Estimating minke whale relative abundance in the North Atlantic using passive acoustic sensors

Shahideh Kiehbadroudinezhad, S. Bruce Martin and Joanna Mills Flemming

Citation: The Journal of the Acoustical Society of America **150**, 3569 (2021); doi: 10.1121/10.0007063
View online: https://doi.org/10.1121/10.0007063
View Table of Contents: https://asa.scitation.org/toc/jas/150/5
Published by the Acoustical Society of America

JASA ARTICLE

Estimating minke whale relative abundance in the North Atlantic using passive acoustic sensors

Shahideh Kiehbadroudinezhad,[1,a)] S. Bruce Martin,[2,b)] and Joanna Mills Flemming[1]
[1]*Department of Mathematics and Statistics, Dalhousie University, 6299 South Street, Halifax, Nova Scotia B3H 4R2, Canada*
[2]*JASCO Applied Sciences, 32 Troop Avenue, Suite 202, Dartmouth, Nova Scotia B3B 1Z1, Canada*

ABSTRACT:
Estimates of abundance and their changes through time are key elements of marine mammal conservation and management. Absolute marine mammal abundance in a region of the open ocean is often difficult to attain. However, methods of estimating their abundance based on passive acoustic recordings are becoming increasingly employed. This study shows that passive acoustic monitoring of North Atlantic minke whales with a single hydrophone provides sufficient information to estimate relative population abundance. An automated detector was developed for minke whale pulse trains and an approach for converting its output into a relative abundance index is proposed by accounting for detectability as well as false positives and negatives. To demonstrate this technique, a 2 y dataset from the seven sites of the Atlantic Deepwater Ecosystem Observatory Network project on the U.S. east coast was analyzed. Resulting relative abundance indices confirm pulse train-calling minke whale presence in the deep waters of the outer continental shelf. The minkes are present December through April annually with the highest abundance near the site offshore of Savannah, Georgia. © *2021 Acoustical Society of America*.
https://doi.org/10.1121/10.0007063

(Received 20 November 2020; revised 8 September 2021; accepted 15 October 2021; published online 11 November 2021)
[Editor: Klaus Lucke] Pages: 3569–3580

I. INTRODUCTION

Most odontocetes and some mysticetes, including minke whales (*Balaenoptera acutorostrata*), form social networks (Risch *et al.*, 2013) in which acoustic communication plays an important role (reviewed in Erbe *et al.*, 2016). Baleen whales communicate with each other over long distances through low-frequency (15–500 Hz) acoustic signals known as vocalizations. These vocalizations are often used to study their behavior (Risch *et al.*, 2013). Some baleen whales, such as humpback whales (*Megaptera novaeangliae*) (Payne and McVay, 1971; Garland *et al.*, 2013) and bowhead whales (*Balaena mysticetus*) (Stafford *et al.*, 2018), combine individual sound units to generate a series of notes in a detectable temporal pattern similar to bird songs (Broughton, 1963; Payne and McVay, 1971). Broughton (1963) defined the concept of the song: "…a series of notes, generally of more than one type, uttered in succession and so related as to form a recognizable sequence or pattern in time." Among baleen whale species, humpback whales generate complex songs (Payne and McVay, 1971), bowhead whales produce elaborate songs (Stafford *et al.*, 2008), and blue (*Balaenoptera musculus*) and fin (*Balaenoptera physalus*) whales sing high intensity song units at very low frequencies (about 15–30 Hz) (McDonald *et al.*, 2001; Croll *et al.*, 2002). In several species, such as humpback, fin, and blue whales, only males produce songs (Glockner, 1983; Croll *et al.*, 2002; Oleson *et al.*, 2007). In most baleen whale species, both sexes produce a range of distinct call types in numerous contexts (Risch *et al.*, 2014b). North Atlantic minke whales produce low-frequency pulse trains (50–400 Hz) with varying inter-pulse interval (IPI) structure (slow-down, constant, and speed-up), low-frequency down-sweeps (118–80 Hz), and series of clicks in the 5000–6000 Hz range (Beamish and Mitchell, 1973; Martin *et al.*, 2013; Risch *et al.*, 2013; Davis *et al.*, 2017; Risch *et al.*, 2019). The behavioral understanding of the vocalizations of North Atlantic minke whales and whether they are produced by a distinct sex or age is still unexplored (Risch *et al.*, 2014b).

Most baleen whales have large home ranges [some covering up to 10 000 km annually (Davis *et al.*, 2020)], making it problematic to gather information on their ecology in a complete life-history (Risch *et al.*, 2014a). Although all baleen whales do not migrate annually, migratory populations move for different reasons, such as foraging and social behaviors. Variation in migratory behaviors with gender, age, and reproductive state, complicates our understanding of their movement patterns (Davis *et al.*, 2020).

The ranges of migratory marine mammals are currently being affected by climate change. For instance, the Gulf of Maine is one of the fastest warming bodies of water in the world and is an important feeding ground for many baleen whales. The warming waters are likely to affect the distribution of their food sources, resulting in seasonal shifts in mammal distributions throughout the western North Atlantic (Davis *et al.*, 2020). Even though migration and seasonal distributions of some species, such as humpback whale, are well studied, little is known about the winter distribution of

[a)]Electronic mail: shahideh.k@dal.ca
[b)]ORCID: 0000-0002-6681-9129.

other large baleen whale species, such as minke whales in the North Atlantic (Risch *et al.*, 2014b). The humpback whales migrate between breeding and feeding grounds within the North Atlantic. Although the migration pattern of sei whales (*Balaenoptera borealis*) is not well understood, they were observed to move from mid to low latitudes in the western North Atlantic. While fin whales are usually sighted in the western North Atlantic, blue whales are generally observed off eastern Canada in the western North Atlantic (Davis *et al.*, 2020). In many cases, such as with North Atlantic minke whales, their abundance and seasonal distribution have been studied across their summer ranges, but data on migration routes and winter habitat are virtually missing (Risch *et al.*, 2014a).

Serious shifts in the distribution and abundance of several baleen whale species have been reported. While the abundance of Central North Atlantic (CNA) humpback and fin whales has increased in the period from 1987 to 2007, the abundance of minke whales on the Icelandic continental shelf decreased in the period from 2001 to 2009 (Vakingsson *et al.*, 2015). North Atlantic minke whales are a species of Least Concern under the IUCN Red List; this species is commercially hunted in summertime (Risch *et al.*, 2019).

Minke whales are the smallest baleen whale species in the Northeast Atlantic (Risch *et al.*, 2019). Minke whales are seen in northern latitudes in the summer and migrate to warmer tropical waters to breed in the winter (Horton *et al.*, 2011; Martin *et al.*, 2013; Silva *et al.*, 2013). Minke whales have been observed throughout the North Atlantic from Baffin Bay to the Caribbean in the western North Atlantic and from the Barents Sea to the west African continental shelf in the eastern North Atlantic (Risch *et al.*, 2013).

Most information on minke whale abundance has been based on analyzing visual detections with techniques like distance sampling (Buckland, 2001), capture–recapture (Martin *et al.*, 2013), and spatially explicit capture–recapture (Marques *et al.*, 2012). Visual methods generally rely on humans to detect the animals and hence, are poorly suited to situations where the animals are difficult to see, such as during nighttime or adverse weather, as well as if the species is difficult to spot (Marques *et al.*, 2013; Risch *et al.*, 2013; Risch *et al.*, 2014b). Conversely, passive acoustic monitoring (PAM) is an indirect observation method that is effective for monitoring marine mammal populations if their vocalization types are well understood and they vocalize regularly (Marques *et al.*, 2013). PAM uses underwater hydrophones and loggers to record acoustic data over long periods of time. The recording contains contributions from different sources, such as marine life, human activities, and natural sources like wind, ice, and waves. Since many species generate readily detectable and distinguishable sounds (Risch *et al.*, 2013; Risch *et al.*, 2014b), analysis of the PAM data can yield information on the species distributions.

Manual analysis of PAM data to detect distinguishable sounds produced by marine mammals is extremely labor-intensive and time consuming, even for experienced analysts (Kowarski and Moors-Murphy, 2020). It is far more efficient to detect marine mammals' vocalizations in large data sets using automated detectors. A detailed understanding of automated detector performance is essential when using their outputs to obtain estimates of abundance.

Abundance estimation methods may provide estimates of population size (absolute abundance) or indices (relative abundance) (Eberhardt and Simmons, 1987; Chen *et al.*, 2004). As it is often difficult or impractical to estimate absolute abundance, relative abundance indices are more commonly used (Dice, 1941; Nichols and Pollock, 1983; Chen *et al.*, 2004; Hopkins and Kennedy, 2004). Relative abundance indices are of two general types: (1) an incomplete count of animal numbers, or (2) a measure of some trait related to abundance that can more easily be monitored in their environment. For marine mammals, vocalizations are such a trait (Skalski *et al.*, 2005). A fundamental assumption that must be made when using abundance methods is that the index is correlated to true population size; preferably, directly proportional to population size with a proportionality that is (relatively) constant. Validating this assumption for a relative abundance method is often impractical because it also requires a known population size (Allen and Engeman, 2015). Dice (1941:402) expressed the concept succinctly: "The difficulty of obtaining accurate counts of the number of individual mammals present on a given area has led to attempts to develop indices of abundance for the species concerned. Such indices may or may not be convertible into terms of population density. For many practical uses, however, it is sufficient to know the relative abundance of a particular species in different areas or at different times without having an exact count of the population". It demonstrates the importance of developing indices of abundance and highlights that it has many practical uses.

