, 2007; WCA5 (M), AL3 (M) Cleator *et al.*, 1989 Budelsky, 1992
Linear trill (*N*.33)	4.8% 828	10.3 ± 0.6	9.4–11.9	1048.1 ± 110.2	252.1 ± 30.7	650.1 ± 62	568.3 ± 56.3	796 ± 103.8	Present	0	Risch *et al.*, 2007; SV2 (T)	Risch *et al.*, 2007; WCA3 (T) Cleator *et al.*, 1989
Strange trill (*N*.79)	1.3% 229	7.9 ± 5.2	0.8–22.9	4598.9 ± 1614.9	2092 ± 674.9	3354.4 ± 1051	—	2506.8 ± 1307	Present	0		
Slope trill (*N*.34)	10.3% 1776	4.5 ± 0.7	3.1 – 6	808.2 ± 68.9	277 ± 34.6	542.6 ± 45	488.2 ± 42.1	531 ± 61.6	Present	0	Risch *et al.*, 2007; SV2 (T)	Risch *et al.*, 2007; WCA3 (T) Cleator *et al.*, 1989
Linear sweep (*N*.32)	6.2% 1062	3.6 ± 0.7	2.5–6	985.1 ± 62.5	294.6 ± 52.1	639.9 ± 47.4	527.7 ± 63.6	664.8 ± 68	Present	0		

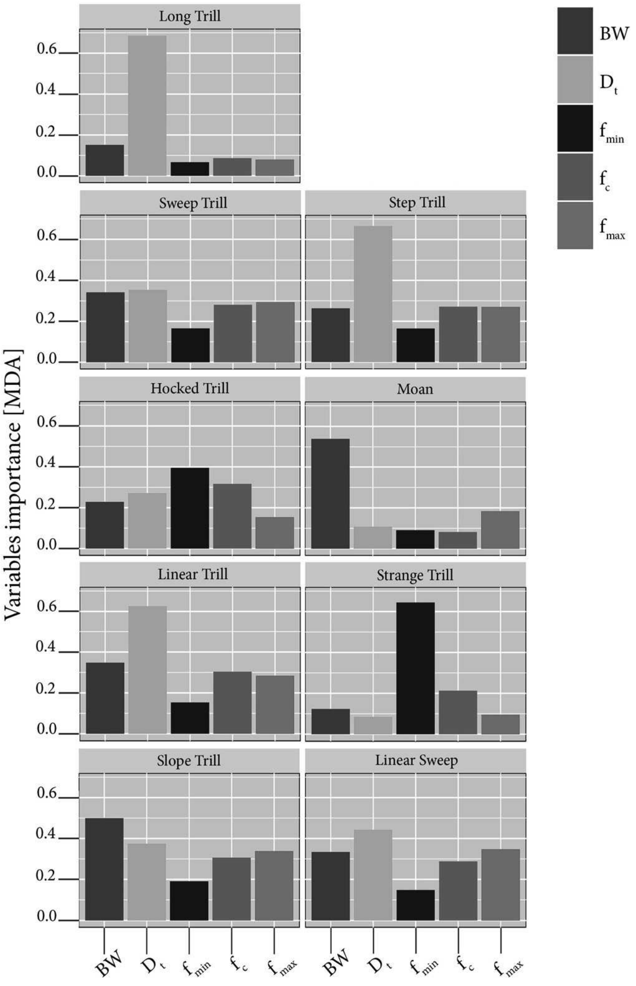


FIG. 4. (Color online) Importance, expressed in terms of MDA, of the acoustic variable considered in the RF model: BW, bandwidth; D_t, total duration; $f_{\min}$, minimum frequency; f_c, central frequency, and $f_{\max}$, maximum frequency. The parameters in the legend follow a descending importance in the overall discrimination among vocal classes, while histogram indicates the variable importance in the discrimination of singularly considered classes.

