

Temporal and spatial characteristics of the Beaufort Sea ambient noise environment

R. Chen and H. Schmidt

Citation: [The Journal of the Acoustical Society of America](#) **148**, 3928 (2020); doi: 10.1121/10.0002955

View online: <https://doi.org/10.1121/10.0002955>

View Table of Contents: <https://asa.scitation.org/toc/jas/148/6>

Published by the [Acoustical Society of America](#)

ARTICLES YOU MAY BE INTERESTED IN

[Quantifying the contribution of ship noise to the underwater sound field](#)

The Journal of the Acoustical Society of America **148**, 3863 (2020); <https://doi.org/10.1121/10.0002922>

[A seminal paper linking ocean acoustics and physical oceanography](#)

The Journal of the Acoustical Society of America **148**, R9 (2020); <https://doi.org/10.1121/10.0002761>

[Vertical directionality and spatial coherence of the sound field in glacial bays in Hornsund Fjord](#)

The Journal of the Acoustical Society of America **148**, 3849 (2020); <https://doi.org/10.1121/10.0002868>

[Source depth estimation using spectral transformations and convolutional neural network in a deep-sea environment](#)

The Journal of the Acoustical Society of America **148**, 3633 (2020); <https://doi.org/10.1121/10.0002911>

[An acoustic remote sensing method for high-precision propeller rotation and speed estimation of unmanned underwater vehicles](#)

The Journal of the Acoustical Society of America **148**, 3942 (2020); <https://doi.org/10.1121/10.0002954>

[Virtual head waves in ocean ambient noise: Theory and modeling](#)

The Journal of the Acoustical Society of America **148**, 3836 (2020); <https://doi.org/10.1121/10.0002926>



**Advance your science and career
as a member of the
ACOUSTICAL SOCIETY OF AMERICA**

[LEARN MORE](#)



Temporal and spatial characteristics of the Beaufort Sea ambient noise environment^{a)}

R. Chen^{b)} and H. Schmidt

Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

ABSTRACT:

Underwater ambient noise from the Beaufort Sea, collected in March 2016, is analyzed to investigate the effect that environmental changes in the region have on the spatial and temporal characteristics of the ice-generated ambient noise. In particular, the influx of warm Pacific water, so called the Beaufort Lens, has dramatically altered the level and vertical directionality of ambient noise, creating a low noise zone in the ~75–250 m depth interval with a noise notch at low grazing angles. It is also demonstrated how the observed noise vertical directionality is consistent with ice-mechanical activity along an active pressure ridge ~30–50 km from the recording array, as shown by satellite imagery. The discrete ranges to this ice activity explain the peak arrival angles between –10 to –15 degrees in noise vertical directionality, as well as the horizontal noise notch. Transient noise events associated with the ice activity are analyzed using an image processing approach with hierarchical clustering applied to the recorded spectrograms. The observed events are grouped into three categories—short-time-broadband, long-time-narrowband, and long-time-broadband—each likely generated through a different mechanism by the ice cover. The spectral and temporal distribution of these transients are discussed. © 2020 Acoustical Society of America.

<https://doi.org/10.1121/10.0002955>

(Received 31 August 2020; revised 3 November 2020; accepted 1 December 2020; published online 22 December 2020)

[Editor: Peter F. Worcester]

Pages: 3928–3941

I. INTRODUCTION

The Beaufort Sea region of the Arctic ocean is a dynamic environment undergoing significant changes in recent decades. Most notably, an influx of warm Pacific water¹ known as the Beaufort Lens (BL) has dramatically altered the underwater sound speed profile (SSP). This warm layer is neutrally buoyant at ~60–80 m depth and causes a local sound speed maximum in the water column (Fig. 1). As a result, this SSP (hence referred to as the BL SSP) creates a double duct propagation environment with channels above and below the local maximum. In the surface channel, acoustic waves encounter frequent interactions with the ice-water interface, resulting in severe attenuation. In contrast, the lower duct has been shown to promote long range propagation by effectively trapping sound above 300 Hz.^{2–5} The effect of this ducted environment on acoustic propagation will likely increase in significance in the future. For instance, with the modeled rise in Arctic Ocean pH, sound is predicted to propagate even further within the lower duct due to decreased attenuation.⁶ Thus, continued study of this phenomenon and its effects are of critical importance to improving our ability to characterize the Arctic Ocean soundscape and successfully conduct underwater acoustic operations in the region.

Another major environmental shift in the Beaufort Sea, as with all of the Arctic Ocean, is the diminishing ice cover. Both the spatial extent and thickness of Arctic ice cover have decreased in recent decades.^{7–9} The percentage of multi-year ice has declined, replaced by first-year ice that is more fragile and susceptible to ridging.¹⁰ These changes consequently affect the underwater ambient noise environment as the ice cover is the most significant source of noise generation in the region. Historically, ambient noise in the Arctic Ocean is primarily generated through a myriad of mechanisms by the ice cover. These include small scale cracking due to temperature, wind, or current forcing,^{11–13} pure or varying tonals from rubbing and shearing between different ice floes,^{14–17} broadband transients from large scale lead openings,¹⁴ as well as short-time squeaks and transients from ice breakup near pressure ridges or marginal ice zones.¹⁸ More recently, analysis of ambient noise data from the 2016 U.S. Navy Ice Exercise in the Arctic Ocean (ICEX16) shows that the younger and thinner ice cover during the experiment may be less capable of withstanding ice sheet shearing to produce sustained tonal noise features.¹⁹ Furthermore, noise generation in the ice cover may have become more spatially discrete, focused predominately along pressure ridge formations rather than distributed over the ice sheet.¹⁹ In turn, transient noise generation may be affected as well, particularly in their temporal distribution and frequency composition. These findings, paired with the effects of the BL SSP on propagation, highlight the significant impact that environmental change has on the Arctic underwater noise environment.

^{a)}This paper is part of a special issue on Ocean Acoustics in the Changing Arctic.

^{b)}Electronic mail: ruic@mit.edu, ORCID: 0000-0001-7790-2553.

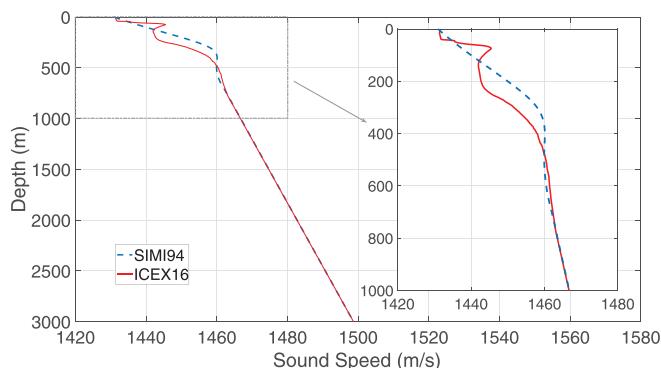


FIG. 1. (Color online) ICEX16 SSP compared with SIMI94 SSP. The ICEX16 case demonstrates the effect of the BL with local sound speed maximum near 75 m depth. The upper 800 m of both SSPs are based off of measured data, the rest are extrapolated by a monotonic increase with depth. Figure from Chen *et al.* (Ref. 19).

This paper examines some of the observed effects on ambient noise due to the described environmental changes. It complements a previous study¹⁹ on Beaufort Sea noise collected during ICEX16 by verifying and extending its analyses. In Sec. II, we provide a brief overview of the ICEX16 experimental setup. This is followed by a more in-depth description of the underwater SSP and surface ice cover during the experiment. In Sec. III, we use acoustic modeling to demonstrate the effects that the BL SSP and ICEX16 ice cover have on underwater ambient noise and explain features observed in the collected noise data. In Sec. IV, we propose an alternative approach to transient event detection and characterization, as is presented in the previous study.¹⁹ This new method allows for classification of different event types based on bandwidth and duration, which provides us with more insight into the mechanism by which transient events are generated by ice cover as well as the statistics of specific event types.

II. ENVIRONMENTAL FACTORS DURING ICEX16

As part of ICEX16, underwater ambient noise data were collected in the Beaufort Sea with the goal of characterizing climate induced effects on the region's underwater acoustic environment. The data were recorded with a 32-element vertical line array (VLA) with nested spacing ~ 1.5 m between the outer 10 elements and 0.75 m between the inner 22 elements. The sampling frequency was 12 000 Hz. Approximately eight hours of ambient noise was recorded with the array at 38 m depth on March 13, 2016, and another ~ 3.5 h was recorded with the array varying in depth down to 238 m at 25 m increments on March 14, 2016. A more detailed description of the experiment can be found in Chen *et al.*, 2019.¹⁹

A. SSP

The SSP measured during ICEX16 is shown in Fig. 1. It demonstrates the effect of the BL with a local sound speed maximum at ~ 75 m depth. This is in contrast to the

traditional monotonically increasing SSP observed in earlier Arctic environments. For example, Fig. 1 also shows the SSP measured during the SIMI94 experiment, which took place near the exact same location as ICEX16 in the Beaufort Sea, but 22 years prior.

