Objective classification of North Atlantic right whale (*Eubalaena glacialis*) vocalizations to improve passive acoustic detection

Ildar R. Urazghildiiev

The Lab of Ornithology, Cornell University, Ithaca, New York, 14850

Susan E. Parks

Department of Biology, Syracuse University, Syracuse New York, 13244

Correspondence: Ildar R. Urazghildiiev, The Lab of Ornithology, Cornell University, Ithaca, New York, 14850. Phone: +1 (607) 257 4773. E-mail: ildar.urazghildiiev@gmail.com

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.322v1 | CC-BY 4.0 Open Access | received: 29 Mar 2014, published: 29 Mar 2014

ABSTRACT

Passive acoustic monitoring is playing an increasing role in the detection of endangered North Atlantic right whales (NARW). Previous acoustic monitoring has relied on a single stereotyped vocalization, the upcall. Here the entire repertoire produced by NARW during the winter and early spring in Cape Cod Bay, Massachusetts is described. An objective sound classification scheme and automatic classification algorithm were developed. Nine days of acoustic recordings were used for the data analysis and a total of 9,611 right whale sounds were identified. The objective classification scheme of right whale sounds allowed for rapid identification of a diversity of right whale sounds. These sounds were assigned to 6 classes of narrowband upcalls, downsweep, complex and high frequency calls, wideband gunshot sounds and complex sounds. Results indicate that the prevalence of upcalls varied from 28% of detected calls in January to 80% in April. Other classes of signals were also well represented in the repertoire including the narrowband complex (10-36%) and high frequency calls (1-26%), wideband gunshot sounds (4-25%) and wideband complex sounds (0 – 25%). The prevalence of non-upcall signals suggests that including more signals classes may improve rates of detection for right whales in the Cape Cod Bay habitat.

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.322v1 | CC-BY 4.0 Open Access | received: 29 Mar 2014, published: 29 Mar 2014

INTRODUCTION

Research into the vocal behavior of the critically endangered North Atlantic right whale *(Eubalaena glacialis)* (NARW) has been conducted over the past decade (Matthews *et al*., 2001; Mellinger *et al*., 2007; Parks and Tyack, 2005; Parks *et al*., 2011; Urazghildiiev *et al*., 2009; Vanderlaan *et al*., 2003). These studies are important to provide both baseline behavioral data and to inform conservation measures for this species. Long-term passive acoustic monitoring of NARW habitats has been used to successfully determine seasonal presence of individuals to contribute to near-real-time implementation of conservation measures (e.g., Van Parijs *et al.*, 2009; Clark *et al.*, 2010).

One of three known areas which receive federal designation as a critical habitat area of NARW (Federal Register 59 FR 28793) is the Cape Cod Bay ecosystem. This area is used primarily for feeding, socializing, and as nursery area for cows and their calves (Clark *et al.*, 2010). To acoustically detect the presence of NARW in this area, the upcalls, or contact calls, are typically used because this type of calls is the most stereotyped right whale vocalization that is known to be produced by both males and females (Parks and Tyack, 2005; Parks *et al.*, 2011). However, different call types classified as “tonals” and “gunshot sounds” have also been detected in Cape Cod Bay in March and April 2004 (Van Parijs *et al.*, 2009). Other call types that have been described from digital acoustic tag (DTAG) based recordings from North Atlantic right whales include “moans” (Matthews *et al.*, 2001), as well as “screams,” (corresponding to “high,” “hybrid,” and “pulsive” calls produced by Southern right whales (Clark, 1983)), “blows,” “warbles,” and “downcalls” (Parks *et al.*, 2011). The relative proportion of different right whale vocalization types produced varies seasonal and between habitats (Van Parijs *et al*. 2009). Therefore, ignoring call types other than upcalls may result in missing essential information about the presence or behavior of NARW in a particular area. Therefore, a quantitative evaluation of the usage of the vocal repertoire of NARW and more objective methods for classification of their other call types in different habitats and times of year are necessary to improve passive acoustic detection for this species.

This study focused on characterizing the multiple call types recorded from NARW during the late winter and spring in Cape Cod Bay using data collected by a sparse array of bottom-mounted synchronized hydrophones. Due to the high variability in the parameters of detected NARW vocalizations, their classification is not a trivial problem. The choice of a practical classification scheme depends on the goals of classification as well as on the presence of well separated signal parameters. In this work, a classification scheme based on measurable signal parameters was developed to allow for the design of an automatic classifier. The goal of this work was to design a practical sound classification scheme, to develop an automatic classifier, and to obtain empirical distributions of signal classes to better understand right whale vocal behavior and to improve passive acoustic monitoring applications.

METHOD

Acoustic recordings

Data were collected using a passive acoustic monitoring system consisting of 11 synchronized marine autonomous recording units (MARUs) located at known positions. The MARUs were approximately 8 km apart (Fig. 1). Each MARU consisted of an HTI-94-SSQ hydrophone (High Tech, Inc., Long Beach, Mississippi, USA) with a sensitivity of -168 dB re 1 V/µPa, an amplifier with a gain of 23.5 dB and A/D converter with a sensitivity of 103 Bit/V (Parks *et al.*, 2009). The units had a flat (± 1.0 dB) frequency response between 10 – 2500 Hz.

All MARUs recorded continuously at a sampling rate of 5 kHz. Acoustic recordings were made from February 8 to May 3, 2012.