Fig. 2(a)]. Observing the amplitudes profile (see blue oscillogram), the head, characterized by a rapid attack (the time in which the amplitude varies between 0 and its maximum value; Roads *et al.*, 1996) and subsequent decay, represents the part with greater intensity [Fig. 2(a)]. Contrary to what happens in other vocalizations with steps, there are no intensity reductions in correspondence with the upsweep and first part of downsweep. The ending hook is comprised of a flat upward sweep and a flat descendant sweep of shorter duration. The bandwidth values of upsweep (BW_{us}) and downsweep (BW_{ds}) are, respectively, 208.2 ± 10.86 Hz and 108.9 ± 11.5 Hz (mean ± SD, $N = 56$). The $f_{\min}$ and the f_{end} are, respectively, 103.6 ± 7.6 Hz and 205.6 ± 8.8 Hz (mean ± SD, $N = 56$); the f_{maxh} is 314.6 ± 10.8 Hz (mean ± SD, $N = 56$) and hook D_{th} is 3.0 ± 0.3 s [mean ± SD, $N = 56$; see hook in Fig. 2(a)].

The moans [Fig. 2(b)] are the most frequent vocalization (Table I). They are characterized by a low frequency, short duration, a decreasing trend in frequency, and a low FM. The head is the portion with the greatest intensity, and can be found without tail. The moan shows a similar pattern found in the head of the hooked trill. This vocalization sometimes is repeated more times in a minute together with strange trills [Fig. 2(c)].

The linear trills [Fig. 2(b)] are composed of an initial portion (head), generally lasting less than 2 s, with a rapid non-linear decay in frequency, and a tail with a linear and lower decrement of frequency. The tail is otherwise characterized by the progressive reduction of the FM_d [with a maximum value at the beginning of the tail and reaching zero at the end of the vocalization; Fig. 2(b)]. In the tail, the FM_d decreases with the frequency, while the FM_r remains constant during the entire duration of vocalization.

The strange trills [Fig. 2(b)] are frequency-modulated vocalizations that, contrary to what was found in other classes, do not show a regular internal articulation, proceeding with a discontinuous and irregular trend both in the spectral and amplitude profile. This vocalization showed the highest values of $f_{\min}$ (Table I). The strange trill can be observed within moan events [Fig. 2(c)].

The slope trill [Fig. 2(b)] follows a similar trend to the linear trill but they differ significantly in D_t, BW, and $f_{\max}$ (Table I).

The linear sweeps [Fig. 2(b)] are vocalizations without any detectable FM, and their overall trend draws a logarithmic downward curve. This vocalization occurs usually as a single tone even if in the initial part some harmonics can be present. Otherwise, this vocalization shows remarkable similarities with the second part of sweep trill [Figs. 2(a) and 2(b)].

A very high discrimination of the nine different classes emerged from the results of RF classification, which showed a low OOB estimate of misclassification rate and correctly assigned the 98.58% of the observations. Specifically, the model did not exhibit errors for 6/9 of classes, whereas it showed low error values (≤9%) in the other cases (Table II), confirming the clear separation of the vocal classes. According to the MDA, the most important variables for the overall classification of the vocal classes were BW and D_t, followed by $f_{\min}$, f_c, and $f_{\max}$ in this order. Specifically, variables importance estimated for the correct classification of each vocal type allowed to identify the acoustic parameters that mostly characterized each group, in agreement with the call structure. Results are summarized in Fig. 4.

Moreover for each class, in support of our description, partial dependence analysis identified the ranges of the most important variables (i.e., MDA > 0.2) for which the prediction was maximized, allowing the recognition of the values of the acoustic parameters that better characterized the identified vocal groups (Table III).

In only four cases vocalizations do not fall into the categories described in this study; these are constituted by sequences of short vocalizations (4–7 events) at low frequency (starting from 80 Hz) with an evident correspondence with groans described by Cleator *et al.* [1989; Fig. 2(c)]. However, the small number of these events has not allowed a detailed characterization.

Finally, in 1.8% of the total vocalizations, a low frequency signal, simultaneous to other signals belonging to other classes, was recorded [see the deep trill in Fig. 2(b)].

TABLE II. Confusion matrix as revealed by RF classification on *E. barbatus* vocal classes. Observations are compared with the model predictions and the differences are expressed in terms of classification error (Class error).