While the ICEX16 SSP is a good representation of the new BL SSP, it is only a snapshot of a dynamic environment. To better understand the temporal variability of the BL SSP, we examined 2331 profiles measured near the ICEX16 camp location between January 2014 and January 2020 as part of the Woods Hole Oceanographic Institution's ice-tethered profiler program.²⁰ The position of each SSP measurement is shown in Fig. 2(a) with respect to the ICEX16 camp location. The yearly mean of the measured profiles, as well as the sound speed standard deviation with depth, are shown in Fig. 2(b). We observe that in most years, the largest temporal variability in the BL SSP occurs around the depth of the local sound speed maximum and just below the lower duct. We further note that while the BL SSP typically has a single, sharp local maximum, it is possible for the profile to exhibit a broader peak or even multiple peaks. These features can be seen in the more recent profile averages from 2018 to 2020. In fact, the mean of the 26 profiles collected in January 2020 shows two distinct peaks instead of just one. In deciding what SSP to use for our modeling analyses in this study, we selected the measured ICEX16 SSP when analyzing data collected during the experiment and used the mean of all measured profiles shown in Fig. 2(b) when modeling the Beaufort Sea environment in general. However, the observed increase in the number of peaks seen in the more recently measured profiles demonstrates that the Beaufort Sea is a dynamic region and invites continued observation to ensure we can accurately describe its undersea environment in the future.

B. Ice cover

A general trend in the Arctic ocean has been the decline in ice cover thickness and the extent of multi-year ice.¹⁰ This was certainly the case during ICEX16. The camp itself was positioned on top of first year ice with thickness of $\sim 1\text{--}2$ m. In addition, unlike the traditional assumption that ambient noise generation in the ice cover is distributed uniformly over the ice sheet, noise generation during ICEX16 was likely focused at discrete locations along a pressure ridge formation.¹⁹ This ridge can be observed on the ice-temperature satellite imagery of the camp site from March 13, 2016.²¹ Figure 3 shows that at $\sim 30\text{--}50$ km to the north-northeast of the ICEX16 camp, a line of higher ice temperature indicates the presence of an ice ridge. As a result, in modeling noise generation in our analyses, we use a discrete noise approach separate from and in addition to a distributed noise approach. For the distributed noise case, we use the Kuperman-Ingenito (KI) noise distribution,^{22,23} which consists of an uniform probability distribution of monopole sources. The simulation is implemented using OASES,²⁴ in which the source distribution is placed at depth of

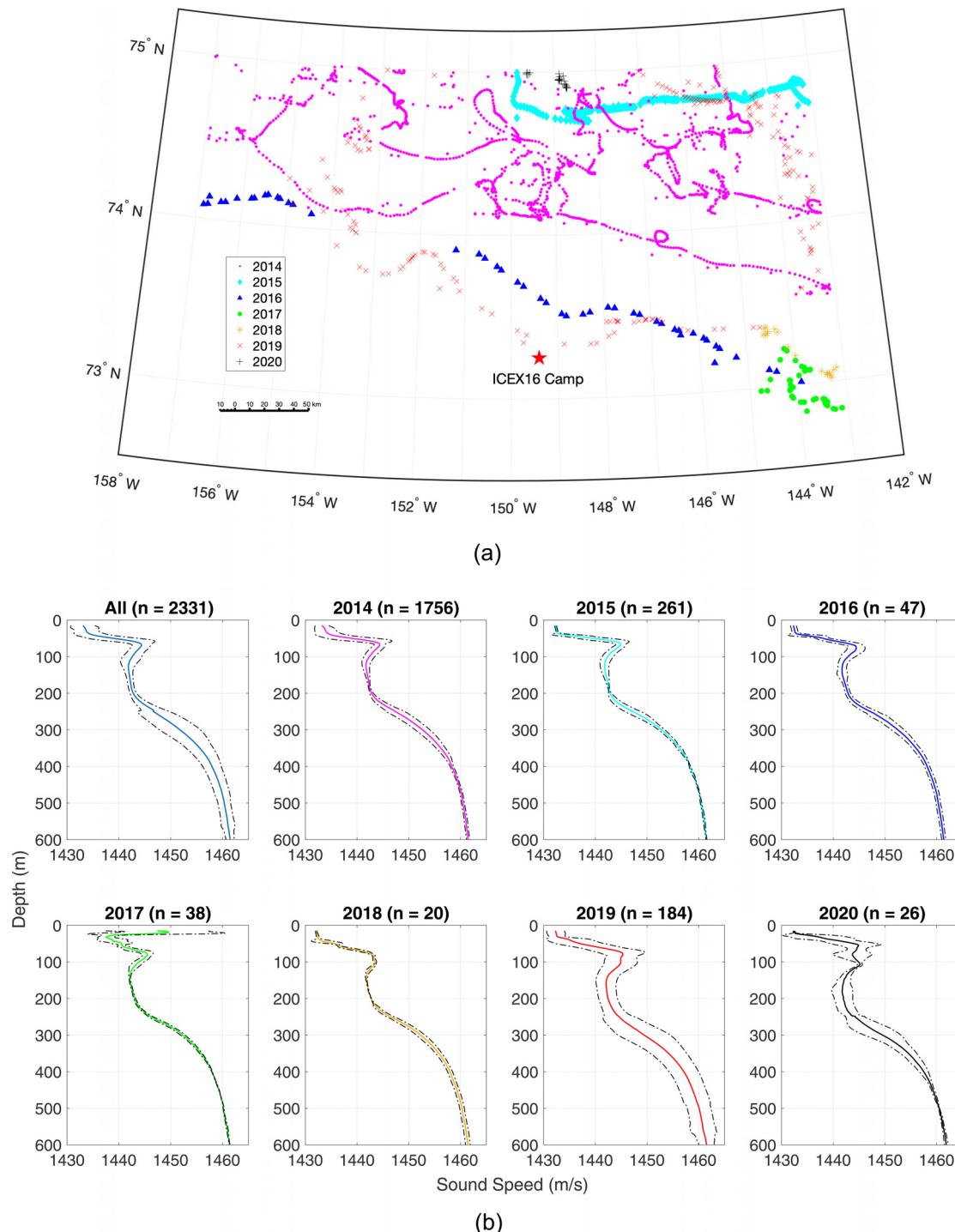


FIG. 2. (Color online) (a) Locations of measured SSPs as part of the ice-tethered profiler program from 2014 to early 2020. Most of the profiles near the ICEX16 location were collected north of the camp site. (b) Mean (solid lines) and ± 2 standard deviation (dashed-dotted lines) of the measured profiles. The first panel shows all profiles while the rest show profiles collected in a single year. The n values in the panel titles denote the number of profiles that were collected during that year.

$\lambda_{top}/30 + \eta_{rms}$ in the water column, where λ_{top} is the surface wavelength and η_{rms} is the root-mean-square roughness of the ice-water interface. For the discrete noise case, we follow a similar approach as the uniform case and place a single, near surface (0.26 m below the ice), monopole source at the range of expected noise generation. Again, the simulation is done using OASES. In this study, our goal is not to

model the exact value of noise level at different frequencies; instead, we are interested in how the noise level changes, relatively, with depth. Thus, the particular choice of the modeled source strength is not important and we use an arbitrary source strength for both cases. For the modeled environment, we used a range independent setup that consists of a 1 m surface ice layer, a 3000 m water column, and a solid

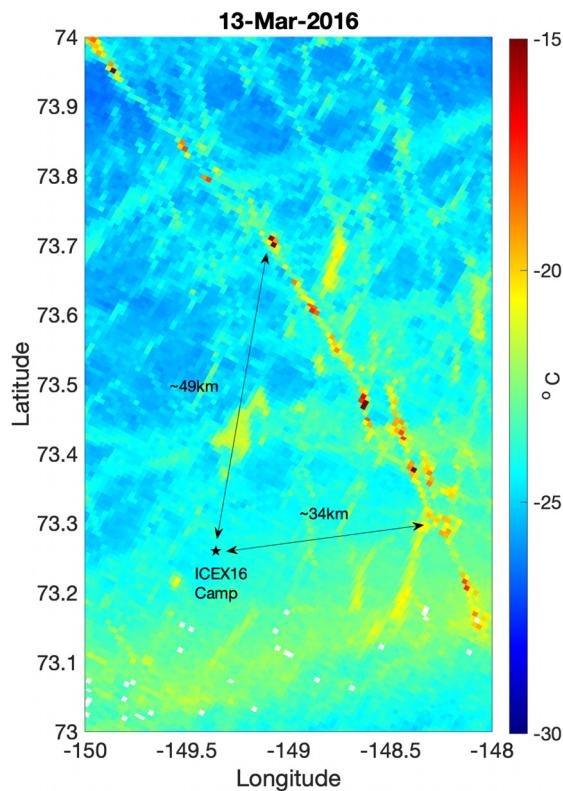


FIG. 3. (Color online) Ice-temperature satellite imagery of the ICEX16 camp site on March 13, 2016. A line of higher ice temperature to the north-northeast of the camp indicates the presence of a pressure ridge.

bottom halfspace. The parameters within these layers are shown in Table I.

III. IMPACT ON AMBIENT NOISE

The changes to the Beaufort Sea SSP and surface ice cover have pronounced impacts on the ambient soundscape. In this section, we explore the effects on noise level with depth and vertical directionality using acoustical modeling. We further compare our modeling results with ICEX16 data to explain features we observe in the collected data.