To estimate abundance from detected vocalizations, information on individual vocal behavior and a knowledge of the detection range for the animals' vocalizations are required (Marques *et al.*, 2013; Risch *et al.*, 2014b; Harris *et al.*, 2018). The detection range depends on the source level of the vocalizations, the local acoustic propagation conditions, the frequency characteristics of the vocalization, and the background noise level against which vocalizations can be detected (Küsel *et al.*, 2011). Marine mammal vocalization source levels are poorly understood and can be highly variable as the animals themselves respond to changes in background noise (Fournet *et al.*, 2018; Miksis-Olds *et al.*, 2019; Helble *et al.*, 2020; Thode *et al.*, 2020). The low-frequency pulse trains of minke whales (50–400 Hz) have an average source level (SL_{rms}) ranging between 164 and 168 dB re 1 μPa (Risch *et al.*, 2014b), with an estimate of a median detection range up to 30 km at the lowest median noise levels (93.09 dB re 1 μPa) for a particular habitat (i.e., Stellwagen Bank) (Risch *et al.*, 2014a). Even though this suggests that PAM devices can monitor a large area from a single site, the size of the monitored area will fluctuate over time due to the proximity of human activities, changing wind-driven ambient sound levels and/or changes in vocal behavior (Tyack, 2008; Carey and Evans, 2011; Pine *et al.*, 2018).

https://doi.org/10.1121/10.0007063

Abundance or density estimation methods, based on PAM data, need to account for these effects.

Some marine mammal density and abundance estimation methods based on PAM data have relied on propagation modeling to predict the detection volume and thus, needed to go through different analyses to arrive at estimated densities. For instance, Marques *et al.* (2009) applied digital time, acoustic, and movement DTAGs (Johnson and Tyack, 2003) to a number of beaked whale species to estimate different parameters, such as probability of detecting a cue as a function of distance between the animal and the sensor. Gathering such data is difficult and expensive and therefore, these methods are not practical for every species or study area.

Here, a method for estimating a relative abundance index of the North Atlantic minke whales calling with pulse trains on the outer continental shelf (OCS) of the North American east coast is proposed. We introduce a detectability coefficient from the hydrophone's perspective to account for noise and propagation conditions and to assess auditory masking [unwanted masking noise which prohibits a biologically important sound to be detected (Erbe *et al.*, 2016; Pine *et al.*, 2020)] at the receiver. Here, only information about the detector performance and environmental noise is used to obtain a relative index of abundance for pulse train-calling minke whales; no knowledge of the actual area/volume/detection range is required.

The paper is organized as follows: Sec. II describes the Atlantic Deepwater Ecosystem Observatory Network (ADEON) data and presents: (1) a manual analysis performed to assess the automated detector performance, (2) the automated detection algorithm, and (3) the derivation of the relative abundance index. Section III contains the results and presents: (1) detector performance and (2) relative abundance index estimates. Discussion is found in Sec. IV. Finally, Sec. V presents our conclusions.

II. METHODS

A. ADEON passive acoustic monitoring data

Data from the Atlantic Deepwater Ecosystem Observatory Network (ADEON), located along the U.S. east OCS, were used for this study. The ADEON project deployed seven autonomous long–term observatory landers in waters between 100 and 1000 m deep from Florida to Virginia (see Table I). The sensors on the landers included passive acoustics, active acoustics, conductivity-temperature-dissolved oxygen loggers, and a fish tag logger. The PAM data from November 2017–November 2019 were employed in this analysis. The duty cycle and sample rate varied by year at all the receivers: November 2017–November 2018; –44 min/h at 8000 Hz (using AMAR G3, JASCO Applied Sciences, Dartmouth, Canada) and November 2018–November 2019; –42 min/h at 16 000 Hz (using AMAR G4, JASCO Applied Sciences, Dartmouth, Canada). All data were collected using GeoSpectrum M36-V35–100 hydrophones with a nominal sensitivity of −165 dB V/μPa. Hydrophones were calibrated using a G.R.A.S. 42AA pistonphone (G.R.A.S. Sound and Vibration, Holte, Denmark) before deployment and at retrieval.

TABLE I. Deployment information at the outer continental shelf (OCS) of the North American east coast.

Station Name	Abbreviation	Location	Depth (m)
Virginia Inter-Canyon	VAC	37.246° N–74.514° W	212
Charleston Bump	CHB	32.070° N–78.374° W	404
Jacksonville	JAX	30.493° N–80.003° W	317
Hatteras South	HAT	35.200° N–75.020° W	296
Savannah Deep	SAV	32.042° N–77.348° W	790
Blake Escarpment	BLE	29.251° N–78.351° W	872
Wilmington	WIL	33.585° N–76.451° W	461

B. Manual data analysis

To evaluate the performance of the detector, manual analysis was performed to create a truth data set, consisting of minke whale vocalizations mixed with local noise. The Automated Data Selection for Validation (ADSV) algorithm [based on capturing the distributions of the number and types of automated detections per unit time (the unit of time for analysis relies on how the data were recorded, merged, or split) in the full data set (Kowarski *et al.*, 2020)] was used to select a sample of acoustic files that represent the diversity of acoustic conditions in the data for manual review. These files were fully analyzed by a trained analyst using PAMlab (JASCO Applied Sciences, Dartmouth, NS, Canada). The analyst drew annotation boundaries to include the entire vocalization in both time and frequency. Figure 1 depicts an example of a file containing minke whale pulse trains from the SAV site on 7 February 2018 at 03:35 (UTC) that was fully annotated by the analyst (yellow boxes).

C. Automated data analysis

Here, a new automated algorithm for detecting minke whale pulse trains was developed to find both strong vocalizations (from whales close to the hydrophone that were recorded with high amplitudes in the data) and faint vocalizations (from whales further away from the recorders). The strong and faint vocalizations have different characteristics, in particular, pulse train duration, for multiple reasons including attenuation and overlapping arrivals of calls from multiple individuals. Since the detectors use these characteristics to separate vocalizations from false alarms, the detector first detects strong vocalizations and then searches for faint ones to improve the probability of detecting the whales.

This algorithm requires two parameters that may depend on local conditions: the strong and faint analysis window durations (W_S and W_F). These parameters are employed to analyze the signals and their energy levels in each window, which is used to detect the desired signals from undesired ones. For this study, we obtained these values from the mean duration of strong and faint pulse trains of the manually analyzed data. A duration of $W_S = 20$ s was

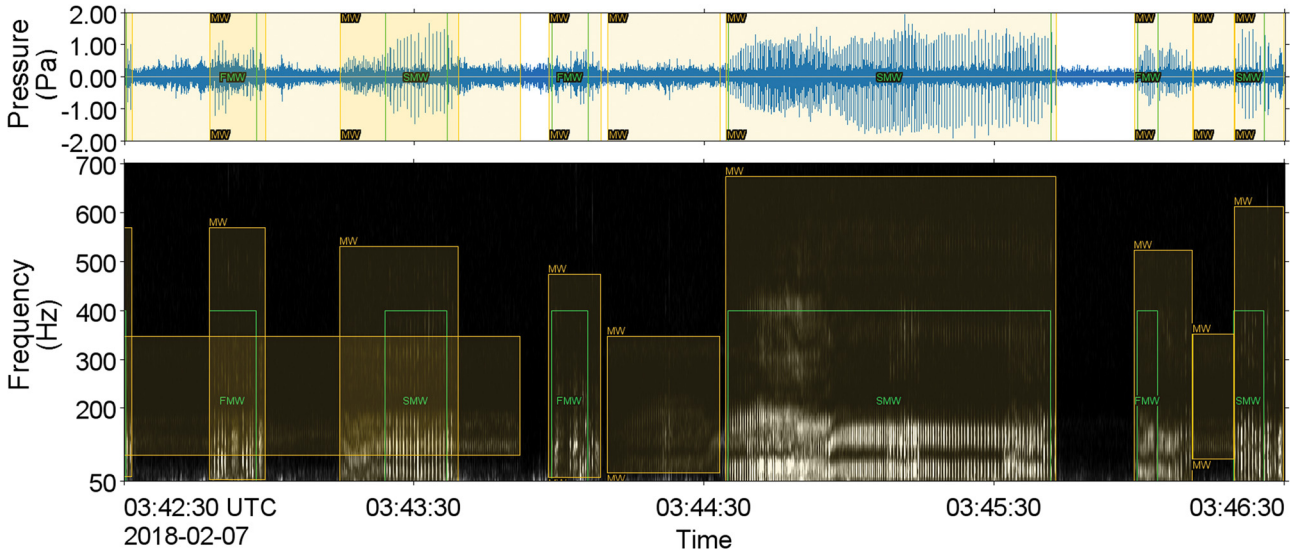

FIG. 1. (Color online) Example of a file from the Savannah Georgia (SAV) site on 7 February 2018 at 03:42:30 (UTC) that was fully annotated by the analyst (yellow boxes). Detected pulse trains annotated with green boxes overlaid on a 2.5 min section of the same manually annotated data (yellow boxes). MW on each yellow box stands for minke whale. SMW and FMW on each green box stand for strong and faint minke whales, respectively.

used for the strong vocalizations and $W_F = 8$ s was employed for the faint pulse trains. Since the pulse train durations for strong/faint signals can be longer or shorter than the mean values (20 s and 8 s), the detector used 5 s and 70 s for the minimum and maximum pulse train durations, respectively [e.g., see the pulse train duration of the second strong minke whale (SMW) in Fig. 1]. A block diagram of the algorithm is shown in Fig. 2.