	Long trill	Sweep trill	Step trill	Hooked trill	Moan	Linear trill	Strange trill	Slope trill	Linear sweep	Class error
Long trill	50	0	0	0	0	0	0	0	0	0
Sweep trill	0	50	0	0	0	0	0	0	0	0
Step trill	0	0	46	0	0	0	0	0	0	0
Hooked trill	0	0	0	56	0	0	0	0	0	0
Moan	0	0	0	0	104	0	1	0	0	0.009
Linear trill	0	0	0	0	0	33	0	0	0	0
Strange trill	0	0	0	0	1	0	79	0	0	0.01
Slope trill	0	0	0	0	1	2	0	31	0	0.09
Linear sweep	0	0	0	0	0	0	0	3	29	0.09

This signal could be the low frequency part of a biphonical signal.

IV. DISCUSSION AND CONCLUSION

E. barbatus is known to produce a complex vocal repertoire in Arctic regions like Alaska (Cleator *et al.*, 1989; Jones *et al.*, 2014; Risch *et al.*, 2007), Canadian high Arctic (Cleator *et al.*, 1989; Risch *et al.*, 2007), and western Canadian Arctic (Risch *et al.*, 2007). Previous studies in Svalbard Islands recognized only four vocalization types (Van Parijs *et al.*, 2001), showing that bearded seals in this area would have limited underwater vocal behaviour if compared with other populations. This study, carried out in the same area 15 years later, has allowed both the improvement of the distinction of the previous four categories defined by Van Parijs *et al.* (2001) and the recognition of other different vocal classes. These results, with the classification of nine call types, are in agreement with the vocal complexity described for other Arctic bearded seal populations (Cleator *et al.*, 1989; Jones *et al.*, 2014; Risch *et al.*, 2007). This key result could be reached due to year-round, high-resolution passive acoustic monitoring. Indeed, the emission rate of the different classes changes among months (see Fig. 3, de Vincenzi *et al.*, 2017), and consequently the possibility to record a determinate class is strictly depending on the recording period. The Van Parijs *et al.* (2001) study was restricted both in time and space, and these restrictions probably caused a missing of some kind of vocalizations in their recordings.

While the classes identified in this study can be included into the macro-distinction defined by Van Parijs *et al.* (2001), we did not find any presence of vocalizations type "flat tone." Considering the trend in frequency, no ascending vocalizations were found in our study. This result is in agreement with previous studies in Svalbard (Van Parijs *et al.*, 2001; Risch *et al.*, 2007), confirming that this characteristic is crucial for the discrimination of the Svalbard populations among the others in the Arctic. In the study of Charrier *et al.* (2013), the authors, using playback and synthetic trills from two distinct populations, showed the ability by bearded seal males to perceive geographic variation in their vocal emissions. In our study, the observation of signal's inner articulation permits to split vocalizations in two principal groups: vocalizations with marked FM (trill) and vocalizations without a marked FM (flat). Within the first category, there are vocalizations without and with steps. Further, among the latter, two vocalizations, step trill and sweep trill, presented two distinct parts with the second one lacking of steps. It was found that basic units compose some vocalizations classes, i.e., flat up- and downsweep and trilled downsweep. These basic units can be found in isolation or composed in greater lexical units, e.g., the second part of step trill and sweep trill would seem to be formed, respectively, by linear trill and linear sweep. Future works, basing on the interdisciplinary study of animal communicative system (termed zoosemiotic; Sebeok, 1968, 1990), could be direct to explore the

TABLE III. Range of values of the most important variables ($\text{MDA} > 0.2$) for which the RF model prediction of the *E. barbatus* vocal classes is maximized, as emerged by the partial dependence analysis. For each class, a quantitative description of the acoustic features that better provide the discrimination among classes is shown.