A. Ambient noise level

First, to isolate the effect of the BL SSP on ambient noise level, we compare modeled noise levels with depth in two cases. The first uses the monotonically increasing SIMI94 SSP, while the second uses the BL SSP [average of all measured profiles in Fig. 2(b)]. For surface noise generation, both cases use the KI (uniform) noise distribution with

source frequency at 850 Hz as an example (the modeled noise profiles at other frequencies show similar shapes). Figure 4(a) shows the modeled noise levels in the top 250 m for these two cases. The left side of this plot shows that with the SIMI94 SSP, noise level decreases consistently with depth. A slight knee in the SSP at 100 m depth may trap some acoustic rays that propagate at very shallow angles near the surface; however, in general, the water column becomes quieter with depth in a mostly linear trend. In contrast, The right side of Fig. 4(a) shows that with the BL SSP, the KI-generated noise profile (black line) has higher noise level above the sound speed local maximum and dips quickly with depth near the top of lower duct (outlined by dashed red lines). This result makes sense as the sharp local maximum in the BL SSP more effectively traps surface generated noise near the top of the water column than the SIMI94 profile. Consequently, less noise is able to propagate into the lower duct, creating a more distinct drop in noise level in comparison.

Next, we analyze the effect of discrete surface noise generation on the ambient noise level. By using a discrete noise source, the modeled noise level becomes inherently range dependent. As a result, we plot the outputs for various ranges as shown by the magenta lines on the right side of Fig. 4(a). The BL SSP is used so that we can compare the modeled noise profiles with the uniform noise case (black line). From this plot, we see that the shapes of the discrete source noise level profiles are similar to the uniform noise case. Due to the BL SSP, the noise level is again higher in the surface duct and dips in the lower duct. However, one key difference between the two cases is that for the discrete case, depending on the source range, the noise level within the lower duct changes. For example, at 3.5 km range (solid magenta line), the noise level within the lower duct shows a larger decrease from the surface noise level than the profiles for the source at 15 and 35 km (dotted and dashed-dotted lines). Furthermore, the profile for the source at 5 km shows an increase in noise level about half way down the lower duct near 150 m. These features can be explained by examining the transmission loss plot of a near surface source in the BL SSP environment, shown in Fig. 4(b). At 3.5 km, there is a distinct shadow zone in the lower duct. Below the duct, the noise level increase again near 225 m depth due to convergence zone propagation. Similarly, at 5 km, the increase in noise level within the lower duct is due to convergence zone propagation refracting back to the surface. Further out in range, at 15 and 35 km, the shadow zone in the lower duct becomes less prominent and the convergence

TABLE I. Parameters used in environmental modeling. C_p denotes compressional speed, C_s denotes shear speed, α_p denotes compressional attenuation, α_s denotes shear attenuation, ρ denotes density, λ denotes spatial wavelength, η_{rms} denotes root-mean-square roughness, η_{cl} denotes roughness correlation length. Parameter values are selected based on discussions in Hope *et al.* (Ref. 25) and Jensen *et al.* (Ref. 26).

Medium	C_p	C_s	α_p	α_s	ρ	η_{rms}	η_{cl}
Ice layer	3600 m/s	1800 m/s	0.216 dB/ λ	0.648 dB/ λ	0.9 g/cm ³	0.2 m	19.1 m
Water column	SSPs in Fig. 4	—	1.68×10^{-5} dB/ λ	—	1.0 g/cm ³	—	—
Bottom halfspace	2200 m/s	1500 m/s	0.5 dB/ λ	0.5 dB/ λ	2.9 g/cm ³	—	—

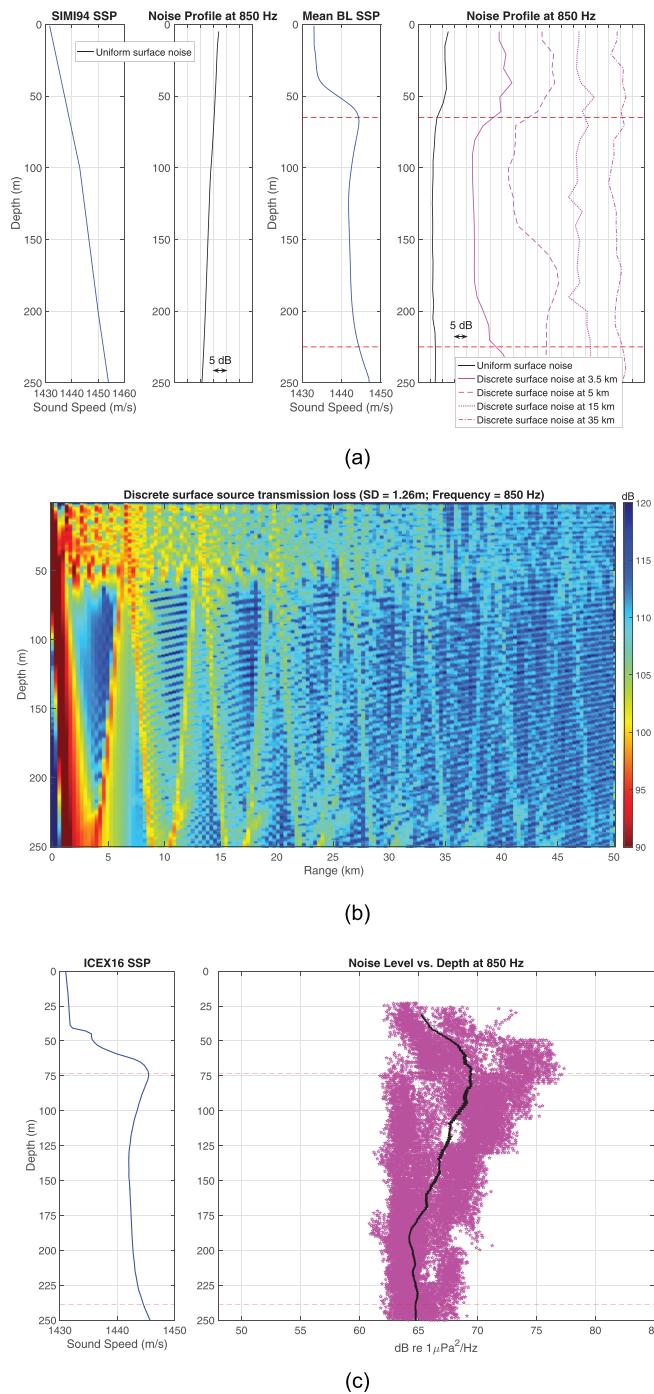


FIG. 4. (Color online) (a) Modeled ambient noise level vs depth with SIMI94 SSP and uniform surface noise (left); BL SSP and uniform surface noise (right, left-most profile); BL SSP and discrete surface noise (right, all other profiles). The distance between the x axis tick marks on the profile plots denotes a 5 dB change in noise level. (b) Transmission loss of a discrete near surface source with BL SSP. (c) Noise level vs depth of data collected during ICEX16 on March 14. Dots represent calculated noise level for individual data windows; solid line represents moving average over every 750 windows.

zone propagation is weaker as well due to attenuation, which is why the decrease in noise level compared to the surface is less dramatic. Consequently, the noise level profiles at these further ranges become more similar to the profile generated with uniform surface noise.

Figure 4(c) shows the noise level vs depth calculated with data collected during ICEX16 on March 14. As noted previously, the recording VLA center moved down to 238 m depth at 25 m increments during this recording session. Data from all 32 array elements are shown here. Each dot represents the calculated power spectral density (PSD) at 850 Hz of a 15 s data window. For each window, PSD is calculated in segments of 2048 points with 50% overlap, then averaged over time. The solid line represents the moving average over every 750 data windows. The noise profile here exhibits the effect of the BL SSP. The noise level of the collected data peaks just above the sound speed local maximum, then decreases within the lower duct. However, it is difficult to say whether this noise profile resembles more closely the BL SSP modeled profiles with uniform surface noise or discrete surface noise. This is partly because the modeled profiles in both cases are very similar. In addition, the estimated surface noise range during the time of ICEX16 data collection is ~ 40 – 50 km.¹⁹ As noted previously, at these further ranges, the range dependency of the discrete source noise profiles is much weaker compared to closer ranges and the profiles more closely resemble the uniform source distribution noise profile. Thus, in regards to underwater ambient noise level during ICEX16, the BL SSP appears to be a much more influential factor than the surface noise distribution in the ice cover.

B. Ambient noise vertical directionality

Since the ICEX16 ambient noise data were recorded with a VLA, we can use beamforming to characterize the vertical directionality of the noise during the experiment. Following Chen *et al.*, 2019,¹⁹ data from three depths are presented here: 38 m (in the surface duct), 138 m (in the lower duct), and 238 m (just below the lower duct). Figure 5 shows the noise vertical directionality profiles for the three depths. They are generated by applying conventional beamforming to the collected data between 800 and 900 Hz and then averaging the output over time and frequency.¹⁹ The

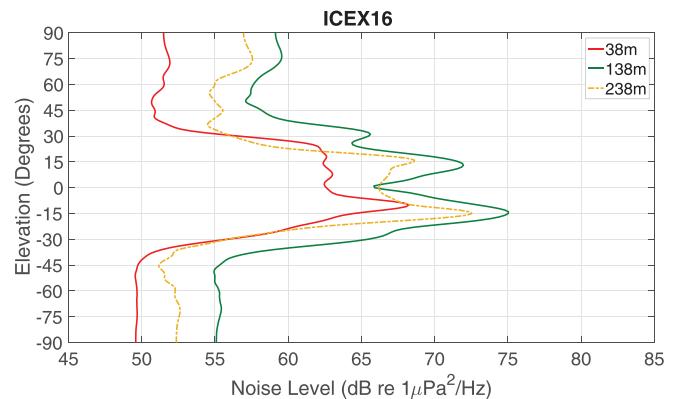


FIG. 5. (Color online) Noise vertical directionality profiles for ICEX16 data collected at 38, 138, and 238 m depths between 800 and 900 Hz. All profiles have a peak around -10 to -15 degrees and the profiles for the deeper depths show a noise notch near horizontal elevation angles. Figure from Chen *et al.* (Ref. 19).

beamforming frequency range is chosen to provide the best spatial resolution without encountering aliasing. The resultant directionality profiles show two distinct features. One is that there is a noise notch near the horizontal elevation angles. The notch is not as evident at 38 m depth but is pronounced at deeper depths. The second feature is that the peak in the profiles occurs from around -10 to -15 degrees. In this section, we will take a closer look at how the BL SSP and the discrete surface noise model in the ice cover can explain these features.