Signal detection

Signal detection was performed in three stages. Upcalls are considered to be the most stereotyped distinctive NARW vocalization; therefore an upcall detector (Urazghildiiev *et al.*, 2009) was applied in the first stage to identify periods of right whale acoustic activity. The detector thresholds were selected in a way to detect >95% of upcalls clearly visible on the spectrograms in three or more channels. All signals detected in this stage were checked by the human operator and false detections occurring due to noise transients were removed from the testing data set. To detect other types of NARW vocalizations, an energy detector was applied in the second stage. The NARW vocalizations were identified from all automatically detected sounds by visual inspection of the signal spectrograms. Signals with SNR > 10 dB were used in this study. Signals were identified as NARW calls if they fit any of the call types described in the literature, such as moans, screams, downcalls, high calls, gunshot sounds, and other calls observable within the frequency range below 2.5 kHz (see, e.g., Matthews *et al.*, 2001; Parks *et al.*, 2011; Parks and Tyack, 2005; Van Parijs *et al.*, 2009; Clark, 1982; Clark *et al.*, 2010).

To avoid misclassification of NARW calls with similar calls produced by humpback whales, data collected after April 8, 2012, when the first song produced by humpback whales was detected in the North part of Cape Cod Bay by recorders 1 and 8, have not been used for the analysis. All selected calls were also checked to make sure that other humpback whale vocalizations, such as patterned repetitive vocalizations or humpback whale stereotyped calls described in Stimpert *et al.*, (2011) were not included in the testing data. In the third stage, visual inspection of the spectrogram by a human operator was performed and all NARW vocalizations not detected in in the previous stages were added to the testing data set.

Signal classification

Impulsive signals produced by NARW were classified based on types of distribution of signal energy in time and frequency domain. One of the basic features for classification is the local bandwidth of signals. Using this parameter, two main classes of NARW vocalizations were introduced. The first class comprised locally narrowband frequency-modulated (FM) signals whose energy is distributed in a vicinity of an instantaneous frequency and its harmonics. The FM signals can be represented as

$$s(t) = \sum_i A_i(t) cos\big(2\pi i f(t)\big), \qquad (1)$$

where function $f(t)$ specifies frequency modulation of signals, and $A_i(t)$ is an amplitude modulation of the *i*th harmonic.

The second main class combined wideband signals having wider distribution of energy in frequency domain, such as gunshot sounds, exhalations, and slaps (Matthews *et al.*, 2001; Parks and Tyack, 2005; Clark, 1982).

The next feature used for classification of NARW vocalizations was the frequency band occupied by signals. The major part of the first harmonics of locally narrowband NARW vocalizations is distributed within two frequency bands situated lower and higher than 200 Hz. Correspondingly, NARW tonals were classified as low frequency (upcalls, downcalls, moans, < 200 Hz) and high frequency (high calls, > 200 Hz) locally narrowband signals.

Classification of locally narrowband vocalizations was also made based on behavior of the FM function $f(t)$. This function can either be monotonic or have a finite number of inflection points, i.e., the points $t_k, k = 1 \dots K$, where $df(t)/dt = 0$. The more complicated a locally narrowband vocalization is, the higher number K of inflection points its FM function has. By using this parameter, locally narrowband vocalizations were classified as simple ($K \leq 1$) and complex ($K \geq 2$) signals. Finally, simple FM signals were classified as upsweep, constant sweep, and downsweep signals depending on the number of points t_k where $df(t)/dt$ is larger, equal or lower than zero, respectively.

The energy distribution of wideband NARW vocalizations in the time-frequency plane is less structured. The most distinctive and important type of wideband calls is the gunshot sound, which typically has shorter duration than other types of wideband vocalizations. Therefore two subclasses of wideband sounds were considered: gunshot sounds (and other slaps), and complex wideband vocalizations.

RESULTS

Classification scheme

Using the features of NARW vocalizations described in section II.B, we introduced the following classes of signals:

Class NU: locally narrowband low frequency upsweep signals. The FM function $f(t)$ of NU signals has one or less local minimum or maximum and increases in time such that the number of points where $df(t)/dt \geq 0$ is larger than 50% of total points associated with signal. The peak frequency of NU signals $F_p < 200$ Hz and instantaneous bandwidth of the first harmonic $B < 100$ Hz.

Class ND: locally narrowband low frequency constant sweep and downsweep signals. Signals associated with this class were similar to NU calls except for the instantaneous frequency of ND signals, $f(t)$, is a non-ascending function, i.e., the number of points where $df(t)/dt \leq 0$ is larger than 50% of total points associated with signal. Constant sweep and downsweep signals were combined into one class since the proportion of this type of vocalizations is relatively small as compared to other types of calls.

Class NC: locally narrowband low frequency complex signals. Signals of this class were similar to NU calls except for the number of local minimums or maximums is greater than 1.

Class NH: locally narrowband high frequency signals. This class comprised all locally narrowband calls having peak frequency $F_p \geq 200$ Hz and any FM function $f(t)$.

Class WG: wideband gunshot sounds. The instantaneous bandwidth of WG signals $B \geq 100$ Hz and duration $D < 1.5$ s. Separate frequency components of WG signals were typically indistinguishable on the spectrogram.

Class WC: wideband complex signals. The instantaneous bandwidth of WC signals $B \geq 100$ Hz and duration $D \geq 1.5$ s. Frequency components of some WC signals were distinguishable on the spectrogram, but they were typically separated by less than 50 Hz.

Example spectrograms of signals from each class are shown in Figs. 2 – 7. The spectrograms were obtained using 1024 point FFT (frequency resolution is 4.88 Hz) with Hann window and 75% overlap. Note that the wideband signal displayed in Fig. 6, right, does not look like typical gunshot (see Fig. 6, left and middle). However, this signal was attributed to WG class due to its parameters' fit to this class.