	Variables				
Class	D_t (s)	$f_{\min}$ (Hz)	$f_{\max}$ (Hz)	f_c (Hz)	BW (Hz)
Long trill	>42.2	—	—	—	—
Sweep trill	7.4–9.1	—	1672.9–4249.4	787.5–1157.3	1899.6–2994.3
Step trill	15.7–40.7	—	1350.8–1511.8	787.5–1157.3	961.2–1743.2
Hooked trill	<42.2	<197.7	—	<540	<840.8
Moan	—	—	—	—	179.3–335.7
Linear trill	10.7–14.1	—	1028.7–1189.8	<787.5	<961.2
Strange trill	—	719.7–4374	—	2513.2 – 6211.2	—
Slope Trill	<7.4	—	<1028.7	<664.2	< 648.5
Linear Sweep	<7.4	—	<1028.7	< 787.5	< 804.8

perception of these different units by the seals, and to understand if there is a semiotic organization inside the acoustic signals (Charrier *et al.*, 2013). The basic units that form vocal compositions are the basis for the micro- and macrogeographic variations in vocalizations, forming dialects, and can have significant impact on the ability of conspecific to communicate, and thus on their breeding behaviour (Searcy *et al.*, 2002).

At the current time, there is a debate regarding the existence of two subspecies of bearded seals, *E. b. barbatus* and *E. b. nauticus*. Several studies have used vocalizations to complement morphological and molecular data to infer phylogenetic relationships within species (Martens, 1996; Risch *et al.*, 2007; Stanger, 1995). Comparing bearded seal vocal repertoires throughout four Arctic study sites (Alaska, High Canadian Arctic, West Canadian Arctic, and Svalbard), Risch *et al.* (2007) attests that the Svalbard population has the most distinct repertoires. This conclusion is based on the results of the precedent study of Van Parijs *et al.* (2001) attesting only four call types, while all other repertoires consisted of about ten call types. Thanks to the year-round high-resolution passive acoustic monitoring implemented in different sites of the studied area, in this study some vocalizations has been described for the first time in the bearded seal Svalbard population (Table I), and some of these present similarities with vocalizations recorded in others Arctic regions (Budelsky, 1992; Cleator *et al.*, 1989; Jones *et al.*, 2014; Risch *et al.*, 2007). Specifically, the step trill and the hooked trill show remarkable similarities with the trill category of High Canadian Arctic, while linear trill and slope trill show remarkable similarities with the trill category of West Canadian Arctic (Cleator *et al.*, 1989; Risch *et al.*, 2007). In the light of this new cataloging for Svalbard population vocalizations, a future acoustic comparative study could find out different population distances with regard to the Risch *et al.* (2007) results. This comparative study could be coupled with a genetic study that in the Arctic region is lacking because of high costs and/or difficult logistics.

The visual inspection of the spectrograms allowed the identification of a particular vocalization defined deep trill [see Fig. 2(b)]. This is very low frequency sound modulated in frequency and follows an overall slightly downward trend. This low frequency signal is always emitted simultaneously with other vocal classes, resulting in probable biphonic events. In all the cases seen, the deep trill started in conjunction with one other vocal class. This event has been recorded here for the first time in bearded seal, even if is a quite common feature for other marine mammals (Papale *et al.*, 2014). Even if it is not possible conferring the deep trill emission to a single subject or to the entire community, the absence of overlapping of more than one deep trill and the absence of simultaneous registration in the two recording sites let us suppose that it can be ascribable to an anatomical-structural characteristic related to a single specimen.

In conclusion, with this study we identified nine call classes using multivariate statistical analyses of call properties, and the detailed acoustic characteristics of each distinctive call type were described. This improved our current knowledge on the underwater vocalization of bearded seals in Svalbard. Description of the acoustic repertoire is crucial to understanding the social interactions between individuals since vocal communication is at the base of complex social structure. Detailed study of the vocal complexity represents a first step to improve the understanding of the behaviour and the social function of these calls (Serrano and Terhune, 2002). Considering that different biotic factors such as sex, age, anatomical-structural variation, and behaviour may influence the acoustic structure of underwater vocalization of bearded seals, future studies tied to the observation of species behaviour and the relationship between the animals and the environment will help interpret the function of these different seal calls. Furthermore, the time extension of the recording season allowed recording calls emitted in different time periods. These results could not have been achieved without using a long-term acoustic recording that permitted the analysis of acoustic behaviour over a longer period than before. Further, long time period studies imply the possibility to investigate through acoustics any relationships with ecological factors, and improve the conservation of the species in the framework of the environmental changes ongoing in the Arctic. Therefore, this opens the possibility to extend the study to other ice breeding seals.