Using raytracing, we explain why horizontal noise notches are observed at 138 and 238 m depths but not so much at 38 m depth. For our raytracing environment, we use the ICEX16 SSP. First, we place a monopole source at 38 m depth. The raytraces for rays with departure angles between ± 7 degrees are shown in Fig. 6(a). Green lines represent rays that interact with the surface, while red lines represent rays that do not interact with surface. We see that with the source at 38 m, all shallow angle rays interact with the surface. By reciprocity, this means that rays generated by a discrete surface noise source at any range can arrive at a receiver at 38 m depth at very shallow arrival angles. Hence, this is why we do not see a strong horizontal noise notch in the data collected at 38 m depth. With a source placed at 138 and 238 m depths, the raytraces show that there are rays with very shallow departure angles that will never interact with the surface. Again, by reciprocity, this means that fewer rays generated by a discrete surface noise source would arrive at a receiver at those depths at shallow angles. In the real environment, this would translate to a reduction in noise arriving from the horizontal, which is why we observe the noise notches in the directionality profiles for 138 and 238 m depths.

The second feature in the ICEX16 directionality profiles is the peak in noise level from around -10 to -15 degrees. This can again be explained using raytracing and by invoking reciprocity. In Fig. 6(b), we show the raytraces for a monopole source placed at 38, 138, and 238 m depths, respectively, for departure angles between -8 and -15 degrees. Based on Fig. 3, we assume that the discrete surface noise source is located between 30 and 50 km away from the VLA. With this in mind, the rays highlighted by the dashed-dotted purple lines in Fig. 6(b) are the ones that arrive at the surface between 30 and 50 km away from the VLA via a direct path. The departure angles of these highlighted rays are from -12 to -15 degrees. Through reciprocity, this result means that for a discrete surface noise source place between 30 and 50 km away from the origin, the noise level received at 0 km range at 38, 138, and 238 m depths should be the highest between -12 and -15 degrees because these angles correspond to the arrival angles of direct paths. For steeper or shallower arrival angles, the noise level would be weaker because the ray paths would experience at least one surface or bottom bounce. Thus, these raytracing results suggest that the noises directionality profiles should peak near -12 to -15 degrees, which is very similar to what we see in the ICEX16 directionality profiles.

This explanation also further reinforces the hypothesis that the noise collected during ICEX16 was predominately generated ~ 30 – 50 km away from the VLA.

IV. TRANSIENT NOISE EVENTS IN NEW BEAUFORT SEA ENVIRONMENT

Detection and characterization of transient ambient noise events provide useful information on the temporal distribution of ice cover activity and the mechanisms by which noise is generated. In Chen *et al.*, 2019,¹⁹ an amplitude based detection algorithm is applied in the time domain to the ICEX16 data time series. While this approach is a straight-forward method for detecting high amplitude transients, it does not provide much information on the frequency characteristics of the detected events, which are helpful in hypothesizing the mechanism by which the event may have been generated. In the section, we present an alternative approach that more easily allows us to categorize detected events based on their bandwidth and duration. We then apply this method to ICEX16 data collected on March 13, 2016. Because this dataset is only an ~ 8 h snapshot of the Beaufort Sea noise environment, we design our analysis approach to focus on detecting short time transient events (on the order of seconds to tens of seconds) that result from small scale ice cracking or floe interactions. These types of events commonly occur near pressure ridges where weaker ice readily fractures and breaks off into floes due to temperature, wind, or current stress.

A. Transient event detection algorithm

Our method takes an image processing approach to detect transient events. The input to the algorithm is the spectrogram of a data snapshot. Spectrograms of ICEX16 dataset snapshots are computed using a 4096-points Hanning window with 50% overlap for each hydrophone channel in the VLA. All channels are then averaged to form a composite spectrogram. We focus our analysis to below 2048 Hz and place more emphasis on lower frequencies by transforming the frequency axis to mel scale with 128 bins. The conversion from Hertz to mel scale is linear for frequencies ≤ 1 Khz and logarithmic above 1 Khz. This step serves three purposes. First, noise generated by small scale and transient ice events has most of their spectral energy below 2 kHz, with much of that focused below 1 kHz.^{18,27} The use of the mel scale conveniently places more focus on lower frequencies during event detection. Second, while the ICEX16 VLA has the best beamforming spatial resolution at ~ 900 Hz, it would not encounter any aliasing between ± 30 degrees for frequencies up to ~ 2048 Hz. Since much of the ambient noise arrives at the VLA between ± 30 degrees (Fig. 5), we can safely perform beamforming on any transient event that we detect to determine its vertical directionality. The final purpose is that the decreased number of frequency bins as the result of using the mel scale decreases the total number of pixels in the input spectrograms. This allows our algorithm to process the data more efficiently.

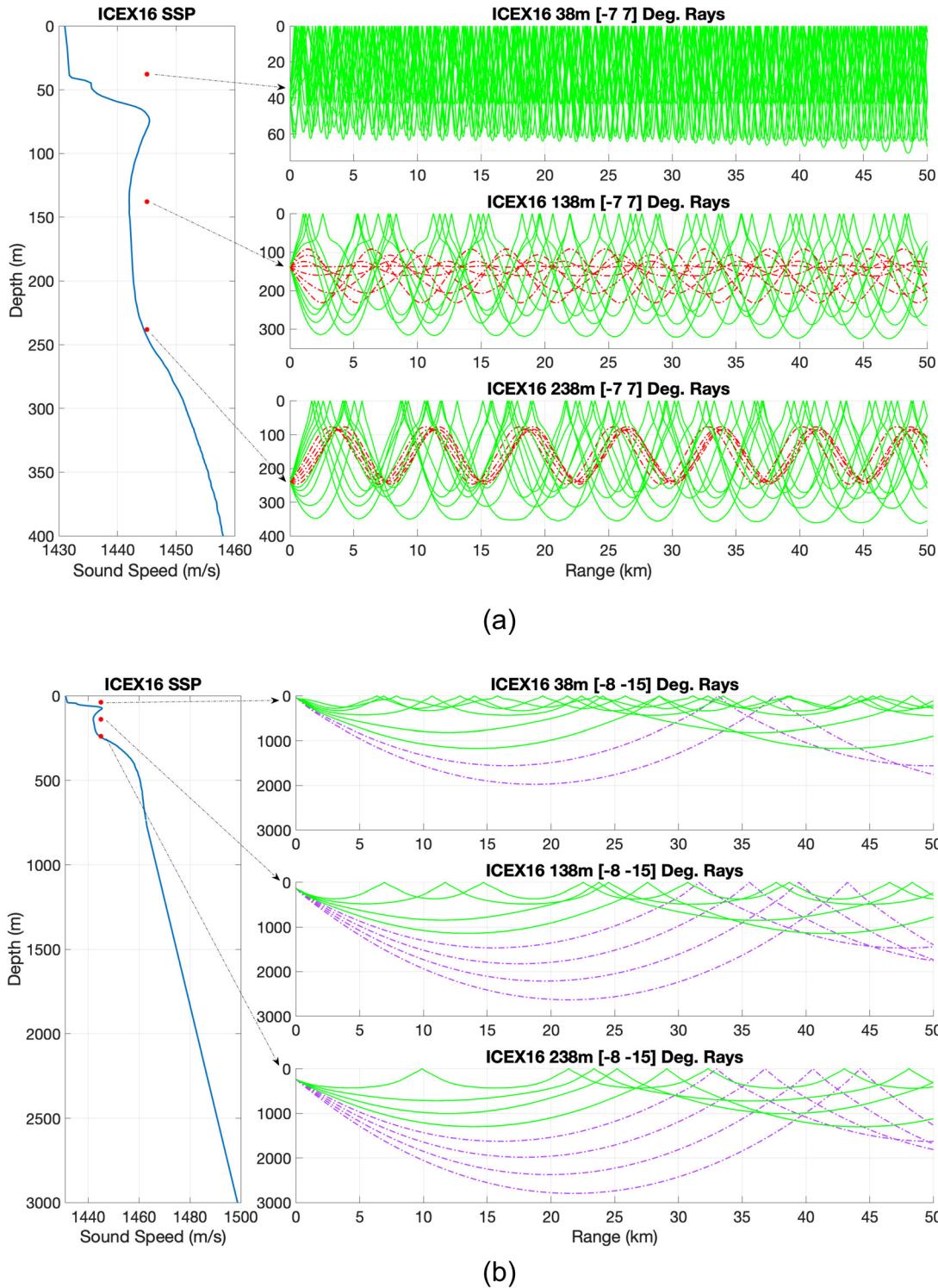


FIG. 6. (Color online) Raytraces for monopole source placed at 38, 138, and 238 m depths in the ICE16 environment. Solid lines denote rays that interact with the surface. (a) For 138 and 238 m depths, some rays with shallow departure angles would never interact with surface (dashed-dotted lines). (b) Rays that reach the surface 30–50 km away from the source via direct paths (dashed-dotted lines) have departure angles from -12 to -15 degrees.