To count the number of pulse trains per file, raw data were first bandpass filtered to remove energy not associated with the minke whale pulse trains. While the pulse train energy peaks in the 50–300 Hz range (Risch *et al.*, 2014b), bandpass filtering in the range of 120–400 Hz better removes environmental noise and sounds from other species while preserving the energy from the pulse trains. The lower cut-off frequency of 120 Hz decreases the proportion of false positives from vessels and thus, improves the detection results. The upper cut-off frequency of 400 Hz was selected based on the bandwidth of the manual analysis results. The bandpass filter was a minimum-order finite impulse response filter and performs zero phase filtering with a stop band attenuation of 60 dB (designed in MATLAB, Natick, MA).

The detector identifies pulse trains by detecting consecutive periods of high amplitudes interspersed with periods of low amplitudes. The threshold for high amplitude signals is adaptively set for each data file. To find the threshold, the data were divided into windows, each W_S seconds long, and the peak sound pressure levels in each window was computed. The peak values were then sorted in descending order. The mean value of the first n_1 peak sound pressure levels (here $n_1 = 5$) was found and stored as parameter T_1. The mean value of the lowest n_2 peak values (here $n_2 = 5$) was stored as parameter T_2. Based on the concept of information entropy, the threshold for the strong signal ($Thresh_S$) is calculated as

$$Thresh_S = T_2 \times \log 2\left(1 + \frac{T_1}{T_2}\right). \quad (1)$$

The number of peaks to average (n_1 and n_2) depends on the file length analyzed (here 9 or 11 min) and the length of the analysis windows (W_S, W_F). The n_1 and n_2 parameters allow the analyst to reduce the proportion of false positives by moderating the effects of outlier peak sound pressure levels.

The filtered time series signal was compared to the threshold $Thresh_S$, and all values below this threshold ($Thresh_S$) were set to zero. The time duration of consecutive zeros was compared to the inter-pulse interval (IPI, t_{IPI}) which is between 0.29 ± 0.02 and 0.83 ± 0.04 s (mean values $\pm$ SD) (Risch *et al.*, 2013). If the duration was within the bounds for consecutive minke whale pulses, the location was considered inside a pulse train; otherwise, the location was considered outside the pulse train. The number of pulses (N_P) and pulse train duration (T) were compared to the bounds from Risch *et al.* (2013) and the manually analyzed data, respectively, and accordingly, the strong pulse train counter increased by one for each valid train.

To find the faint pulse trains, all periods that contained strong pulses were removed from the original time series, and the process was repeated using W_F rather than W_S. For the faint case, the threshold $Thresh_F$ is the mean of all peak sound pressure levels (P_F), since there was a slight variation in the signal amplitudes,

$$Thresh_F = mean(P_F). \quad (2)$$

Figure 1 depicts a 2.5 min window from the annotated file on 7 February 2018 at 03:42:30 (yellow boxes) overlaid with detected pulse trains (green boxes) and displayed PAMlab.

https://doi.org/10.1121/10.0007063

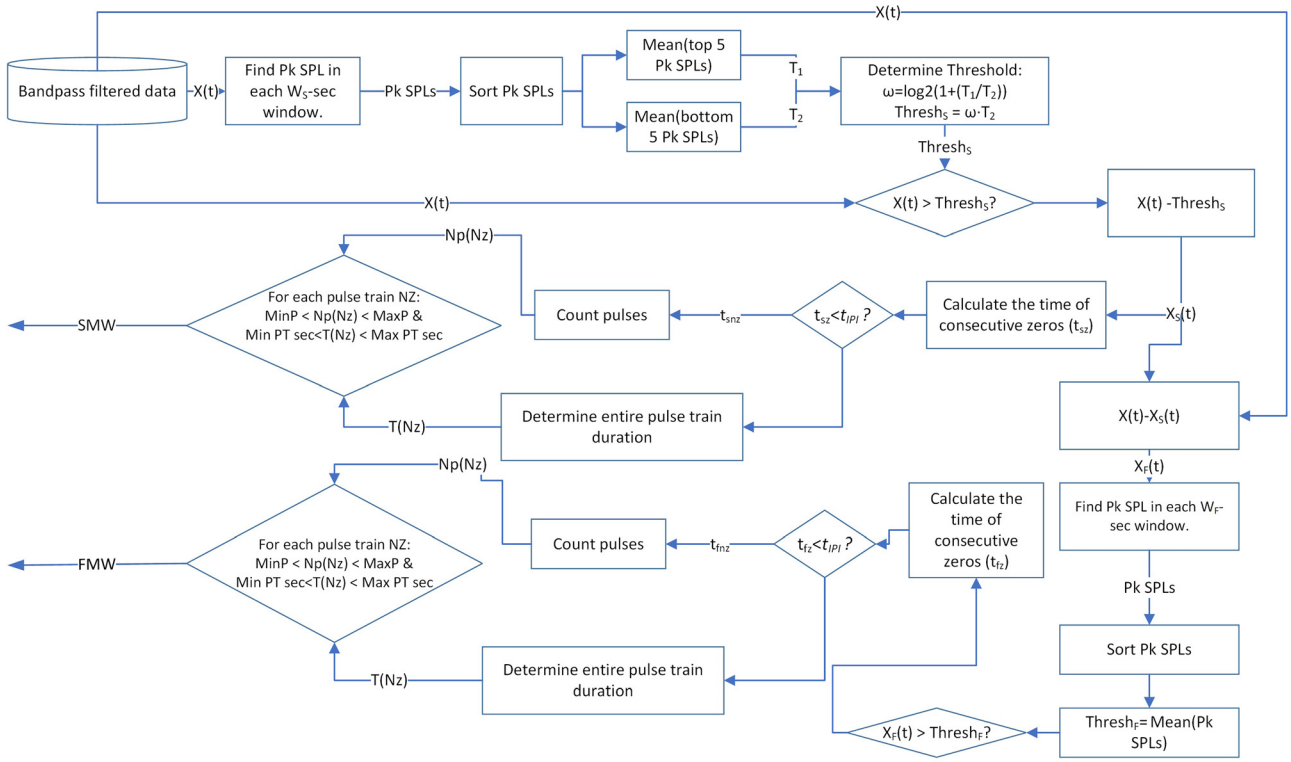

FIG. 2. (Color online) Block diagram of the automated detector algorithm. SMW and FMW stand for strong and faint minke whales, respectively.

D. Abundance indices

In recent years, counts of animal vocalizations in PAM recordings have been used to estimate animal density using a canonical density estimator equation (Mellinger *et al.*, 2007; Marques *et al.*, 2009; Küsel *et al.*, 2011; Martin *et al.*, 2013; Hildebrand *et al.*, 2019),

$$\hat{D} = \frac{n_c\,(1-\hat{c})}{k\pi w^2 \hat{P} T \hat{r}}. \tag{3}$$

In this approach, the estimated density of marine mammals ($\hat{D}$) is calculated from the number of vocalizations (n_c) corrected by the estimated average proportion of false positives ($\hat{c}$), divided by the number of sensors (k), study time (T), the estimated mean probability of detecting the vocalization in the area surveyed by each sensor ($\hat{P}$), the area surveyed (πw^2 where w is the maximum detection radius), and the estimated vocalization production rate ($\hat{r}$). In this formula, $n_c\,(1-\hat{c})/\hat{P}$ is the number of detected vocalizations corrected for estimated false positives and estimated detection probability, hereafter referred to as the estimated abundance of vocalizations. The estimated abundance of vocalizations divided by $kT\hat{r}$ gives the estimated number of marine mammals per unit time and recorder. The last factor, $1/\pi w^2$, normalizes the estimated number of animals by the detection area to obtain a density of animals. Sometimes, the estimated average proportion of false positives ($\hat{c}$) is called a false positive ratio; specific issues related to the use of this parameter are discussed in Sec. II D [refer to Eqs. (7) and (8)]. Estimating the various parameters in Eq. (3) is complex and requires substantial levels of data collection and analysis or simulation and modeling (Küsel *et al.*, 2011; Martin *et al.*, 2013).