ACKNOWLEDGMENTS

This research has been carried out under the Arctic Field Grant project KUAM (Grant No. 235878/E10) and the CALVINGSEIS project (Grant No. 244196/E10) funded by The Norwegian Research Council. The authors thank Elvio Ceci for his valuable support and assistance in understanding the linguistic and semiotic traits developed in this text; Dr. Angelo Viola, Emiliano Liberatori, Roberto Sparapani, and the Department of Earth Systems Science and Environmental Technologies of the Italian National Research Council for the logistic support in Ny-Ålesund.

Azzolin, M., Gannier, A., Lammers, M. O., Oswald, J. N., Papale, E., Buscaino, G., Buffa, G., Salvatore, M., and Cristina, G. (**2014**). "Combining whistle acoustic parameters to discriminate Mediterranean odontocetes during passive acoustic monitoring," J. Acoust. Soc. Am. **135**, 502–512.

Breiman, L. (**1996**). "Bagging predictors," Mach. Learn. **24**, 123–140.

Breiman, L. (**2001**). "Random forests," Mach. Learn. **45**, 5–32.

Brudzynski, S. (**2010**). "Handbook of mammalian vocalization: An integrative neuroscience approach," in *Handbook of Behavioral Neuroscience*, edited by S. Brudzynski (Academic, London), pp. 1–526.

Budelsky, R. A. (**1992**). "Underwater behavior and vocalizations of the bearded seal (*Erignathus barbatus*) off Point Barrow, Alaska," Dissertation, University of Minnesota, Minneapolis, MN, 99 pp.

Charrier, I., Mathevon, N., and Aubin, T. (**2013**). "Bearded seal males perceive geographic variation in their trills," Behav. Ecol. Sociobiol. **67**, 1679–1689.

Cleator, H. J. (**1996**). "The status of the bearded seal, *Erignathus barbatus*, in Canada," Can. Field-Nat. Ott. ON **110**, 501–510.

Cleator, H. J., and Stirling, I. (**1990**). "Winter distribution of bearded seals (*Erignathus barbatus*) in the Penny Strait area, Northwest Territories, as determined by underwater vocalizations," Can. J. Fish. Aquat. Sci. **47**, 1071–1076.

Cleator, H. J., Stirling, I., and Smith, T. G. (**1989**). "Underwater vocalizations of the bearded seal (*Erignathus barbatus*)," Can. J. Zool. **67**, 1900–1910.

Cottier, F., Tverberg, V., Inall, M., Svendsen, H., Nilsen, F., and Griffiths, C. (**2005**). "Water mass modification in an Arctic fjord through cross-shelf exchange: The seasonal hydrography of Kongsfjorden, Svalbard,"

J. Geophys. Res. Oceans **110**, 2156–2202, https://doi.org/10.1029/2004JC002757.

Derocher, A. E. (**2005**). "Population ecology of polar bears at Svalbard, Norway," Popul. Ecol. **47**, 267–275.

Derocher, A. E., Wiig, Ø., and Andersen, M. (**2002**). "Diet composition of polar bears in Svalbard and the western Barents Sea," Polar Biol. **25**, 448–452.

de Vincenzi, G., Parisi, E., Torri, M., Papale, E., Mazzola, S, Nuth, C., and Buscaino, G. (**2017**). "Influence of environmental parameters on the use and spatio-temporal distribution of Arctic seal *Erignathus barbatus* vocalizations in Kongsfjorden, Svalbard Islands," Polar Biol. (submitted).

Friedman, J. H. (**2001**). "Greedy function approximation: A gradient boosting machine," Ann. Stat. **29**(5), 1189–1232.