Note that classification of locally narrowband signals by the complexity of the FM function is applicable to both low frequency and high frequency sounds. However, this kind of classification was applied to low frequency sounds only since upcalls traditionally used for passive acoustic detection of NARW belong to this class of locally narrowband signals. The need for more detailed classification of low frequency sounds is explained by the fact that the detection performance of the automatic upcall detectors (Gillespie 2004, Urazghildiiev *et al.*, 2009) degrades as the complexity of upcalls increases, so considering simple and complex low frequency tonal calls as separate classes can be useful for improving the efficiency of the automatic NARW detectors. For the sake of simplicity, all high frequency locally narrowband calls were associated with the same class.

Automatic classifier

Measurable signal parameters used for the design of the classification scheme described above allowed developing an automatic classifier. The block-diagram of the automatic classifier is represented in Fig. 7. The classifier was implemented as a custom-build Matlab program. It was used to aid the human operator to classify NARW vocalizations. The human operator checked classes computed by the automatic algorithm for all detected signals and changed classes of signals incorrectly classified automatically. Testing the performance of the automatic classifier was outside the scope of this work.

Empirical distribution of classes

A total of 9 chunks of acoustic recordings, 24 hours each, collected on February 9, 16, 23 and 27; March 4, 12, 17, 23; and April 1 and 7, 2012, were used for the data analysis. These days were selected because the results of visual survey indicate the presence of NARW in the vicinity of the array and the days were temporally separated to allow for sampling of different individual whales.

The results of classification of all detected NARW vocalizations are represented in Table I. Daily distribution of detected NARW calls and their classes are displayed in Fig. 9. Fig. 10 displays empirical distribution of classes of all detected NARW vocalizations.

DISCUSSION

The vocalization repertoire of NARW during their spring migration in Cape Cod Bay is highly variable. Using the distribution of signal energy in the time and frequency domains of detected right whale signals has led to a new scheme for the classification of NARW vocalizations. An important advantage of the proposed classification scheme is that it is based on a finite set of measurable signal parameters: instantaneous bandwidth, peak frequency, duration, and the number of inflection points in the FM function. Using these parameters, an automatic classifier was designed and implemented in this study

Statistical analysis and classification of the diversity of signals produced by right whales in this habitat yielded six robust, quantifiable classes to provide additional signals to aid in passive acoustic monitoring of this species. These classes included upcalls (NU), three additional tonal categories (downsweep (ND), narrowband complex (NC), high (NH)), and two broadband categories (gunshot/slaps (WG), and wideband complex (WC)). All previously reported call types (see, e.g., Matthews *et al.*, 2001; Parks *et al.*, 2011; Parks and Tyack, 2005; Van Parijs *et al.*, 2009; Clark 1982; Clark *et al.*, 2010) can be classified using the proposed classification scheme and an automatic classifier. Using distribution of signal energy in time and frequency domain, an automatic classifier was designed and implemented in this study.

The sound production activity levels of NARW changed significantly over the course of the observation period. Sound production rates were highest in February and decreased in March. The relative usage of the components of the vocalization repertoire also changed. The relatively simple upcalls (NU class) typically used for passive acoustic monitoring of NARW comprise 48.7% of all detected NARW vocalizations, and daily occurrence of upcall changes from 28 % to 80.2%. This finding alone is intriguing, as it indicates that while upcalls were always detected (by default based on our methods for detection of other right whale signals), their prevalence within the repertoire of signals produced on a given day varied widely. This variation is likely due to a combination of different individual whales and variability in the behavioral context of signal production.

Signal types that are not typically considered in passive acoustic monitoring or detection or right whales comprised a significant portion of the detected signals over the study period. Signals with more complex frequency modulation (NC) and high calls (NH) comprise 9.9 – 35.8% and 0.5 – 25.8% of total detected calls, respectively. The percentage of high calls (NH class) was greater in February and March. The production rate of wideband gunshot sounds (WG) varied from 4.1 to 24.6%. These values are comparable to those reported in Van Parijs *et al.*, 2009, which found that in the Cape Cod Bay habitat, upcalls were the predominant call type, with fewer more complex tonal signals and low rates of gunshot sound production during spring months.

These types of analysis may also provide insight into the behavioral activity of right whales in different habitat areas or in different times of years (Van Parijs *et al.*, 2009). Previous studies have indicated different behaviors associated with the production of these different sound types (Parks *et al.*, 2005; Parks and Tyack 2005; Parks *et al.*, 2011). Analysis of seasonal variation in the production of a single call-type, the upcall, in the Gulf of Maine showed both seasonal and diel trends in sound production (Mussoline *et al.*, 2012). Future studies could expand this type of analysis to include additional call types produced by right whales.

This study indicates that applying NARW upcall detectors (e.g., Dugan *et al.*, 2010; Gillespie, 2004; Urazghildiiev and Clark, 2006; Urazghildiiev *et al.*, 2009) may result in missing a large number of non-upcall vocalizations and potentially, in the worst case scenario, result in missing the detection of vocalizing NARW or groups of whales if their vocalization session does not contain upcalls. Therefore, we suggest that the design and testing of more complicated detection techniques sensitive to non-upcall vocalizations should be an area of future research and will potentially allow for more detailed behavioral information to be obtained from these passive acoustic monitoring systems.

This study should be considered preliminary, as it is based on the acoustic behavior of right whales from 9 days in a single year. Further studies are necessary to assess the added benefit of adding detection of these additional signals for right whale passive acoustic monitoring. Further studies should also involve processing more data, evaluation the individual statistical properties of vocalization sessions produced by each vocalizing NARW or groups of closely spaced NARW and the design of the tools for automatic detection and classification of all NARW vocalizations.