The probability of detection of a marine mammal mainly depends on its distance from the sensor. The following equation is used to estimate the probability of detection (Marques *et al.*, 2009; Küsel *et al.*, 2011):

$$\hat{P} = \int_0^w \hat{g}(y)h(y)dy, \tag{4}$$

where y is the horizontal distance, $\hat{g}(y)$ is the detection function, and $h(y)$ is the distribution of horizontal distances to all animals, detected or not.

The estimated probability of detecting an animal can be estimated by a few methods. Possible ways to estimate this parameter include employing hydrophones capable of localizing the vocalizations or using tagged animals to know how far they are from the recorder when detected. Due to complexity and cost, these methods generally yield small sample sizes, and thus, $\hat{P}$ is generally estimated once and treated as fixed for a full data set (Marques *et al.*, 2012; Martin *et al.*, 2013). Modeling approaches, such as coupling Monte Carlo simulations with acoustic propagation models (Küsel *et al.*, 2011), or bearing estimation and propagation modeling (Harris *et al.*, 2018), are accessible (although complex) methods for estimating $\hat{P}$. However, while they are able to incorporate a representation for the variability of the local environment, they are still often used to estimate a single value for $\hat{P}$. For example, in Harris *et al.* (2018), the 90th percentile of the source distribution and 10th percentile of

the noise distribution were used to define the maximum detection range, and this detector function was employed to determine $\hat{P}$. However, because the noise level fluctuates over short time scales due to human activities and at daily to weekly time scales due to weather (Risch *et al.*, 2014a; Martin *et al.*, 2019), a new estimate of $\hat{P}$ should be considered. In the proposed method, the elements of $\hat{P}$ (the parameters that are used to estimate P) that represent the probability of missed detections are combined with $\hat{c}$ as a part of the estimated abundance of vocalizations. The remaining elements of $\hat{P}$ [refer to Eq. (3) and Marques *et al.*, 2009] are replaced with a detectability coefficient (Ω) that accounts for noise and acoustic propagation conditions.

The proposed equation for the relative abundance index $\hat{A}$ is

$$\hat{A} = \Omega \frac{n_{ca}\,(1 - PFP + PFN)}{T\hat{r}}, \tag{5}$$

where Ω is the detectability coefficient, n_{ca} is the total number of pulse trains obtained from the automated detection algorithm, T is the study time, $\hat{r}$ is the estimated vocalization production rate, PFP is the proportion of false positives, and PFN is the proportion of false negatives. Note that if Ω is set to 1, Eq. (5) estimates the number of audible minke whales. The rationale for using PFP, PFN, a definition for Ω, and the appropriate value for $\hat{r}$ is discussed below.

In Eq. (3), vocalization counts are converted to absolute abundance using a known vocalization rate, $\hat{r}$. These rates are often determined from acoustic tag data (DTAGs) (Johnson and Tyack, 2003). For minke whales, Risch *et al.* (2014b) provided the vocalization rate as the mean with a minimum and maximum value which would lead to estimates of absolute abundance with a wide prediction interval. However, for the relative abundance index developed here, one only requires the mean value of 48.6 (vocalizations/h) [$\hat{r} = 48.6 \pm 27.5$ (vocalizations/h) (mean $\pm$ SD), any value for $\hat{r}$ would work since it simply scales the final result (Risch *et al.*, 2014b)]. This value could even change as a function of time per day for each season and vocalization type (Delarue, 2008; Wang *et al.*, 2015; Garcia *et al.*, 2019). Further explanation is in the Discussion section.

The total number of pulse trains (n_{ca}) is

$$n_{ca} = n_{cas} + n_{caf}, \tag{6}$$

where n_{cas} and $\boldsymbol{n_{caf}}$ represent the number of pulse trains for the strong and faint signals, respectively. The proportion of false positives (PFP) may be calculated as

$$PFP = \frac{n_{FP}}{n_{ca}}, \tag{7}$$

where n_{FP} is the total number of false positives across all annotated files and similarly, n_{ca} is the total number of pulse train vocalizations identified by the automated detector. To account for the elements of $\hat{P}$ that represent the probability of missed detections, we include the proportion of false negatives (PFN) as a second corrector to n_{ca},

$$PFN = \frac{n_{FN}}{n_{cA}}, \tag{8}$$

where n_{FN} is the total number of false negative detections and n_{cA} is the total number of pulse trains obtained from the manual analysis.

It should be mentioned that PFP is different from false positive rate (FPR), defined as

$$FPR = \frac{n_{FP}}{n_{FP} + n_{TN}}, \tag{9}$$

where n_{FP} and n_{TN} are the number of false positives and true negatives, respectively. Since finding n_{TN} is infeasible, PFP was used in this study.

The elements of $\hat{P}$ and w that account for noise and acoustic propagation conditions in Eq. (3) are replaced with the detectability coefficient (Ω) in Eq. (5). The proposed approach to account for variable noise levels is similar to the concept of listening range reduction [LRR, reductions in animal communication space/range (communication space is used when the volume or area surrounding the animal is considered and the term "range" is used for a single distance) (Pine *et al.*, 2020)]. Calculating the communication space/range of a marine mammal requires knowledge of its vocalizing source level, the noise level, and the propagation loss. However, computing the listening range reduction requires only the change in noise level (Δ) and the geometric spreading coefficient [N, Eq. (7), Pine *et al.* (2018)],

$$LRR\,(\%) = 100\left(1 - 10^{-(\Delta/N)}\right). \tag{10}$$

For assessing LRR, Δ is the difference between the current sound level and a masking noise level. Here, we are interested in finding the change in available detection range for sounds at a recorder, so Eq. (10) was slightly modified, giving the detectability coefficient as

$$\Omega = 10^{\Delta/N}, \tag{11}$$

where Δ is the change in sound level defined as the difference between the current received sound level and the nth percentile level of the sound pressure level over the entire recording duration. The 10th percentile of the distribution of measured sound levels was selected as the realistic noise floor based on experience for the sites under study (see Fig. 3) and is the same threshold employed in Harris *et al.* (2018). N is the geometric spreading loss coefficient value that ranges between spherical spreading ($N = 20$) and cylindrical spreading ($N = 10$) (Ainslie *et al.*, 2014); an intermediate term appropriate for many environments of $N = 15$ dB was used in this analysis (Risch *et al.*, 2014b). The noise level data were 1 min sound pressure levels, summed over a frequency band encompassing the six decidecade bands at center frequencies of 125, 160, 200, 250, 315, and 400 Hz, that match the frequencies used for detection (see Sec. II C). The noise levels varied by station (Fig. 3) and in time, which demonstrates why the detectability coefficient needs to

https://doi.org/10.1121/10.0007063

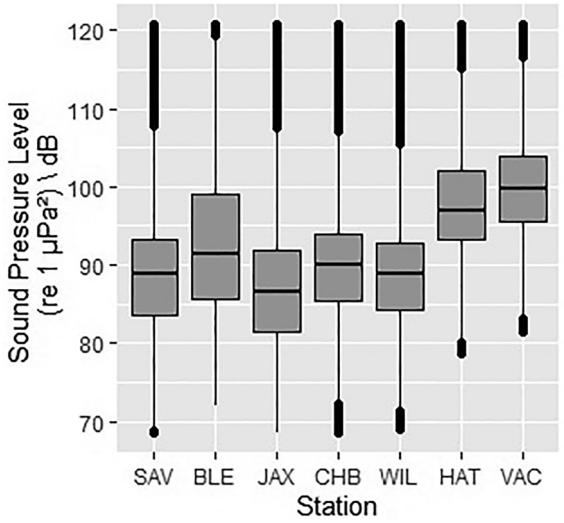

FIG. 3. The distribution received 1 min sound pressure levels over a frequency band encompassing the 125–400 Hz decidecade bands at each station for December 2017–June 2018. The noise levels varied by station and in time; hence, the detectability coefficient was developed to allow a comparison across stations. In this figure, the boxes show the interquartile range (i.e., the middle half of the distribution). The horizontal line in the box is the median value. The vertical lines show the range of values for 25% of the data above or below the middle half. The dots above or below the line indicate outlier values.

adapt to background noise levels. To suppress the short-term effects of passing vessels and focus on variations due to weather, the median value of the detectability coefficient per day for each station was used in this analysis.

III. RESULTS

A. Detector performance

To evaluate the detector performance, all pulse trains in 1% of the Savannah Deep (SAV) (see Table I) 2017–18 dataset and each of the 2018–19 datasets were manually annotated, yielding a wide range of pulse train counts (Table II).

The signal-to-noise ratio for each annotation for SAV 2017–18 was computed by summing the energy over the signal bandwidth for the annotation (signal) and an equal time period just before and just after an annotated pulse train (to define the noise, ratios shown in Fig. 4). Negative SNRs

TABLE II. Manual analysis of the outer continental shelf (OCS) of the North American east coast minke whale data. The total number of detected pulse trains and the total number of files annotated as determined by the ADSV algorithm per station are presented as n_t and n, respectively.