The detection algorithm itself is regulated by three threshold parameters ($\text{Thres}_{\text{mask}}$, $\text{Thres}_{\text{dist}}$, $\text{Thres}_{\text{area}}$). $\text{Thres}_{\text{mask}}$ is a threshold that controls which pixels in the spectrogram we keep based on how many times higher their values are compared to the background noise estimate (it is a unitless

measure of linear signal-to-noise). $\text{Thres}_{\text{dist}}$ is a measure of distance between pixels based on their indices in the spectrogram. $\text{Thres}_{\text{area}}$ is a measure of the number of pixels in a feature (see definition of a feature below) in the spectrogram. Briefly, our algorithm works as follows:

- (1) The background noise level at each frequency, defined as the noise level when no transient events are occurring, is estimated for the first input data snapshot by calculating the spectrogram of a window of data much longer than the duration of the first snapshot and then averaging the spectrogram values over time. For the analysis of the ICEX16 data, the individual input snapshot lengths are set to 30 s while the duration of data used to estimate the background noise is 30 min. This approach of estimating the background noise level works off of two assumptions. One is that for any sufficiently long window of data, there are no transient events during a large majority of the time. The other is that there is enough total data available for such a sufficiently long window. In our testing with the ICEX16 dataset, a 30 min long window is long enough for us to detect even very low signal-to-noise ratio (SNR) transients in the first input snapshot. An example of an input spectrogram is shown in Fig. 7 (top left).
- (2) The frequency-normalized spectrogram (Fig. 7, top middle) is calculated for the current data snapshot window by dividing the original spectrogram values at each frequency by the estimated background noise level at the corresponding frequency. This step helps to mitigate the issue that noise level is typically higher at lower frequencies, which can mask transient events at higher frequencies if the spectrogram is not normalized.
- (3) Image processing techniques such as smoothing, dilation, and erosion are applied to the frequency-normalized spectrogram to filter out noise in the image that may be mistaken for transient events. In image analysis, erosion thins the boundary of the foreground object and gets rid of very small features, or “salt” noise. Contrarily, dilation expands the boundary of the foreground object and eliminates small gaps between

objects, or “pepper” noise. When used together, they help diminish the amount of “salt and pepper” noise. Smoothing is done by applying a 3×3 Gaussian blurring filter over the spectrogram. Dilation and erosion are performed using the Python library OpenCV, also with a 3×3 kernel. After noise reduction, the absolute gradients along the time and frequency axes of the filtered spectrogram are determined and combined to construct a gradient map (Fig. 7, top right).

- (4) A mask is created by thresholding the gradient map—all pixels values in the map that are less than $Thres_{mask}$ are set to 0 while the rest are set to 1 (Fig. 7, bottom left). The mask is then applied to the filtered spectrogram through multiplication and all zero-value pixels are removed from the output. This results in a post-processed image with disconnected pixel groups that we call features (Fig. 7, bottom middle).
- (5) The features in the post-processed image are further grouped together using hierarchical clustering (h-clustering) with a custom distance metric that is designed to conjoin separate features that are likely to be part of a single transient event (see next section). Two features are conjoined if the distance between them is less than $Thres_{dist}$. At the end of the clustering process, very small features with less than $Thres_{area}$ number of pixels are deleted to further eliminate noise. The detected transients are then labeled by type based on their bandwidth and duration (Fig. 7, bottom right, see below for type definitions).
- (6) After transient detection within the current snapshot, background noise is estimated for the next snapshot. Pixels in the detected transients are deleted from the current input spectrogram to create a “noise-only” image. An update to the background noise level is calculated by averaging this “noise-only” image over time. The new

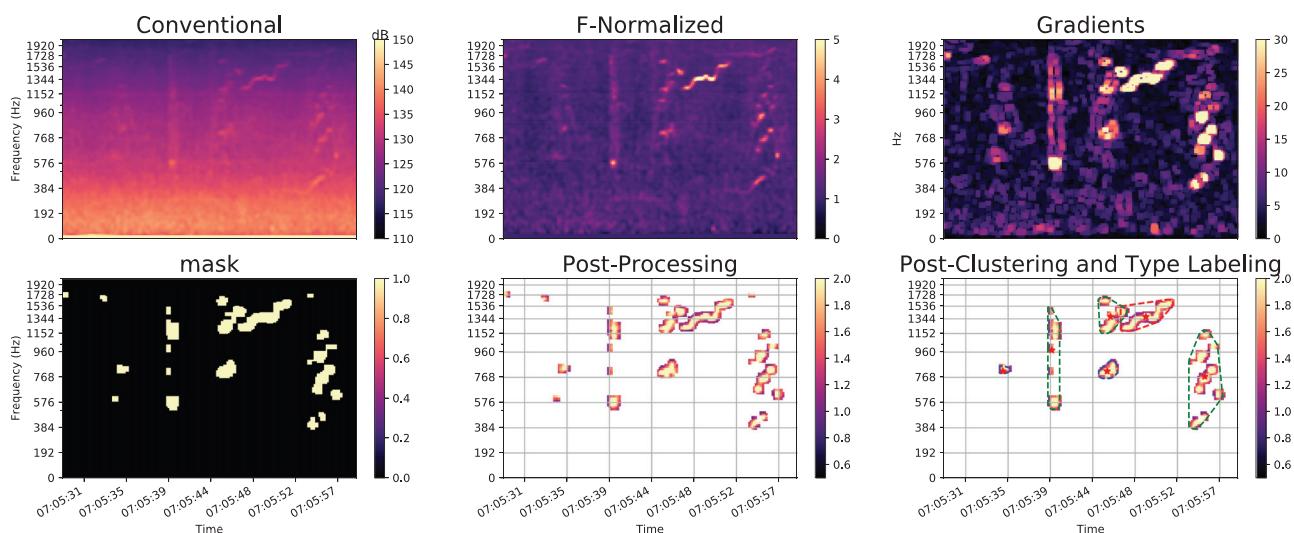


FIG. 7. (Color online) Example of an ICEX16 data snapshot spectrogram that is the input into the detection algorithm (top left). Input spectrogram after frequency normalization (top middle). Gradient map created following image processing to reduce “salt and pepper” noise (top right). Mask created based on the gradient map (bottom left). Post-processed spectrogram after applying the mask; only pixels in detected transients are kept (bottom middle). Outline of detected transients after h-clustering and labeling. Blue, stnb; red, ltnb; green, stbb; black, ltbb. Red star denotes the center of the detected event (bottom right).

background noise level is then calculated by averaging together the previous background noise level and this new update.

- (7) Repeat steps 2–6 on the spectrogram of the next data snapshot window.

For analysis of the ICEX16 data, the three threshold parameters are set as $Thres_{mask} = 17$, $Thres_{dist} = 5.5$, $Thres_{area} = 20$. These values are selected empirically after some testing on a small subset of the data to check the algorithm is detecting all transients that are observed by a human