ACKNOWLEDGMENTS

The authors wish to thank Dr. Christopher W. Clark for help with lab support and useful comments. Thanks also to Christopher Tessaglia-Hymes, Christopher Tremblay and Jason John Michalec for deployment and recovery of the pop-up recording units and Janelle Morano for useful discussions.

REFERENCES

Clark, C. W. (**1982**). "The acoustic repertoire of the southern right whale, a quantitative analysis," Anim. Behav. **30**, 1060-1071.

Clark, C.W., Brown, M.W., and Corkeron, P. (**2010**). “Visual and acoustic surveys for North Atlantic right whales, *Eubalaena glacialis,* in Cape Cod Bay, Massachusetts, 2001-2005: Management implications,” Mar. Mamm. Sci., **26**, 837 – 854.

Dugan, P.J., Rice, A.N., Urazghildiiev, I.R., and Clark, C.W. (**2010**). "North Atlantic Right Whale acoustic signal processing: Part I. Comparison of machine learning recognition algorithms," IEEE Applications and Technology Conference (LISAT), Long Island Systems, 1-6.

Gillespie, D. (**2004**). “Detection and classification of right whale calls using an “edge” detector operating on a smoothed spectrogram,” Canadian Acoustics, **32**, 39–47.

Matthews, J.N. , Brown, S., Gillespie, D., Johnson, M., McLanaghan, R., Moscrop, A., Nowacek, D., Leaper, R., Lewis, T., and Tyack, P. (**2001**). “Vocalization rates of the North Atlantic right whale (*Eubalaena glacialis*)*,”* J. Cet. Res. Man. **3**, 271-282.

Mellinger, D. K., Nieukirk, S. L., Matsumoto, H., Heimlich, S. L., Dziak, R. P., Haxel, J. (**2007**). “Seasonal occurrence of North Atlantic right whale (*Eubalaena glacialis*) vocalizations at two sites on the Scotian shelf,” Marine Mammal Science, **23**(**4**)**,** 856–867.

Mussoline, S.E, Risch, D, Hatch, L.T., Weinrich, M.T. Wiley, D. N., Thompson, M. A., Corkeron, P. J., Van Parijs, S. M. (**2012**)” Seasonal and diel variation in North Atlantic right whale up-calls: implications for management and conservation in the northwestern Atlantic Ocean,” Endang Species Res., **17**, 17-26.

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.322v1 | CC-BY 4.0 Open Access | received: 29 Mar 2014, published: 29 Mar 2014

Parks, S.E., and Tyack, P. L. (**2005**). "Sound production by North Atlantic right whales (*Eubalaena glacialis*) in surface active groups," J. Acoust. Soc. Am. **117**, 3297-3306.

Parks, S.E., Hamilton, P.K., Kraus, S.D., Tyack, P.L. (**2005**) "The Gunshot sound produced by male North Atlantic right whales (*Eubalaena glacialis*) and its potential function in reproductive advertisement," Mar. Mamm. Sci., **21**(**3**)**,** 458-475.

Parks, S.E., Urazghildiiev, I. R., and Clark, C. W. (**2009**). "Variability in ambient noise levels and call parameters of North Atlantic right whales in three habitat areas," J. Acoust. Soc. Am., **125**, 1230-1239.

Parks, S.E., Searby, A., Celerier, A., Johnson, M.P., Nowacek, D.P. and Tyack, P.L. (**2011**). "Sound production behavior of individual North Atlantic right whales: implications for passive acoustic monitoring," Endang. Species Res., **15**, 63 – 76.

Stimpert, A.K, Au, W. W. L., Parks, S. E., Hurst, T., Wiley, D. N. (**2011**). "Common humpback whale (Megaptera novaeangliae) sound types for passive acoustic monitoring," J. Acoust. Soc. Am. **129**, 476–482.

Urazghildiiev, I.R. and Clark, C.W. (**2006**). "Acoustic detection of North Atlantic right whale contact calls using the generalized likelihood ratio test," J. Acoust. Soc. Am. **120**, 1956 – 1963.

Urazghildiiev, I.R., Clark, C. W., Krein, T., and Parks, S. (**2009**). "Detection and recognition of North Atlantic right whale contact calls in the presence of ambient noise," IEEE Trans. on Oceanic Eng., **34**, 358 – 368.

Vanderlaan, A. S. M., Hay, A. E., and Taggart, C. T. (**2003**). "Characterization of North Atlantic right-whale (*Eubalaena glacialis*) sounds in the Bay of Fundy," IEEE Journal of Oceanic Engineering, **28**(**2**), 164–173.

Van Parijs, S.M., Clark, C. W., Sousa-Lima, R. S., Parks, S. E., Rankin, S., Risch, D., and Van Opzeeland, I. C. (**2009**). “Management and research applications of real-time and archival passive acoustic sensors over varying temporal and spatial scales,” Mar. Ecol. Prog. Ser., **395**, 2—36.

TABLE I: The number of detected signals from each class.