Station name	Time period	n_t	n
SAV	November 2017–June 2018	780	137
VAC	November 2018–October 2019	2	113
CHB	November 2018–October 2019	154	173
HAT	November 2018–October 2019	26	166
SAV	November 2018–October 2019	390	168
BLE	November 2018–October 2019	248	181
WIL	November 2018–October 2019	218	167

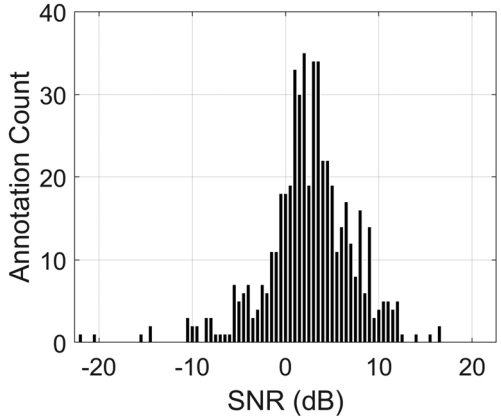

FIG. 4. Distribution of the manual annotations as a function of signal-to-noise ratio (SNR) at the Savannah Georgia (SAV) site.

occurred when there was another pulse train in the noise time window before or after the annotation (e.g., Fig. 1).

To estimate the performance of our detector, the *precision* (positive predictive value), *recall* (sensitivity or true predictive rate), and F_1 *score* (the harmonic mean of the precision and recall) indicators were calculated based on the following equations (Hand and Christen, 2017):

$$\text{Precision} = \frac{n_{TP}}{(n_{TP} + n_{FP})}, \tag{12}$$

$$\text{Recall} = \frac{n_{TP}}{(n_{TP} + n_{FN})}, \tag{13}$$

$$\text{F}_1 = 2 \times \frac{Precision \times Recall}{(Precision + Recall)}. \tag{14}$$

Overall, the precision, recall, and F_1 score of the detector against the SAV data for the period of November 2017–June 2018 were 0.8, 0.8, and 0.8, respectively. The value for *PFP* in Eq. (7) was 0.15 and the value for the false negative scaling parameter (*PFN*) in Eq. (8) was 0.22. The values for the other stations are illustrated in Table III. It should be mentioned that high values of PFP in Table III are due to rarely observed/detected minke whales at the stations, such as VAC and HAT.

B. Relative abundance index estimator

Equation (5) includes PFN, which diverges from the typical canonical Eq. (3). To assess whether PFN improved

TABLE III. Proportion of false positive (PFP) and false negative (PFN) values at outer continental shelf (OCS).

Station Name	Time period	PFP[a]	PFN[a]
VAC	November 2018–October 2019	0.80	0.00
CHB	November 2018–October 2019	0.70	0.37
HAT	November 2018–October 2019	0.96	0.4
SAV	November 2018–October 2019	0.17	0.22
BLE	November 2018–October 2019	0.36	0.17
WIL	November 2018–October 2019	0.51	0.30

[a]Poor results at some stations for PFP and PFN are discussed in Sec. IV.

algorithm performance, the estimated number of audible minke whales ($\hat{N}_M$) was compared using different versions of Eq. (6) (Fig. 5). For simplicity in the comparisons, the detectability coefficient Ω was set to 1. Three versions of Eq. (5) were evaluated: (1) $\hat{N}_{MA}$ was computed using the number of annotated calls (which served as a reference value), (2) $\hat{N}_{Ma1}$ was computed using the number of automated detections and setting the proportion of false negative (PFN) to 0, and (3) $\hat{N}_{Ma2}$ was computed using the number of automated detections and including PFN. The analysis was performed on the 137 fully annotated files from SAV for the period from November 2017–June 2018. The root mean square errors (RMSE) [Eq. (15)] were used to compare performance of the methods to the estimated number of audible minke whales ($\hat{N}_{MA}$), where a lower RMSE indicates better performance (i is a sum over all files analyzed),

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(\hat{N}_{Mai} - N_{MAi}\right)^2}{n}}. \quad (15)$$

Both versions of Eq. (5) closely tracked the manual annotation results (Fig. 5), but the version with PFN was preferred with a lower RMSE of 0.35 compared to 0.38 without PFN. The result in Fig. 5 shows that adding PFN improves the estimated number of minke whales. Therefore, separating the signals into strong and faint and subsequently, considering faint signals, decreases the number of missed detections and makes it closer to the true value, which eventually provides a better estimate for abundance index.

Equation (5) was applied to the full 2 y dataset at site SAV. The 8 d moving average relative abundances of minke whales using the daily median number of audible whales and detectability coefficients are illustrated in Fig. 6. The median value was selected rather than the mean for each day as it suppressed false positives from the detector and thus, avoided overestimating the relative abundance of minke whales. The values of PFP and PFN in Table III were used in the calculation.

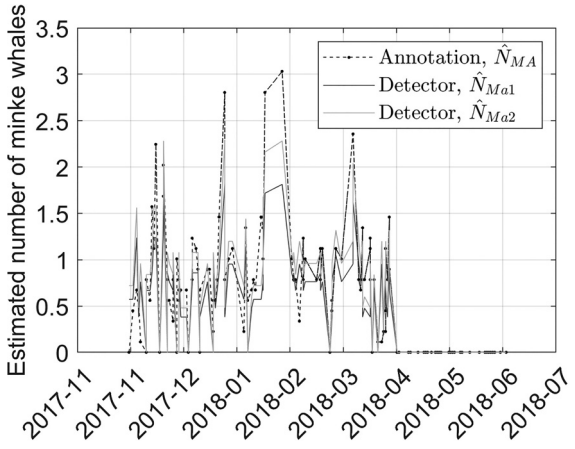

FIG. 5. Estimated number of audible minke whales using annotations and the automated detector algorithm at SAV. $\hat{N}_{MA}$ was computed using the number of annotated calls, $\hat{N}_{Ma1}$ was computed using the number of automated detections and setting the proportion of false negative (PFN) in Eq. (5) to zero. $\hat{N}_{Ma2}$ was computed using the number of automated detections including PFN.

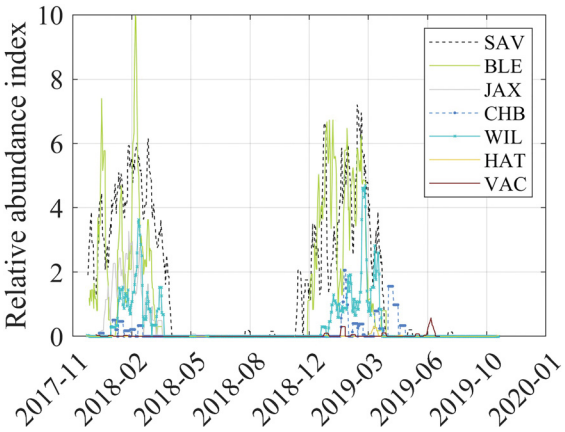

FIG. 6. (Color online) An 8 d moving average for relative abundance index of audible minke whales for 2 y at the seven stations using Eq. (5).

To summarize the results and provide a spatial and temporal interpretation, the relative abundance index was summed by month per year and the average of 2 y was calculated and plotted as variable sized bubbles by station (Fig. 7). Minke whales are present in deep waters of the OCS from December through April and have a peak concentration near the site offshore of Savannah Georgia (SAV) in both years. They also show a high concentration at Blake Escarpment (BLE) station and are present at Jacksonville (JAX), Charleston Bump (CHB), and Wilmington (WIL). They were rarely present at Hatteras (HAT) and Virginia Inter-Canyon (VAC). The results also indicate a northward migration starting in February through April.

IV. DISCUSSION

The goal of this work was to propose a method for converting marine mammal vocalizations detected in passive acoustic recordings at single hydrophones into a comparable index of relative abundance. The method was demonstrated using pulse trains of North Atlantic minke whales. Little is known about how minke whale pulse train vocalizations relate to the age, sex, season, and location of the whales (Risch *et al.*, 2014b), and thus, the detection of audible whales is a relative abundance estimate.