Jones, J. M., Thayre, B. J., Roth, E. H., Mahoney, M., Sia, I., Merculief, K., Jackson, C., Zeller, C., Clare, M., Bacon, A., Weaver, S., Gentes, Z., Small, R. J., Stirling, I., Wiggins, S. M., and Hildebrand, J. A. (**2014**). "Ringed, bearded, and ribbon seal vocalizations north of Barrow, Alaska: Seasonal presence and relationship with sea ice," Arctic, **67**, 203–222.

Liaw, A., and Wiener, M. (**2002**). "Classification and regression by randomForest," R News **2**, 18–22.

Lydersen, C., Assmy, P., Falk-Petersen, S., Kohler, J., Kovacs, K. M., Reigstad, M., Steen, H., Strøma, H., Sundfjord, A., Varpe, O., Walczowski, W., Weslawski, J. M., and Zajaczkowski, M. (**2014**). "The importance of tidewater glaciers for marine mammals and seabirds in Svalbard, Norway," J. Mar. Syst., **129**, 452–471.

MacIntyre, K. Q., Stafford, K. M., Berchok, C. L., and Boveng, P. L. (**2013**). "Year-round acoustic detection of bearded seals (*Erignathus barbatus*) in the Beaufort Sea relative to changing environmental conditions, 2008–2010," Polar Biol. **36**, 1161–1173.

MacIntyre, K. Q., Stafford, K. M., Conn, P. B., Laidre, K. L., and Boveng, P. L. (**2015**). "The relationship between sea ice concentration and the spatio-temporal distribution of vocalizing bearded seals (*Erignathus barbatus*) in the Bering, Chukchi, and Beaufort Seas from 2008 to 2011," Prog. Oceanogr. **136**, 241–249.

Martens, J. (**1996**). "Vocalizations and speciation of Palearctic birds," in *Ecology and Evolution of Acoustic Communication in Birds*, edited by D. E. Kroodsma and E. H. Miller (Cornell University Press, Ithaca, NY), pp. 221–240.

McDonald, M. A., Hildebrand, J. A., and Wiggins, S. M. (**2006**). "Increases in deep ocean ambient noise in the Northeast Pacific west of San Nicolas Island, California," J. Acoust. Soc. Am. **120**, 711–718.

Moore, S. E., and Huntington, H. P. (**2008**). "Arctic marine mammals and climate change: Impacts and resilience," Ecol. Appl. **18**, S157–S165.

Nelson, R. R., Burns, J. J., and Frost, K. J. (**1984**). "The bearded seal (*Erignathus barbatus*)," in *Marine Mammal Species Accounts, Wildlife Technical Bulletin No. 7*, edited by J. J. Burns (Alaska Department of Fish and Game, Juneau, AK), pp. 1–6.

Noad, M. J., Cato, D. H., Bryden, M. M., Jenner, M.-N., and Jenner, K. C. S. (**2000**). "Cultural revolution in whale songs," Nature **408**, 537.

Papale, E., Azzolin, M., Cascão, I., Gannier, A., Lammers, M. O., Martin, V. M., Oswald, J., Perez-Gil, M., Prieto, R., Silva, M. A., and Giacoma, C. (**2014**). "Acoustic divergence between bottlenose dolphin whistles from the Central–Eastern North Atlantic and Mediterranean Sea," Acta Ethologica, **17**, 155–165.

Papale, E., Perez-Gil, M., Castrillon, J., Perez-Gil, E., Ruiz, L., Servidio, A., Tejedor, M., Giacoma, C., and Martin, V. (**2017**). "Context specificity of Atlantic spotted dolphin acoustic signals in the Canary Islands," Ethol. Ecol. Evol. **29**(4), 311–329.

R Core Team (**2016**). "R: A language and environment for statistical computing R Foundation for Statistical Computing," Vienna, Austria 2015, available at www.R-Proj.Org (Last viewed June 17, 2017).

Risch, D., Clark, C. W., Corkeron, P. J., Elepfandt, A., Kovacs, K. M., Lydersen, C., Stirling, I., and Van Parijs, S. M. (**2007**). "Vocalizations of male bearded seals, *Erignathus barbatus*: Classification and geographical variation," Anim. Behav. **73**, 747–762.