Date	Class						Nsig
	Locally Narrowband				Wideband		
	Low frequency			High frequency (NH)			
	Simple		Complex (NC)		Gunshot (WG)	Complex (WC)	
	Upsweep (NU)	Down sweep (ND)					
9-Feb	532 (28.1%)	10 (0.5%)	254 (13.4%)	490 (25.8%)	146 (7.7%)	464 (24.5%)	1896
16-Feb	535 (49.7%)	12 (1.1%)	252 (23.4%)	196 (18.2%)	55 (5.1%)	26 (2.4%)	1076
23-Feb	1002 (60.3%)	36 (2.2%)	317 (19.1%)	176 (10.6%)	99 (6.0%)	33 (2.0%)	1663
4-Mar	1091 (54.5%)	30 (1.5%)	350 (17.5%)	271 (13.6%)	151 (7.5%)	107 (5.3%)	2000
12-Mar	507 (77.5%)	3 (0.5%)	75 (11.5%)	40 (6.1%)	27 (4.1%)	2 (0.3%)	654
17-Mar	474 (32.4%)	22 (1.5%)	266 (18.2%)	299 (20.5%)	360 (24.6%)	40 (2.7%)	1461
23-Mar	286 (61.2%)	14 (3.0%)	92 (19.7%)	16 (3.4%)	57 (12.2%)	2 (0.4%)	467
1-Apr	111 (52.4%)	1 (0.5%)	76 (35.8%)	4 (1.9%)	18 (8.5%)	2 (0.9%)	212
7-Apr	146 (80.2%)	3 (1.6%)	18 (9.9%)	1 (0.5%)	14 (7.7%)	0 (0.0%)	182
Total	4684 (48.7%)	131 (1.4%)	1700 (17.7%)	1493 (15.5%)	927 (9.6%)	676 (7.0%)	9611

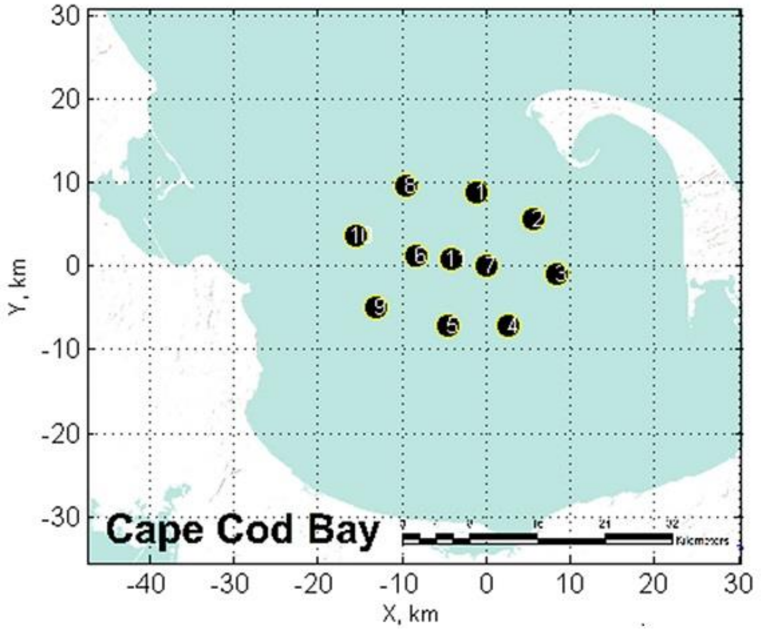


FIG. 1. Map of the sensor array geometry. MARU locations are shown as black, numbered circles.

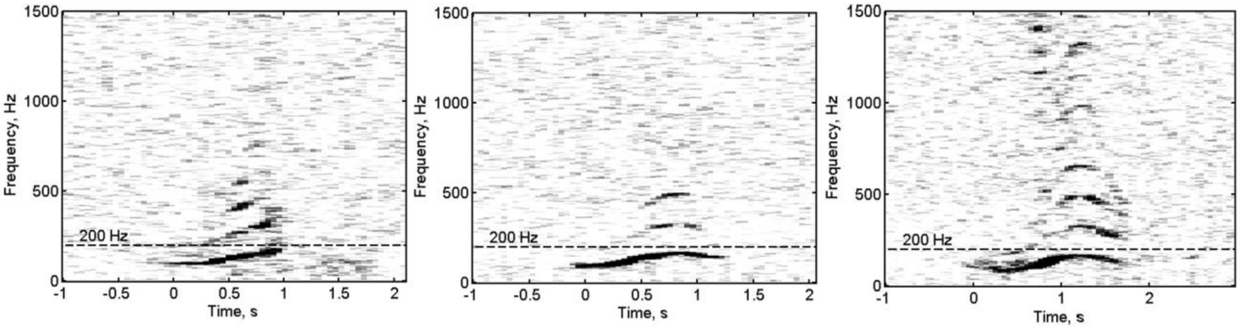


FIG. 2. Spectrograms of narrowband upsweep FM signals (NU). The spectrograms were obtained using 1024 point FFT with Hann window, 75% overlap, and sampling rate of 5 kHz.

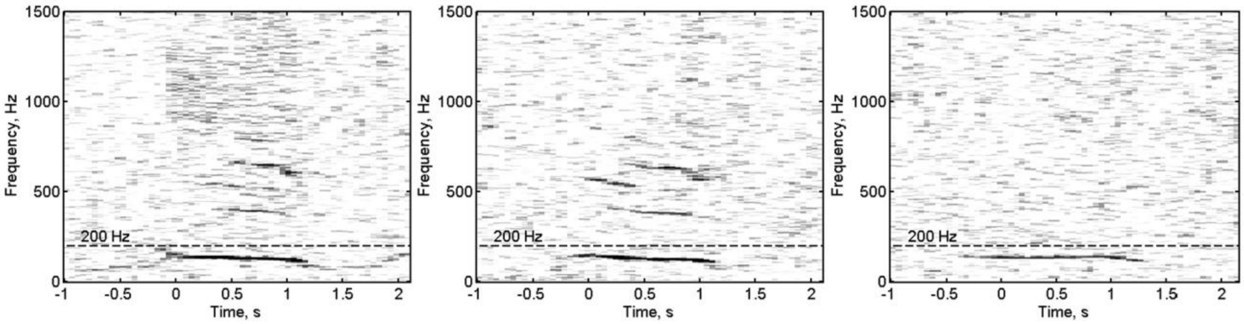


FIG. 3. Spectrograms of narrowband downsweep FM signals (ND).