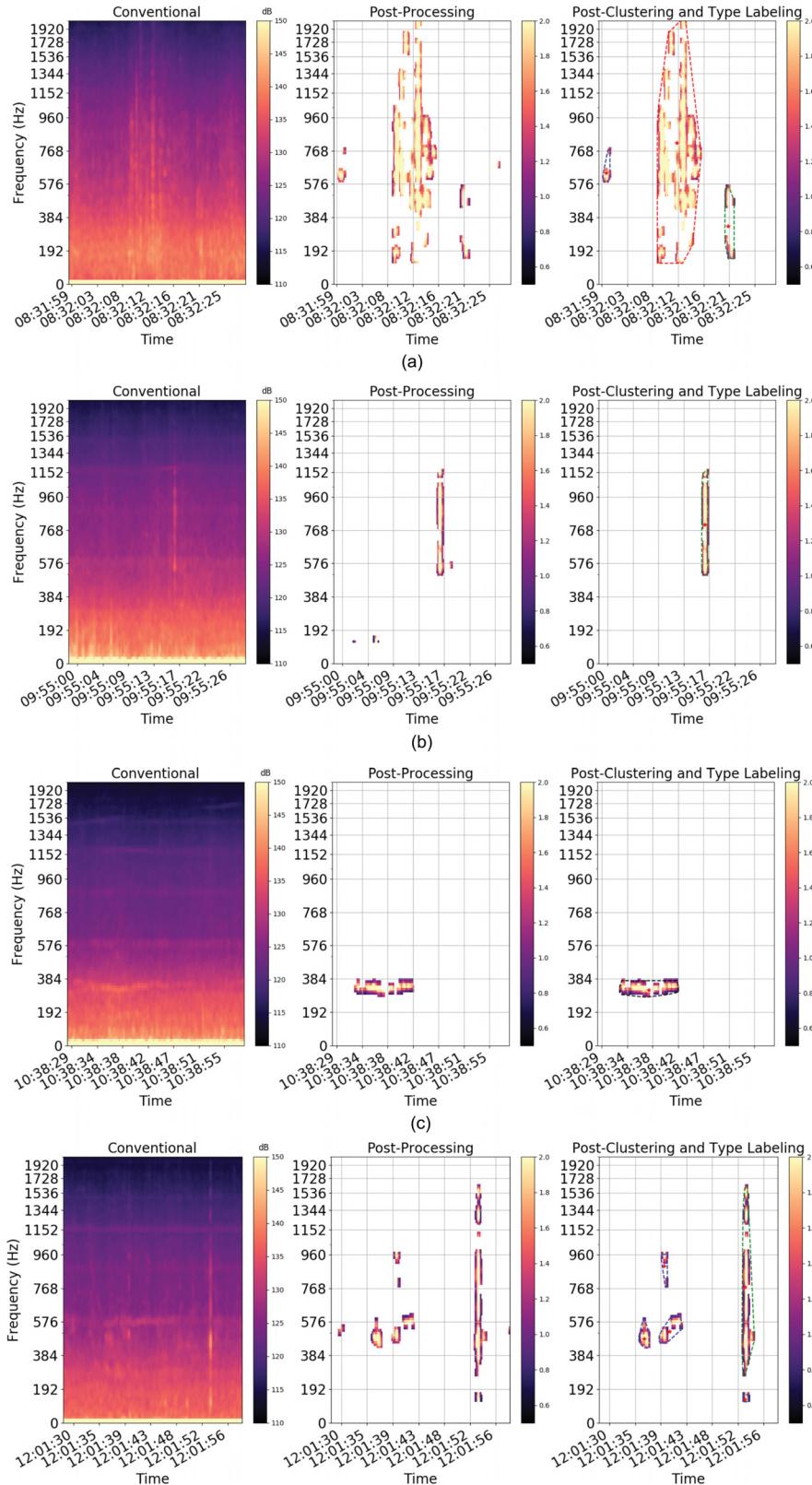


FIG. 8. (Color online) (a) A ltbb event followed by a stbb event. (b) A stbb event. (c) A ltnb event. (d) Some stnb noise followed by a stbb event.

analyst. More details on the performance of this algorithm are shown in the [Appendix](#), where we apply it to a dataset with annotated transient events that serve as ground truth comparison to the algorithm's outputs.

Based on the bandwidth and duration of the detected transients, we define four types of events: short-time-narrowband (stnb, duration ≤ 5 s, bandwidth ≤ 200 Hz), long-time-narrowband (ltnb, duration > 5 s, bandwidth ≤ 200 Hz), short-time-broadband (stbb, duration ≤ 5 s, bandwidth > 200 Hz), and long-time-broadband (ltbb, duration > 5 s, bandwidth > 200 Hz). Examples of each event type are shown in Fig. 8. We believe the stnb events mostly encompass high SNR background noise within the spectrogram that have been falsely detected as events or perhaps small features that may have been part of another event but were incorrectly clustered. Thus, the stnb category serves as another noise removal procedure. The representation of the other three categories is discussed in Sec. [IV C](#).

B. Hierarchical clustering and distance metric design

Hierarchical clustering (h-clustering) is used in our event detection method to group together features that may be part of the same transient event. We take a “bottom-up” approach, which means that each feature in the post-processed image starts off as an individual cluster of one. Then, starting with the largest cluster by the number of pixels, we iterate through each cluster and combine that cluster with its closest neighbor as defined by a distance metric. We stop the clustering process once all remaining clusters are at least $Thres_{dist}$ apart from each other.

The distance metric used in our h-clustering procedure is designed based on two principles. The first is that, intuitively, features that occur at or near the same time are more likely to be part of the same transient event. To implement this principle, we elect to use an average of the absolute pixel distance and the pixel distance along the time axis only as the distance metric between two features. This metric represents a balance between grouping together features that are close together in terms of relative proximity, while also combining more distant features that are close in time only; some of the more distant features may be part of a single broadband event and would not otherwise be grouped.

The second design principle places a constraint on the first and holds that it is unlikely for a horizontal and a vertical feature to be part of the same transient event. This is because a horizontal feature is more likely to be part of a narrowband tonal while a vertical feature is more likely to be part of a broadband impulse. Thus, as part our distance metric calculation, we first determine the slopes of the features. If the slope is greater than 1.2, we classify that feature as vertical. If the slope is less than 0.8, we classify the feature as horizontal. Otherwise, we do not assign an orientation label. Note, a slope of 1 means a feature has an equal number of pixels in height and width. In calculating the distance between two features, if one has vertical orientation while the other has horizontal orientation, we manually

assign their distance to be infinite so that they are never clustered together. This ensures that tonals and broadband events are not grouped as a single event, even if they occur closely in time.

C. Characteristics of detected transient noise events

A total of 3178 transient events are detected by applying our algorithm to the ~ 8 h of ICEX16 data collected on March 13. Of these, 2081 are stnb events, which, as we noted, are likely background noise that are not part of ice generated transients. Thus, we do not consider these in our discussion and instead focus on the other three event categories. Details of the 1097 non-stnb events are shown in Table [II](#). stbb events are the most common event type, with 773 occurrences accounting for 70.5% of the detections. We believe this event type is indicative of small, impulsive ice fractures that may occur in the periphery of larger cracking events. As described by Xie and Farmer,¹⁶ ice breaking can be partitioned into three phases. First, environmental forcing initiates ice breaking near weakened locations such as ridge formations. This process emits short, impulsive, sound pulses, similar to the stbb events that we observe. Second, the initial fractures promote more failures and further cracking. In turn, a positive feedback loop develops and leads to a more extended period of broadband noise generation. We believe this second phase is descriptive of the ltbb events. They are the most prevalent event type with 2366.4 s of total combined duration over 242 detections. This accounts for $\sim 8.3\%$ of the total data recording time. The third phase of ice breaking accounts for ice floe interactions. As small ice floes break off and rub past each other, pure tones are emitted into the water column. We believe that the ltnb events are representative of these tonal features. They are the rarest event type from our analysis. Only 82 are detected and their combined duration represents less than $\sim 2.5\%$ of the total data recording time. The distribution of event durations for the detected events is shown in Fig. 9. From this plot, we see that the majority of stbb events have durations less than 3 s. For ltnb and ltbb events, their durations are longer than 5 s by definition but the majority lasts no longer than 10 s. These relatively shorter event durations compared to findings from previous studies^{14,15} suggest that ice cracking during ICEX16 was likely occurring on a small temporal and spatial scale, absent of any prolonged major activity in the ice cover (e.g., no large lead openings).

After event detection, the peak arrival angle in vertical directionality at the VLA is determined for each transient. This step is done by segmenting the portion of data that

TABLE II. Detected transient events by type.

Event type	# of Detections	% of Total	Combined Duration (s)	% of recording Time	Mean peak Arrival angle (degrees)
stbb	773	70.5	1800.5	6.3	-10.7
ltnb	82	7.5	691.9	2.4	-11.0
ltbb	242	22.0	2366.4	8.3	-10.5

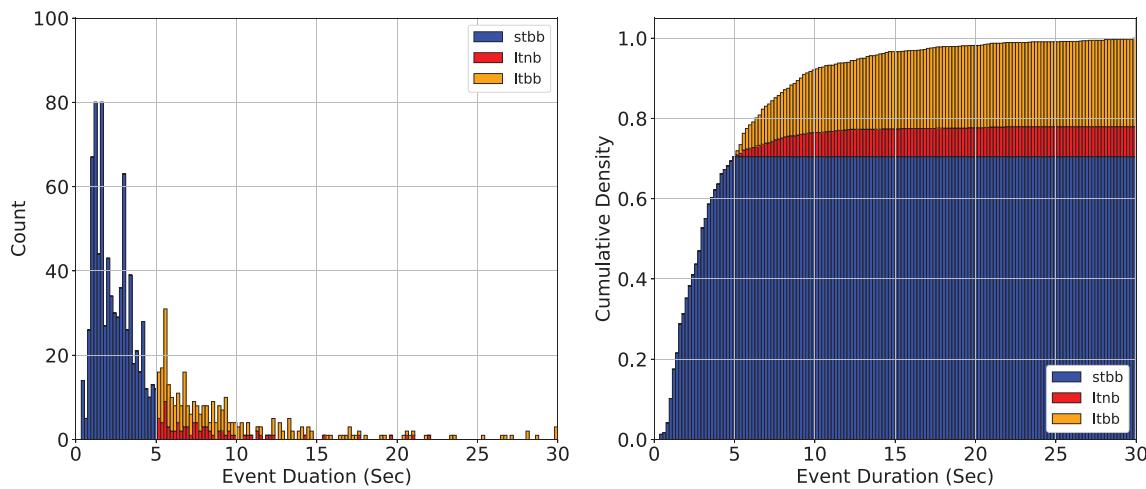


FIG. 9. (Color online) Distribution of event durations by type.

contains the event using a 512-point Hanning window and then beamforming over the frequency range of the event. The beamformed output is then averaged over frequency and time to derive its directionality profile. From there, the peak arrival angle is determined. As shown in Fig. 10, these peak arrival angles mostly cluster near -10 degrees and are fairly consistent over time for all event types. The mean peak arrival angles for different event types are all between -10 and -11 degrees (Table III). These results agree with that of the overall ambient noise. Thus, in accordance with our previous discussion in Sec. III, the peak arrival angle values support that the detected events are generated by the ridge formation $\sim 30\text{--}50$ km away from the ICEX16 camp (Fig. 3).