Results show a distinct spatial distribution that evolved in time. These may be compared to the density surface models for the minke whales in the OCS from Roberts *et al.* (2016). The density surface models indicate an even distribution of minke whales throughout the ADEON area in winter, with very few present in the summer months (Duke University, 2021). The ADEON data show a gradient with fewer whales to the west and north of the study area in winter, a movement north towards the end of the detection period, and replicates the density surface models that show no presence in the summer months. The northward movement also agrees with the previous minke whale presence

https://doi.org/10.1121/10.0007063

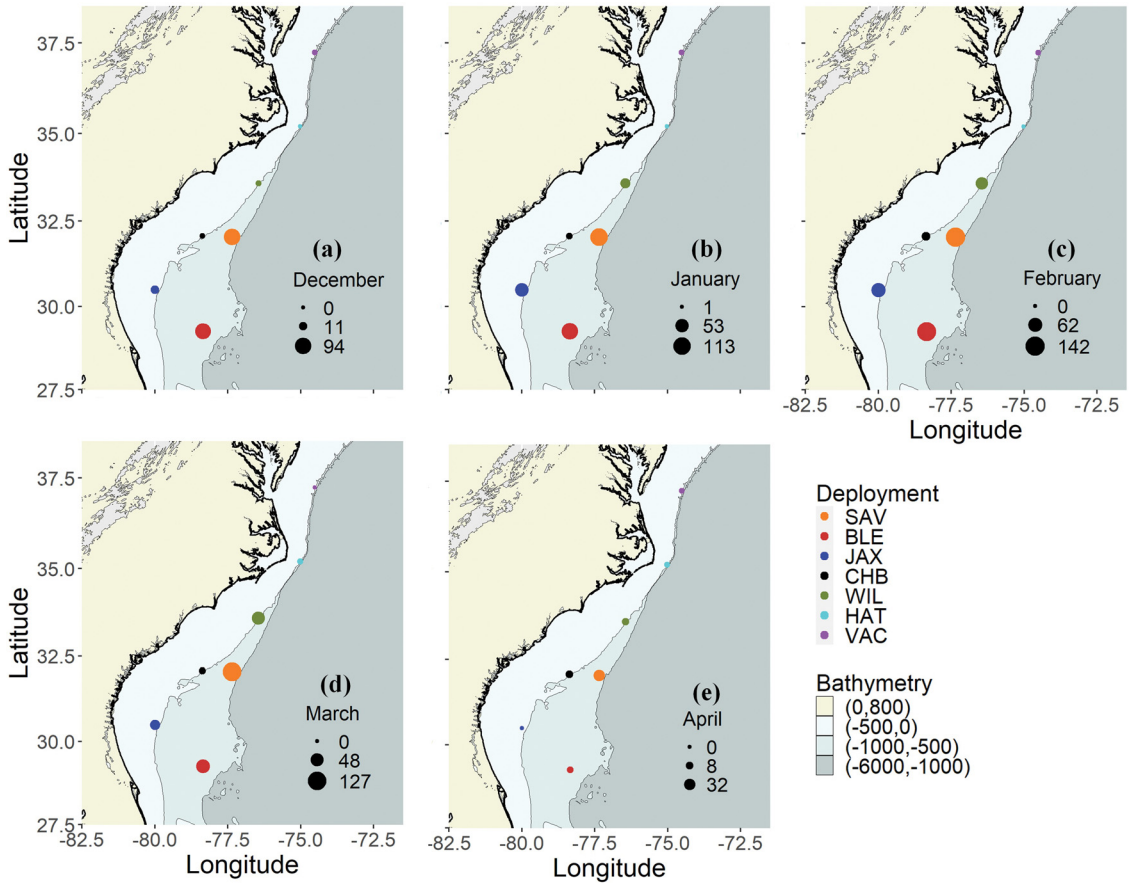

FIG. 7. (Color online) The relative abundance index at the ADEON sites in a) December, b) January, c) February, d) March, and e) April. The maximum relative abundance index values were 94, 113, 142, 127, and 32 for December, January, February, March, and April, respectively. The minimum value of 0 represents the absence of minke whales in March, April, and February at VAC, JAX, and HAT, respectively. The best available minke whales density results from Roberts *et al.*, 2016 were 731, 715, 691, 750, and 1498 in December, January, February, March, and April, respectively, at the east coast (Duke University 2021). The peak concentrations were towards south at the ADEON sites in December–March and was towards north in April.

trends from PAM data (Risch *et al.*, 2013). The cessation of detections in April is likely due to a northward movement of the animals, but also a change in vocalization patterns as the minke whales stop their wintertime mating behaviors and move north to feed for the summer.

There are many refinements and conventions that will need to be developed for this method to become sufficiently standardized so that it can be used for management and conservation. Many of these refinements/conventions will still require extensive research to quantify mammal behaviors. A few of those issues are introduced here in the following paragraphs.

In this analysis, we employed the mean vocalization rate per hour from Risch *et al.* (2014b) to convert the vocalizations counted into an estimated number of audible whales; however, as we are proposing a relative abundance index, any value for $\hat{r}$ would work since it simply scales the final result. To apply the method to different time periods and species, it is important to arrive at an agreed-upon standard relative number of vocalizations per unit time in different seasons that can be employed by all research groups. This value could even change as a function of time per day for each season and vocalization type (Delarue, 2008; Wang *et al.*, 2015; Garcia *et al.*, 2019). In the present analysis, it is likely that a different vocalization rate should be employed in summer than in winter, but further research is needed to determine the relative occurrence of pulse trains in different seasons and habitat areas.

Here, we applied a constant value for the attenuation factor (N) in Eq. (11) since we had *a priori* knowledge that propagation conditions were in a broad sense similar for all the ADEON sites (based on water depths of 200–900 m, and ranges of interest at most 30 km). However, it is simple to imagine PAM survey designs where this would not be the case, such as recorders located on the shelf, on the continental slope at the sound-speed axis, and others in deep water far below the sound-speed axis (e.g., Delarue *et al.*, 2018). To accommodate these differences, site specific values for N could be employed. Further work is needed to make recommendations on how to select a value of N for different situations (Reeder and Chiu, 2010; Erbe *et al.*, 2012; Ainslie *et al.*, 2014; Pine *et al.*, 2020). There are other situations in which it may make sense to reduce the attenuation factor by a fixed multiplier, such as if a recorder were located very close to a steep shelf break such that it could only hear over one-half of a circle. Moreover, the Δ parameter, which is the

change in sound level, is an important multiplier required for any sound level–based method of abundance/density estimation. This parameter can be as varied as any event, such as noise or signals produced by the animals, that might be detected acoustically by our deployed hydrophones.

The pulse train algorithm used in this analysis is fast, straightforward, and flexible. It was well suited to the detection of the choruses of minke whale pulse trains that often lasted for hours to days at a time. The pulse trains were also separated into two groups: faint and strong pulse trains. Faint signals possess different characteristics than strong ones, and a detector that can account for this variation improves the estimate of the number of pulse trains. Overlapping vocalizations between singers (e.g., Fig. 1) complicated the assessment of detector performance and led to the recommendation to include the proportion of false negatives as a parameter in the abundance equation. We also expect that the false negative parameter will be important when this method is applied to non-chorusing species, such as North Atlantic right whales, and when an algorithm suffers from a high proportion of false negatives. Occasionally, the algorithm generated large numbers of false positives over short periods of time, for example, when vessels passed by. In the current situation, we know that when minke whales are present, vocalizations are detected for extended periods. Therefore, we choose to use the median number of vocalizations per day as our vocalization count as it was able to greatly reduce the influence of the false positives; similar results have been employed to reduce false alarms during real-time monitoring with gliders (Baumgartner *et al.*, 2013; Kowarski *et al.*, 2020). A different approach would be required if the false positives could be more common than true positives. North Atlantic right whale (NARW) is a species where this occurs, since both vessels and humpback whales are common false triggers for the NARW upcall detectors. In such cases, either improvements in the detector or human-in-the-loop protocols are employed (Kowarski *et al.*, 2020).

Long minke whale pulse trains occur at regular intervals and cause an overlapping in frequency and timing of vocal activity, as overlapping vocalizations are likely from different individuals. This variable vocalization structure, associated with high underwater noise levels, proved to be a major challenge for detection algorithms (Risch *et al.*, 2013). Here, with low proportions of false negatives, and a relatively reasonable low proportion of false positives, at stations in where minke whales are abundant (see the Results and Discussion sections), the pulse train detector used in this study performed well for our spatial occurrence patterns in the survey area.

The performance of the algorithm could be improved by including other vocalizing rate parameters that identify the different types of North Atlantic pulse train-calling minke whales. These could include start and end inter-pulse interval (IPI) (s; averaged over the first and last 20% of pulses, respectively), change in pulse rate [1/IPI (1/s); difference of averaged values for first and last 20%], and change in frequencies over time (Hz; difference of averaged first and last 20% mean spectrum measures). Future work is needed on minke whale call types, individual calling behavior, and source levels in the OCS. This information is essential for interpreting the social interaction of minke whales and estimating their numbers. To significantly improve abundance estimates for this cryptic species, is it necessary to attempt a better insight when determining the strength and weaknesses of the approach. For instance, determining what parameters must be added to the detector to develop its performance to enhance the results, or how to improve the results through obtaining the relative abundance index of minke whales by separating strong and faint pulse trains, would be beneficial.

The original canonical Eq. (3) and the proposed extension/simplification in Eq. (5) both depend on the performance of automated detectors that count marine mammal vocalizations. The minke whale vocalizations do not have standardized detectors; hence, the development of a new one was warranted. Detectors that work well for one location often need to be tuned to work in another, or they will have a substantially different performance (see Table III). Thus, we must expect variable performance from detectors, that different groups will use different detectors, and that even within research groups, detectors will evolve. To combine an abundance index from different research groups, years, and locations, we need a standardized manner for logging detections and describing detector performance. This is analogous to the concepts of collecting standardized data about visual detections and measuring the detection function "g" during line-transect studies (Roberts *et al.*, 2016).

V. CONCLUSION

To implement and enhance management and conservation policies for marine mammals, reliable and effective abundance estimates are necessary. Some density/abundance estimation methods for marine mammals that are based on passive acoustic detections employ propagation modeling to estimate the area over which animals may be heard. Accurate acoustic propagation modeling depends on environmental data that are difficult to obtain, and hence requires highly specialized software and expertise. These resources are not available to many practitioners. Therefore, this study aimed to demonstrate a new method for obtaining a relative abundance index from single hydrophone data that does not need source level and propagation modeling. The method was demonstrated using North Atlantic pulse train-calling minke whales.