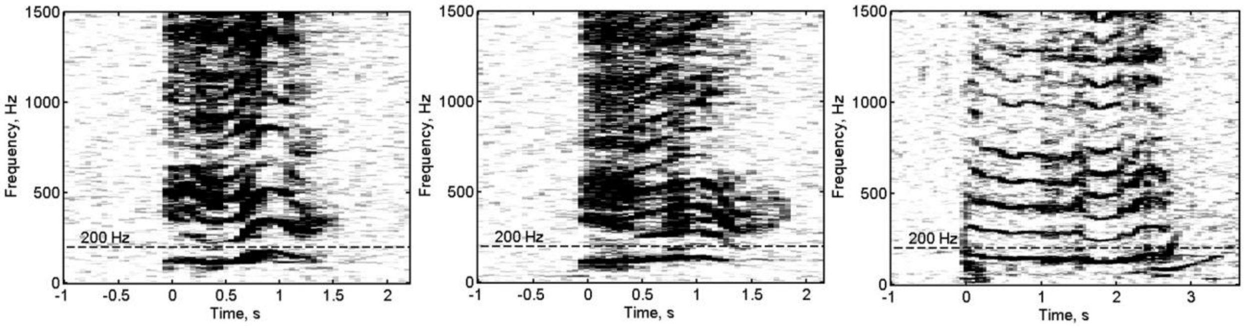


FIG. 4. Spectrograms of narrowband complex FM signals (NC).

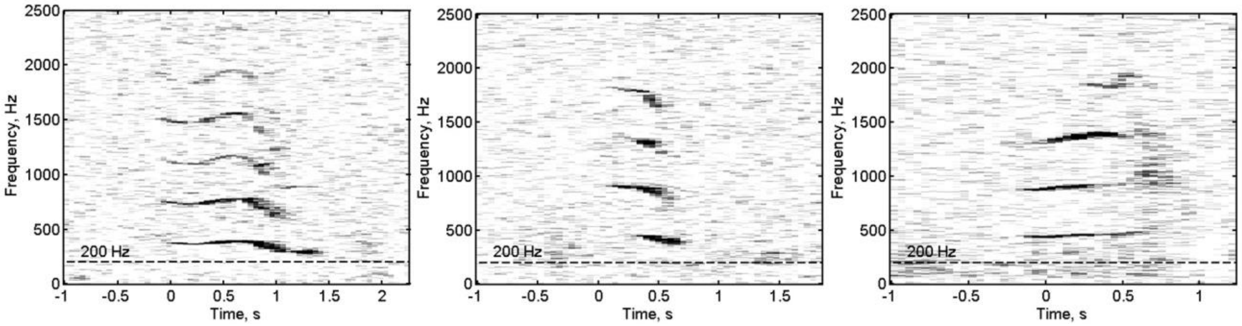


FIG. 5. Spectrograms of narrowband high calls (NH).

PeerJ PrePrints | http://dx.doi.org/10.7287/peerj.preprints.322v1 | CC-BY 4.0 Open Access | received: 29 Mar 2014, published: 29 Mar 2014

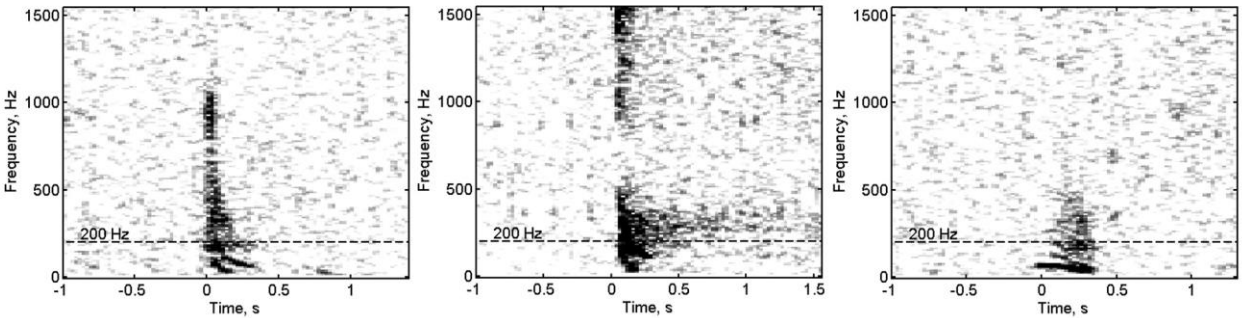


FIG. 6. Spectrograms of wideband gunshot sounds (WG).