Table III shows the frequency content of the detected events. Five frequency intervals are defined at octave ranges (<160 , $160\text{--}320$, $320\text{--}640$, $640\text{--}1280$, and $1280\text{--}2048$ Hz). Of these intervals, $640\text{--}1280$ Hz and $320\text{--}640$ Hz are the most active with 79.8% and 70.9% of events having some frequency content within these two octaves, respectively. In addition, the bandwidth distribution of the events is shown in Fig. 11. For ltnb events, most have a bandwidth of less

than 150 Hz. For broadband events of both types, most have bandwidths of less than 750 Hz. Of course, the event bandwidths are capped by our choice for the analysis frequency limit of 2048 Hz. However, the values in Fig. 11 are reasonable compared with transient events presented in previous studies.^{14–16}

The final noise characteristic that we present is the time gap between the starts of consecutive transients. Figure 12 (left) shows that the majority of events occurred less than 50 s from the previous. This result suggests that transient events were occurring quite uniformly throughout the data recording session. If we examine the time between consecutive events of the same type, Fig. 12 (right) shows that the time gaps can be much longer, especially for ltnb and ltbb events. This makes intuitive sense as these events are rarer than stbb events. However, their rarity does not mean that they can not occur in quick succession of one another. The majority of the ltnb and ltbb events nonetheless occurred within 100 s of the previous event of their respective type. Figure 10 shows the temporal distribution of detected events by type. For ltnb and ltbb events, we observe that there are times when they happen in clusters (e.g., near 08:24:00) and times when there is a longer gap in their occurrence (e.g., after 13:12:00). In contrast, the occurrence of stbb event appears to be more consistent over time. These observations suggest that small fractures may have occurred fairly regularly during ICEX16, whereas larger cracking events and the subsequent ice floe interactions were more sparse in time. However, when larger cracking events did occur, they tend to have happened in clusters, resulting in a sudden increase in transient noise generation.

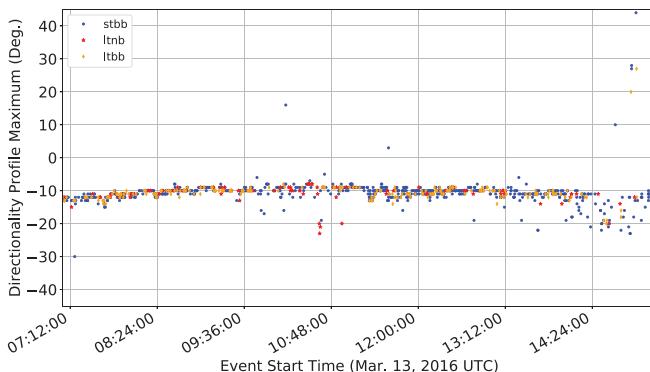


FIG. 10. Peak arrival angles of detected events most cluster near -10 degrees, in agreement with that of the overall ambient noise. This plot also shows the temporal distribution of detected events by type.

TABLE III. Percentage of events that are at least partly within a frequency interval (excludes stnb events).

Frequency interval (Hz)	<160	$160\text{--}320$	$320\text{--}640$	$640\text{--}1280$	$1280\text{--}2048$
% of events	8.8	22.3	70.9	79.8	19.3

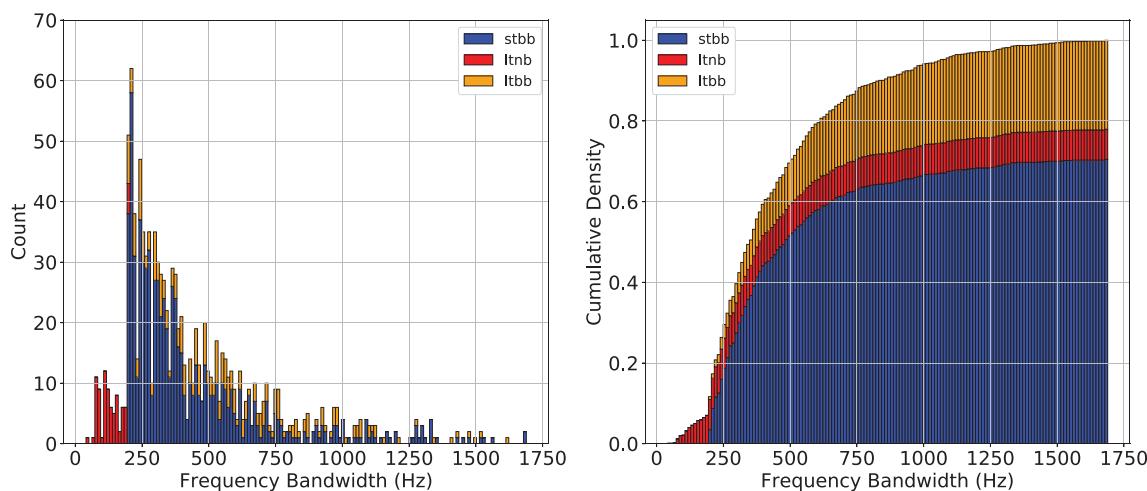


FIG. 11. (Color online) Distribution of event bandwidths by type.

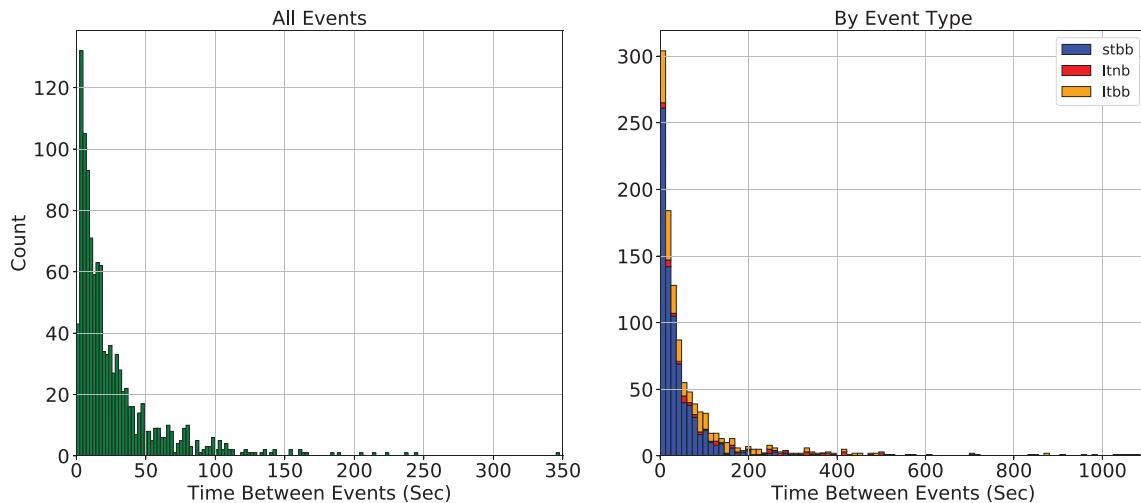


FIG. 12. (Color online) Distribution of time gap between event start times (left) for all events and (right) for events of the same type.

V. CONCLUSIONS

Using data collected during ICEX16, we demonstrate some effects on underwater ambient noise caused by environmental changes in the Beaufort Sea region of the Arctic Ocean. The emergence of the BL SSP is responsible for a decrease in ambient noise level within the lower propagation duct between ~ 75 and 250 m depth. This SSP, paired with localized generation of ambient noise along a pressure ridge ~ 30 – 50 km away from the ICEX16 camp, suppresses noise arrival near horizontal elevations while promoting arrival between -10 and -15 degrees. With a spectrogram based transient event detection algorithm, we characterize transient events that occurred during ICEX16 based on their duration and bandwidth. Our results suggest that ice cracking during this experiment likely happened on small temporal and spatial scales. Small fractures occurred regularly, whereas larger cracking events and ice floe interactions were more rare.

For future work, the BL SSP should continue to be monitored so we can better understand its variability and

how it affects the noise environment. In addition, less empirical metrics for event categorization can be developed for the transient event detection algorithm used in this study. For example, the slope or orientation of the detected events may be used instead of arbitrary boundaries in duration and bandwidth. This would allow for less biased classification of detected events and further increase detection accuracy and utility.