A dataset from the Atlantic Deepwater Ecosystem Observatory Network (ADEON) located along the U.S. east coast outer continental shelf (OCS), was used for the effort. The results facilitated the comparison of the relative importance of the seven sites, with the deep sites at SAV and BLE being preferred. Temporally, the results showed the pulse train-calling minke whales present from winter to mid-spring. This agrees with the known wintering of pulse

train-calling minke whales in southern waters (Risch *et al.*, 2014a).

Future work is needed to capture minke whale vocalization types, individual vocalizing behavior, and source levels in the OCS more accurately. This information is essential for interpreting the social interaction of minke whales and estimating their numbers. Application of the method with different species and integration of the method with directional data will also need further investigation to confirm the usefulness of the approach.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Jennifer L. Miksis-Olds for her helpful advice and comments. This work was supported by Mitacs through the Mitacs Accelerate program under Contract No. IT13606. Study concept, oversight, and funding for ADEON were provided by the U.S. Department of the Interior, Bureau of Ocean Energy Management, Environmental Studies Program, Washington, DC under Contract No. M16PC00003, in partnership with other National Ocean Protection Partnership funding agencies. Funding for ship time was provided under separate contracts by ONR, Code 32.

Ainslie, M. A., Dahl, P. H., de Jong, C. A. F., and Laws, R. M. (**2014**). "Practical Spreading Laws: The Snakes and Ladders of Shallow Water Acoustics," *UA2014 - 2nd International Conference and Exhibition on Underwater Acoustics*, June 22–27, Island of Rhodes, Greece, pp. 879–886.

Allen, L. R., and Engeman, R. M. (**2015**). "Evaluating and validating abundance monitoring methods in the absence of populations of known size: Review and application to a passive tracking index," Environ. Sci. Pollut. Res. Int. **22**(4), 2907–2915.

Baumgartner, M. F., Fratantoni, D. M., Hurst, T. P., Brown, M. W., Cole, T. V. N., Van Parijs, S. M., and Johnson, M. (**2013**). "Real-time reporting of baleen whale passive acoustic detections from ocean gliders," J. Acoust. Soc. Am. **134**(3), 1814–1823.

Beamish, P., and Mitchell, E. (**1973**). "Short pulse length audio frequency sounds recorded in the presence of a Minke whale (*Balaenoptera acutorostrata*)," Deep-Sea Res. Oceanogr. Abstr. **20**(4), 375–380.

Broughton, W. (**1963**). *Acoustic Behaviour of Animals* (Elsevier, Amsterdam).

Buckland, S. T., Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D. L., and Thomas, L. (**2001**). *Introduction to Distance Sampling: Estimating Abundance of Biological Populations* (Oxford University Press, Oxford, UK).

Carey, W. M., and Evans, R. B. (**2011**). *Ocean Ambient Noise: Measurement and Theory* [The Underwater Acoustics Series] (Springer, New York). https://doi.org/10.1007/978-1-4419-7832-5.

Chen, J., Thompson, M. E., and Wu, C. (**2004**). "Estimation of fish abundance indices based on scientific research trawl surveys," Biometrics **60**(1), 116–123.

Croll, D. A., Clark, C. W., Acevedo, A., Tershy, B., Flores, S., Gedamke, J., and Urban, J. (**2002**). "Only male fin whales sing loud songs," Nature **417**(6891), 809–809.

Davis, G. E., Baumgartner, M. F., Bonnell, J. M., Bell, J., Berchok, C., Bort Thornton, J., Brault, S., Buchanan, G., Charif, R. A., Cholewiak, D., Clark, C. W., Corkeron, P., Delarue, J., Dudzinski, K., Hatch, L., Hildebrand, J., Hodge, L., Klinck, H., Kraus, S., Martin, B., Mellinger, D. K., Moors-Murphy, H., Nieukirk, S., Nowacek, D. P., Parks, S., Read, A. J., Rice, A. N., Risch, D., Širović, A., Soldevilla, M., Stafford, K., Stanistreet, J. E., Summers, E., Todd, S., Warde, A., and Van Parijs, S. M. (**2017**). "Long-term passive acoustic recordings track the changing distribution of North Atlantic right whales (*Eubalaena glacialis*) from 2004 to 2014," Sci. Rep. **7**(1), 1–12.

Davis, G. E., Baumgartner, M. F., Corkeron, P. J., Bell, J., Berchok, C., Bonnell, J. M., Bort Thornton, J., Brault, S., Buchanan, G. A., Cholewiak, D. M., Clark, C. W., Delarue, J., Hatch, L. T., Klinck, H., Kraus, S. D., Martin, B., Mellinger, D. K., Moors-Murphy, H., Nieukirk, S., Nowacek, D. P., Dawn, P., Pegg, N., Read, A. J., Rice, A. N., Risch, D., Scott, A., Soldevilla, M., Stafford, K. M., Stanistreet, J. E., Summers, E., Todd, S., and Van Parijs, S. M. (**2020**). "Exploring movement patterns and changing distributions of baleen whales in the western North Atlantic using a decade of passive acoustic data," Glob. Change Biol. **26**(9), 4812–4840.

Delarue, J. J.-Y. (**2008**). "Northwest Atlantic fin whale vocalizations: Geographic variations and implications for stock assessments," Master of Philosophy (Human Ecology) thesis, College of the Atlantic, Bar Harbor, ME.

Delarue, J., Kowarski, K. A., Maxner, E. E., MacDonnell, J. T., and Martin S. B. (**2018**). Acoustic Monitoring Along Canada's East Coast: August 2015 to July 2017. Document No. 01279, Environmental Studies Research Funds Report No. 215, Version 1.0. Technical report by JASCO Applied Sciences for Environmental Studies Research Fund, Dartmouth, NS, Canada.

Dice, L. R. (**1941**). "Methods for estimating populations of mammals," J. Wildl. Manage. **5**(4), 399–407.

Duke University. (**2021**). "Mapping tool for cetacean density for the U.S. Atlantic and Northern Gulf of Mexico [webpage]," https://seamap.env.duke.edu/models/mapper/USECGOM (Last viewed: February 2021).

Eberhardt, L. L., and Simmons, M. A. (**1987**). "Calibrating population indices by double sampling," J. Wildl. Manage. **51**(3), 665–675.

Erbe, C., MacGillivray, A., and Williams, R. (**2012**). "Mapping cumulative noise from shipping to inform marine spatial planning," J. Acoust. Soc. Am. **132**(5), EL423–428.

Erbe, C., Reichmuth, C., Cunningham, K., Lucke, K., and Dooling, R. (**2016**). "Communication masking in marine mammals: A review and research strategy," Mar. Pollut. Bull. **103**(1-2), 15–38.

Fournet, M. E. H., Matthews, L. P., Gabriele, C. M., Haver, S., Mellinger, D. K., and Klinck, H. (**2018**). "Humpback whales *Megaptera novaeangliae* alter calling behavior in response to natural sounds and vessel noise," Mar. Ecol. Prog. Ser. **607**, 251–268.

Garcia, H. A., Zhu, C., Schinault, M. E., Kaplan, A. I., Handegard, N. O., Godø, O. R., Ahonen, H., Makris, N. C., Wang, D., and Ratilal, P. (**2019**). "Temporal–spatial, spectral, and source level distributions of fin whale vocalizations in the Norwegian Sea observed with a coherent hydrophone array," ICES J. Mar. Sci. **76**, 268.

Garland, E. C., Gedamke, J., Rekdahl, M. L., Noad, M. J., Garrigue, C., and Gales, N. (**2013**). "Humpback whale song on the Southern Ocean feeding grounds: Implications for cultural transmission," PLoS One **8**(11), e79422.

Glockner, D. (**1983**). "Determining the sex of humpback whales (*Megaptera novaeangliae*) in their natural environment," in *Communication and Behavior of Whales*, edited by R. S. Payne (Westview Press, Boulder, CO) Vol. 76, pp. 447–464.

Hand, D., and Christen, P. (**2018**). "A note on using the F-measure for evaluating record linkage algorithms," Stat. Comput. **28**(3), 539–547.

Harris, D. V., Miksis-Olds, J. L., Vernon, J. A., and Thomas, L. (**2018**). "Fin whale density and distribution estimation using acoustic bearings derived from sparse arrays," J. Acoust. Soc. Am. **143**(5), 2980–2993.

Helble, T. A., Guazzo, R. A., Martin, C. R., Durbach, I. N., Alongi, G. C., Martin, S. W., Boyle, J. K., and Henderson, E. E. (**2020**). "Lombard effect: Minke whale boing call source levels vary with natural variations in ocean noise," J. Acoust. Soc. Am. **147**(2), 698–712.