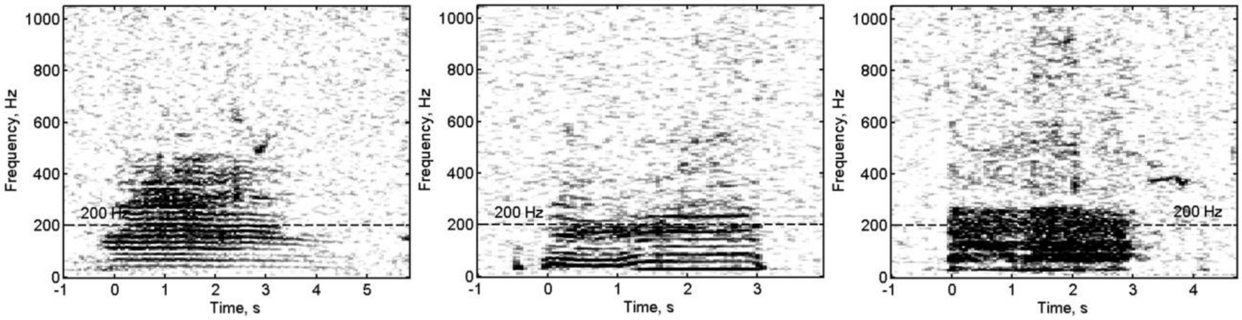
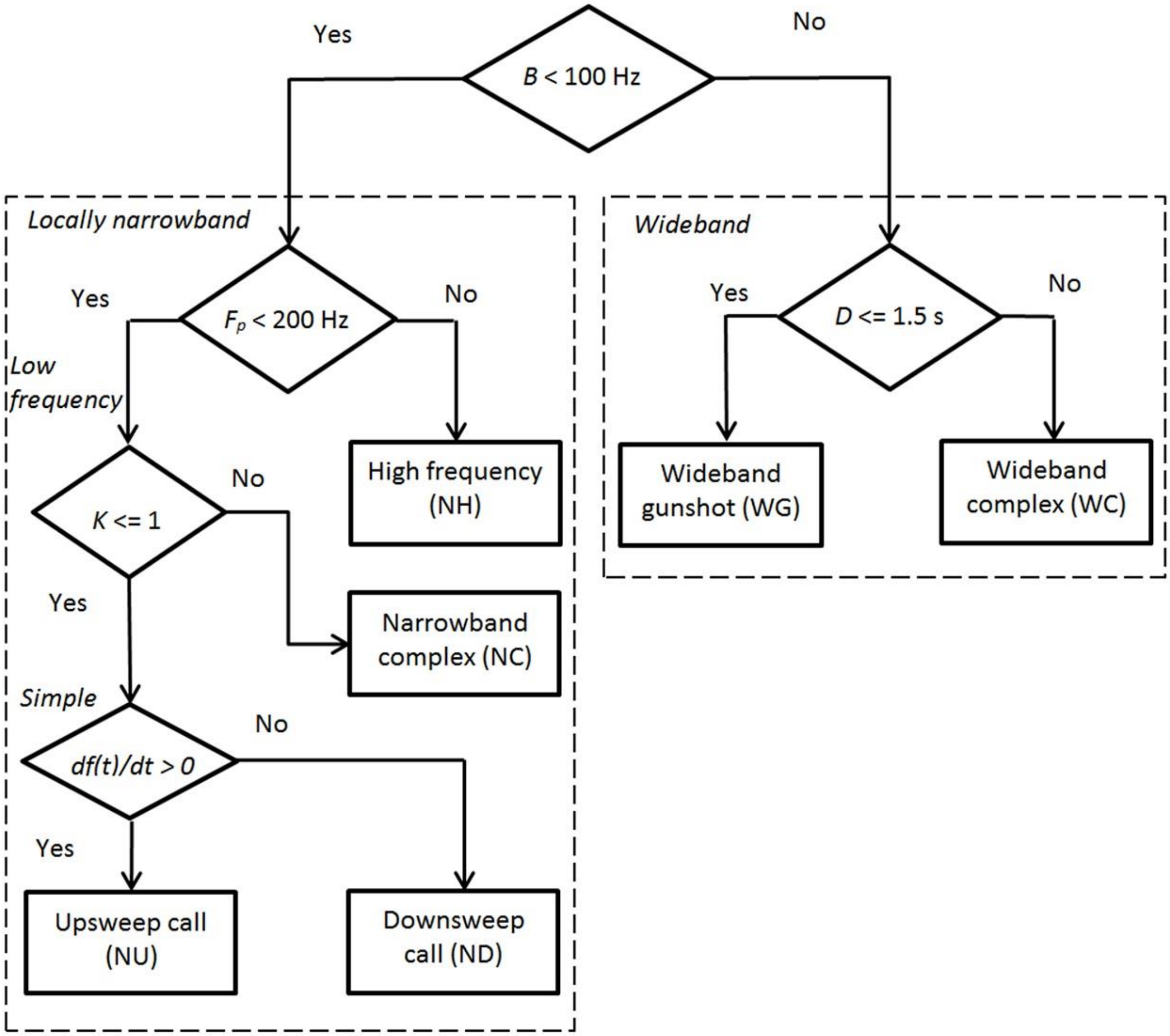


FIG. 7. Spectrograms of complex wideband signals (WC).