Roads, C., Strawn, J., Abbott, C., Gordon, J., and Greenspun, P. (**1996**). *The Computer Music Tutorial* (Mass. Inst. Technol., Cambridge, MA), pp. 1–1256.

Sciacca, V., Caruso, F., Beranzoli, L., Chierici, F., De Domenico, E., Embriaco, D., Favali, P., Giovanetti, G., Larosa, G., Marinaro, G., Papale, E., Pavan, G., Pellegrino, C., Pulvirenti, S., Simeone, F., Viola, S., and Riccobene, G. (**2015**). "Annual acoustic presence of Fin Whale (*Balaenoptera physalus*) offshore Eastern Sicily, central Mediterranean Sea," PloS One, **10**, e0141838.

Searcy, W. A., Nowicki, S., Hughes, M., and Peters, S. (**2002**). "Geographic song discrimination in relation to dispersal distances in song sparrows," Am. Nat. **159**, 221–230.

Sebeok, T. A. (**1968**). "Zoosemiotics," Am. Speech **43**, 142–144.

Sebeok, T. A. (**1990**). *Essays in Zoosemiotics*, edited by Toronto Semiotic Circle (Victoria College in the University of Toronto, Toronto), pp. 1–147.

Serrano, A., and Terhune, J. M. (**2002**). "Antimasking aspects of harp seal (*Pagophilus groenlandicus*) underwater vocalizations," J. Acoust. Soc. Am. **112**, 3083–3090.

Simon, M., Nuuttila, H., Reyes-Zamudio, M. M., Ugarte, F., Verfub, U., and Evans, P. G. (**2010**). "Passive acoustic monitoring of bottlenose dolphin and harbour porpoise, in Cardigan Bay, Wales, with implications for habitat use and partitioning," J. Mar. Biol. Assoc. U. K. **90**, 1539–1545.

Sousa-Lima, R. S., Norris, T. F., Oswald, J. N., and Fernandes, D. P. (**2013**). "A review and inventory of fixed autonomous recorders for passive acoustic monitoring of marine mammals," Aquat. Mamm. **39**, 23–53.

Stanger, K. F. (**1995**). "Vocalizations of some cheirogaleid prosimians evaluated in a phylogenetic context," in *Creatures of the Dark*, edited by L. Alterman, G. A. Doyle, and M. K. Izard (Springer, Boston, MA), pp. 353–376.

Stirling, I., Calvert, W., and Cleator, H. (**1983**). "Underwater vocalizations as a tool for studying the distribution and relative abundance of wintering pinnipeds in the High Arctic," Arctic **36**(3), 262–274.

Svendsen, H., Beszczynska-Møller, A., Hagen, J. O., Lefauconnier, B., Tverberg, V., Gerland, S., Ørbøk, J. B., Bischof, K., Papucci, C., Zajaczkowski, M., Azzolini, R., Bruland, O., Wiencke, C., Winther, J. G., and Dallmann, W. (**2002**). "The physical environment of Kongsfjorden–Krossfjorden, an Arctic fjord system in Svalbard," Polar Res. **21**, 133–166.

Van Opzeeland, I. (**2010**). "Acoustic ecology of marine mammals in polar oceans," Rep. Polar Mar. Res. **619**, 1–332, available at http:/hdl.handle.net/10013/epic.36260.

Van Parijs, S. M., and Clark, C. W. (**2006**). "Long-term mating tactics in an aquatic-mating pinniped, the bearded seal, *Erignathus barbatus*," Anim. Behav. **72**, 1269–1277.

Van Parijs, S. M., Kovacs, K. M., and Lydersen, C. (**2001**). "Spatial and temporal distribution of vocalizing male bearded seals: Implications for male mating strategies," Behaviour **138**, 905–922.

Van Parijs, S. M., Lydersen, C., and Kovacs, K. M. (**2003**). "Vocalizations and movements suggest alternative mating tactics in male bearded seals," Anim. Behav. **65**, 273–283.