Hildebrand, J. A., Frasier, K. E., Baumann-Pickering, S., Wiggins, S. M., Merkens, K. P., Garrison, L. P., Soldevilla, M. S., and McDonald, M. A. (**2019**). "Assessing seasonality and density from passive acoustic monitoring of signals presumed to be from pygmy and dwarf sperm whales in the Gulf of Mexico," Frontiers. Dataset.

Hopkins, H. L., and Kennedy, M. L. (**2004**). "An assessment of indices of relative and absolute abundance for monitoring populations of small mammals," Wildl. Soc. Bull. **32**(4), 1289–1296.

Horton, T. W., Holdaway, R. N., Zerbini, A. N., Hauser, N., Garrigue, C., Andriolo, A., and Clapham, P. J. (**2011**). "Straight as an arrow: Humpback whales swim constant course tracks during long-distance migration," Biol. Lett. **7**(5), 674–679.

Johnson, M. P., and Tyack, P. L. (**2003**). "A digital acoustic recording tag for measuring the response of wild marine mammals to sound," IEEE J. Ocean. Eng. **28**(1), 3–12.

Kowarski, K. A., Gaudet, B. J., Cole, A. J., Maxner, E. E., Turner, S. P., Martin, S. B., Johnson, H. D., and Moloney, J. E. (**2020**). "Near real-time marine mammal monitoring from gliders: Practical challenges, system development, and management implications," J. Acoust. Soc. Am. **148**(3), 1215–1230.

Kowarski, K. A., and Moors-Murphy, H. (**2020**). "A review of big data analysis methods for baleen whale passive acoustic monitoring," Mar. Mammal Sci. (Early View) **37**(2), 652–673.

Küsel, E. T., Mellinger, D. K., Thomas, L., Marques, T. A., Moretti, D., and Ward, J. (**2011**). "Cetacean population density estimation from single fixed sensors using passive acoustics," J. Acoust. Soc. Am. **129**(6), 3610–3622.

Marques, T. A., Thomas, L., Ward, J., DiMarzio, N., and Tyack, P. L. (**2009**). "Estimating cetacean population density using fixed passive acoustic sensors: An example with Blainville's beaked whales," J. Acoust. Soc. Am. **125**(4), 1982–1994.

Marques, T. A., Thomas, L., Martin, S. W., Mellinger, D. K., Jarvis, S., Morrissey, R. P., Ciminello, C.-A., and DiMarzio, N. (**2012**). "Spatially explicit capture–recapture methods to estimate minke whale density from data collected at bottom-mounted hydrophones," J. Ornithol. **152**(S2), 445–455.

Marques, T. A., Thomas, L., Martin, S. W., Mellinger, D. K., Ward, J. A., Moretti, D. J., Harris, D., and Tyack, P. L. (**2013**). "Estimating animal population density using passive acoustics," Biol. Rev. **88**(2), 287–309.

Martin, S. B., Morris, C., Bröker, K., and O'Neill, C. (**2019**). "Sound exposure level as a metric for analyzing and managing underwater soundscapes," J. Acoust. Soc. Am. **146**(1), 135–149.

Martin, S. W., Marques, T. A., Thomas, L., Morrissey, R. P., Jarvis, S., DiMarzio, N., Moretti, D., and Mellinger, D. K. (**2013**). "Estimating minke whale (*Balaenoptera acutorostrata*) boing sound density using passive acoustic sensors," Mar. Mammal Sci. **29**(1), 142–158.

McDonald, M. A., Calambokidis, J., Teranishi, A. M., and Hildebrand, J. A. (**2001**). "The acoustic calls of blue whales off California with gender data," J. Acoust. Soc. Am. **109**(4), 1728–1735.

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**(4), 36–45.

Miksis-Olds, J. L., Harris, D. V., and Heaney, K. D. (**2019**). "Comparison of estimated 20-Hz pulse fin whale source levels from the tropical Pacific and Eastern North Atlantic Oceans to other recorded populations," J. Acoust. Soc. Am. **146**(4), 2373–2384.

Nichols, J. D., and Pollock, K. H. (**1983**). "Estimation methodology in contemporary small mammal capture-recapture studies," J. Mammal. **64**(2), 253–260.

Oleson, E. M., Wiggins, S. M., and Hildebrand, J. A. (**2007**). "Temporal separation of blue whale call types on a southern California feeding ground," Anim. Behav. **74**(4), 881–894.

Payne, R. S., and McVay, S. (**1971**). "Songs of humpback whales," Science **173**(3997), 585–597.

Pine, M. K., Hannay, D. E., Insley, S. J., Halliday, W. D., and Juanes, F. (**2018**). "Assessing vessel slowdown for reducing auditory masking for marine mammals and fish of the western Canadian Arctic," Mar. Pollut. Bull. **135**, 290–302.

Pine, M. K., Nikolich, K., Martin, B., Morris, C., and Juanes, F. (**2020**). "Assessing auditory masking for management of underwater anthropogenic noise," J. Acoust. Soc. Am. **147**(5), 3408–3417.

Reeder, D. B., and Chiu, C. S. (**2010**). "Ocean acidification and its impact on ocean noise: Phenomenology and analysis," J. Acoust. Soc. Am. **128**(3), EL137–EL143.

Risch, D., Clark, C. W., Dugan, P. J., Popescu, M., Siebert, U., and Van Parijs, S. M. (**2013**). "Minke whale acoustic behavior and multi-year seasonal and diel vocalization patterns in Massachusetts Bay, USA," Mar. Ecol. Prog. Ser. **489**, 279–295.

Risch, D., Castellote, M., Clark, C. W., Davis, G. E., Dugan, P. J., Hodge, L. E. W., Kumar, A., Lucke, K., Mellinger, D. K., Risch, D., Castellote, M., Clark, C. W., Davis, G. E., Dugan, P. J., Hodge, L. E. W., Kumar, A., Lucke, K., Mellinger, D. K., Nieukirk, S. L., Popescu, C. M., Ramp, C., Read, A. J., Rice, A. N., Silva, M. A., Siebert, U., Stafford, K. M., Verdaat, H., and Van Parijs, S. M. (**2014a**). "Seasonal migrations of North Atlantic minke whales: Novel insights from large-scale passive acoustic monitoring networks," Mov. Ecol. **2**, 24.

Risch, D., Siebert, U., and Van Parijs, S. M. (**2014b**). "Individual calling behaviour and movements of North Atlantic minke whales (Balaenoptera acutorostrata)," Behaviour **151**(9), 1335–1360.

Risch, D., Wilson, S. C., Hoogerwerf, M., van Geel, N. C. F., Edwards, E. W. J., and Brookes, K. L. (**2019**). "Seasonal and diel acoustic presence of North Atlantic minke whales in the North Sea," Sci. Rep. **9**, 3571.

Roberts, J. J., Best, B. D., Mannocci, L., Fujioka, E., Halpin, P. N., Palka, D. L., Garrison, L. P., Mullin, K. D., and Cole, T. V. N. (**2016**). "Habitat-based cetacean density models for the U.S. Atlantic and Gulf of Mexico," Sci. Rep. **6**.

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**(10), e76507.

Skalski, J. R., Ryding, K. E., and Millspaugh, J. (**2005**). *Wildlife Demography: Analysis of Sex, Age, and Count Data* (Academic Press, Cambridge, MA).

Stafford, K. M., Moore, S. E., Laidre, K. L., and Heide-Jorgensen, M. P. (**2008**). "Bowhead whale springtime song off West Greenland," J. Acoust. Soc. Am. **124**(5), 3315–3323.

Stafford, K. M., Lydersen, C., Wiig, Ø., and Kovacs, K. M. (**2018**). "Extreme diversity in the songs of Spitsbergen's bowhead whales," Biol. Lett. **14**(4): 20180056.

Thode, A. M., Blackwell, S. B., Conrad, A. S., Kim, K. H., Marques, T., Thomas, L., Oedekoven, C. S., Harris, D., and Bröker, K. (**2020**). "Roaring and repetition: How bowhead whales adjust their call density and source level (Lombard effect) in the presence of natural and seismic airgun survey noise," J. Acoust. Soc. Am. **147**(3), 2061–2080.

Tyack, P. L. (**2008**). "Implications for marine mammals of large-scale changes in the marine acoustic environment," J. Mammal. **89**(3), 549–558.

Vakingsson, G. A., Pike, D. G., Valdimarsson, H., Schleimer, A., Gunnlaugsson, T., Silva, T., Elvarsson, B., Mikkelsen, B., Øien, N., Desportes, G., Bogason, V., and Hammond, P. S. (**2015**). "Distribution, abundance, and feeding ecology of baleen whales in Icelandic waters: Have recent environmental changes had an effect?," Front. Ecol. Evol. **3**, 1.

Wang, Z.-T., Nachtigall, P. E., Akamatsu, T., Wang, K.-X., Wu, Y.-P., Liu, J.-C., Duan, G.-Q., Cao, H.-J., and Wang, D. (**2015**). "Passive acoustic monitoring the diel, lunar, seasonal and tidal patterns in the biosonar activity of the Indo-Pacific humpback dolphins (*Sousa chinensis*) in the Pearl River Estuary, China," PLoS ONE **10**(11), e0141807.