352

353 FIG. 8. Block-diagram of the automatic classifier.

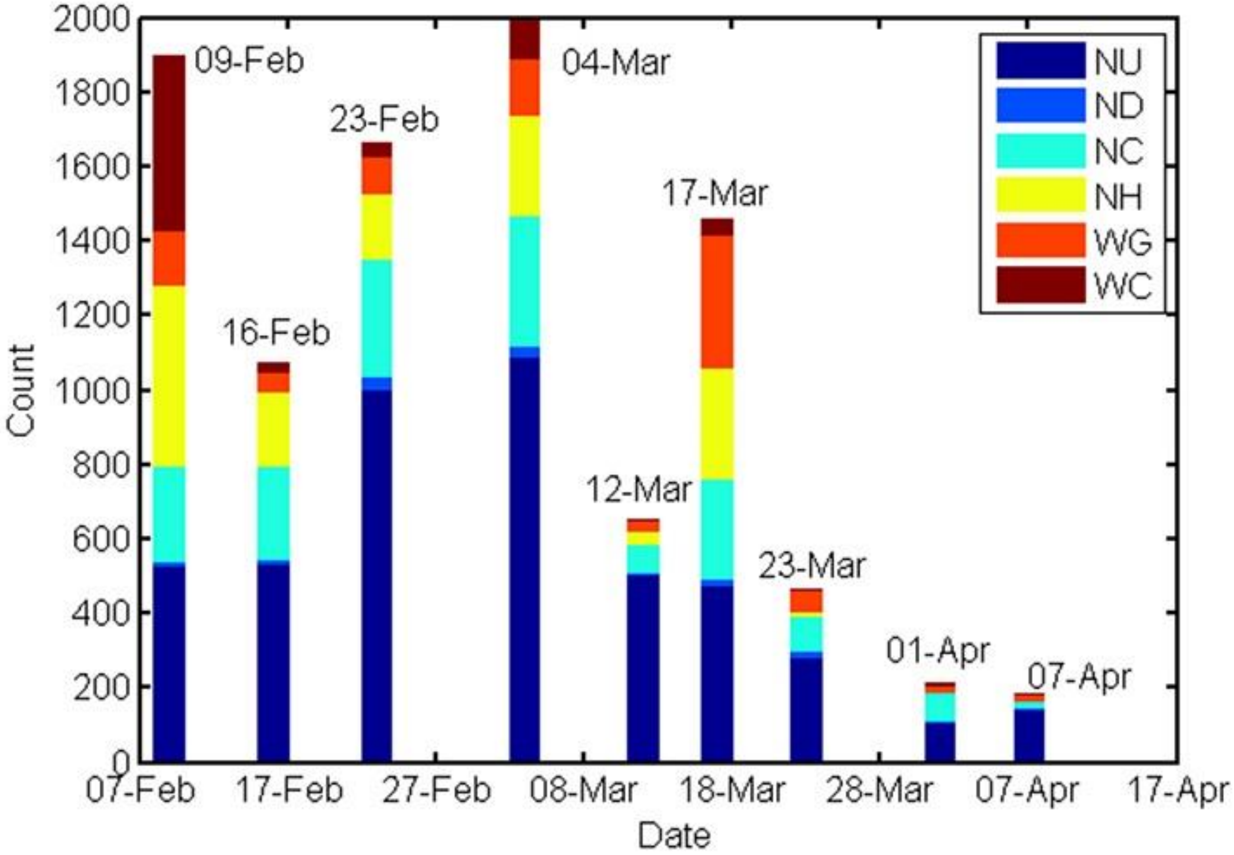


FIG. 9. Daily distribution of detected NARW calls and their classes.

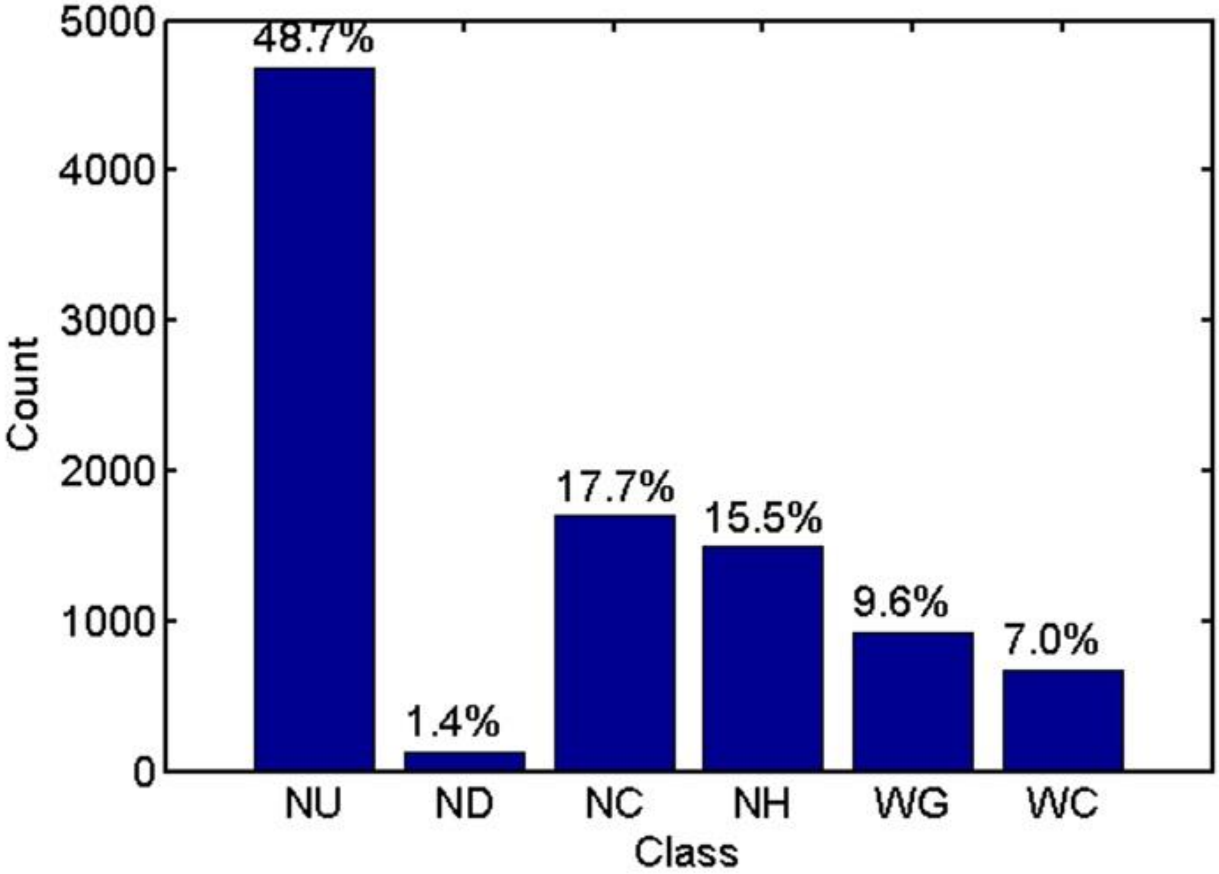


FIG. 10. Empirical distribution of classes of all detected NARW vocalizations.