ACKNOWLEDGMENTS

This paper is part of a special issue on Ocean Acoustics in the Changing Arctic. The authors appreciate the enormous effort by the entire team involved in the ICEX16 deployment that made this research possible. Particular thanks go to LT Scott Carper and Thomas Howe who worked tirelessly in extreme conditions deploying, operating, and recovering the equipment. The support from the Arctic Submarine Laboratory and the logistics team from APL-UW is highly appreciated as well. The research was supported by the National Science Foundation Graduate

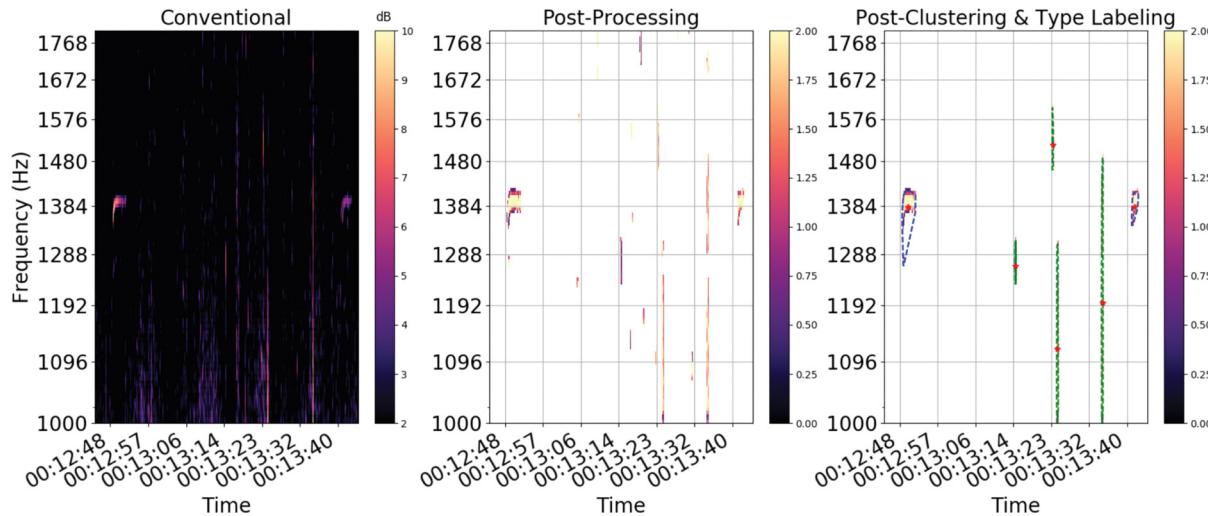


FIG. 13. (Color online) Examples of whale call detections (near 00:12:48 and 00:13:40) and impulsive noise (all others) in HICEAS data. The x axis in the plots denote UTC time on December 1, 2017.

Research Fellowship under Grant No. 2389237 and by the Office of Naval Research Code 322OA under Grant No. N00014-17-1-2474. Additional funding for the data processing and analysis was provided by DARPA STO under the FAST Program. We also thank the editor and reviewers, whose suggestions further improved this work.

APPENDIX: EVENT DETECTION ALGORITHM PERFORMANCE ON HAWAIIAN ISLANDS CETACEAN AND ECOSYSTEM ASSESSMENT SURVEY (HICEAS) DATASET

We present more information on the performance of the event detection algorithm used in this study by applying it to a dataset with ground truth transient event labels.

This dataset is a subset of whale acoustic data collected during the HICEAS in 2017.²⁸ It was collected using a multi-channel towed hydrophone array²⁹ with a sampling frequency of 500 kHz. For our analysis, we focus on detecting Minke whale “boing” vocalizations within a ~ 3.5 h segment of data collected on December 1, 2017. The ground truth labels for these calls were annotated using the software program PAMGuard between 1100 and 1800 Hz.²⁸ To generate the input spectrograms, a Hanning window of length 51 200 is applied to 1 min-long data snapshots with 50% overlap. The initial background noise estimate is calculated using 10 min of data. The application of our algorithm is the same as described in Sec. IV but with some parameters changed to suit this dataset. Most notably, we set $Thres_{mask} = 50$, $Thres_{dist} = 5.5$, $Thres_{area} = 75$. These values are again empirically chosen after some testing on a subset of the dataset. The major cause of false predictions in this dataset is the presence of short, impulsive noise (Fig. 13). Thus, we devise two event categories: short-time (duration < 1.25 s) and long-time (duration ≥ 1.25 s). Only events classified as long-time are considered whale call detections. With this set-up, our algorithm achieves a true positive rate of 70.9% and a false positive rate of 13.9% when compared to the ground truth annotations.

¹J. M. Toole, M. L. Timmermans, D. K. Perovich, R. A. Krishfield, A. Proshutinsky, and J. A. Richter-Menge, “Influences of the ocean surface mixed layer and thermohaline stratification on arctic sea ice in the central Canada basin,” *J. Geophys. Res.* **115**, C10018, <https://doi.org/10.1029/2009JC005660> (2010).

²S. Carper, “Low frequency active sonar performance in the arctic beaufort lens,” Master’s thesis, MIT, Cambridge, MA (2017).

³L. Freitag, K. Ball, P. Koski, and S. Singh, “Long range acoustic communications and navigation in the Arctic,” in *Proceedings of OCEANS 2015—MTS/IEEE*, Washington, DC (May 18–21, 2015), pp. 1–5.

⁴T. Howe, “Modal analysis of acoustic propagation in the changing arctic environment,” Master’s thesis, MIT, Cambridge, MA (2015).

⁵H. Schmidt and T. Schneider, “Acoustic communication and navigation in the new arctic: A model case for environmental adaptation,” in *Proceedings of the 2016 IEEE Third Underwater Communications and Networking Conference (UComms)*, Lerici, Italy (August 30–September 1, 2016), pp. 1–4.

⁶T. F. Duda, “Acoustic signal and noise changes in the beaufort sea pacific water duct under anticipated future acidification of arctic ocean waters,” *J. Acoust. Soc. Am.* **142**, 1926–1933 (2017).

⁷K. E. Frey, G. W. K. Moore, L. W. Cooper, and J. M. Grebmeier, “Divergent patterns of recent sea ice cover across the bering, chukchi, and beaufort seas of the pacific arctic region,” *Prog. Oceanogr.* **136**, 32–49 (2015).

⁸R. Kwok and D. A. Rothrock, “Decline in Arctic sea ice thickness from submarine and icesat records: 1958–2008,” *Geophys. Res. Lett.* **36**, L15501, <https://doi.org/10.1029/2009GL039035> (2009).

⁹R. Lindsay and A. Schweiger, “Arctic sea ice thickness loss determined using subsurface, aircraft, and satellite observations,” *Cryosphere* **9**, 269–283 (2015).

¹⁰J. Maslanik, J. Stroeve, C. Fowler, and W. Emery, “Distribution and trends in arctic sea ice age through spring 2011,” *Geophys. Res. Lett.* **38**, L13502, <https://doi.org/10.1029/2011GL047735> (2011).

¹¹A. R. Milne and J. H. Ganton, “Ambient noise under arctic sea ice,” *J. Acoust. Soc. Am.* **36**, 855–863 (1964).

¹²A. R. Milne, “Statistical description of noise under shore-fast sea ice in winter,” *J. Acoust. Soc. Am.* **39**, 1174–1182 (1966).

¹³P. Zakarauskas and J. M. Thorleifson, “Directionality of ice cracking events,” *J. Acoust. Soc. Am.* **89**(2), 722–734 (1991).

¹⁴G. B. Kinda, Y. Simard, C. Gervaise, J. I. Mars, and L. Fortier, “Arctic underwater noise transients from sea ice deformation: Characteristics, annual time series, and forcing in beaufort sea,” *J. Acoust. Soc. Am.* **138**(4), 2034–2045 (2015).

¹⁵E. Ozanich, P. Gerstoft, P. F. Worcester, M. A. Dzieciuch, and A. Thode, “Eastern arctic ambient noise on a drifting vertical array,” *J. Acoust. Soc. Am.* **142**(4), 1997–2006 (2017).

¹⁶Y. Xie and D. M. Farmer, “The sound of ice break-up and floe interaction,” *J. Acoust. Soc. Am.* **91**(3), 1423–1428 (1992).

- ¹⁷Z. Ye, "Sound generation by ice floe rubbing," *J. Acoust. Soc. Am.* **97**, 2191–2198 (1995).
- ¹⁸W. C. Cummings, O. I. Diachok, and J. D. Shaffer, "Acoustic transients of the marginal sea ice zone: A provisional catalog," Technical Report No. NRL-MR-6408 (Naval Research Lab Washington, DC, 1989).
- ¹⁹R. Chen, A. Poulsen, and H. Schmidt, "Spectral, spatial, and temporal characteristics of underwater ambient noise in the beaufort sea in 1994 and 2016," *J. Acoust. Soc. Am.* **145**, 605–614 (2019).
- ²⁰Woods Hole Oceanographic Institution, "Ice-tethered profiler program," <http://www.whoi.edu/itp> (Last viewed: December 17, 2020).
- ²¹M. Tschudi, G. Riggs, D. K. Hall, and M. O. Román, "VIIRS/NPP ice surface temperature 6-min 12 swath 750 m, version 1," NASA National Snow and Ice Data Center Distributed Active Archive Center, <https://doi.org/10.5067/VIIRS/VNP30.001> (Last viewed July 24, 2020).
- ²²W. A. Kuperman and F. Ingenito, "Spatial correlation of surface generated noise in a stratified ocean," *J. Acoust. Soc. Am.* **67**(6), 1988–1996 (1980).
- ²³H. Schmidt and W. A. Kuperman, "Estimation of surface noise source level from low frequency seismoacoustic ambient noise measurements," *J. Acoust. Soc. Am.* **84**, 2153–2162 (1988).
- ²⁴H. Schmidt, "OASES: Ocean acoustic and seismic exploration synthesis," <http://lamss.mit.edu/lamss/pmwiki/pmwiki.php?n=Site.Oases> (Last viewed July 24, 2020).
- ²⁵G. Hope, H. Sagen, E. Storheim, H. Hobæk, and L. Freitag, "Measured and modeled acoustic propagation underneath the rough Arctic sea-ice," *J. Acoust. Soc. Am.* **142**, 1619–1633 (2017).
- ²⁶F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics* (Springer, New York, 2011).
- ²⁷K. L. Williams, M. L. Boyd, A. G. Soloway, E. I. Thorsos, S. G. Kargl, and R. I. Odom "Noise background levels and noise event tracking/characterization under the Arctic ice pack: Experiment, data analysis, and modeling," *IEEE J. Ocean. Eng.* **43**, 145–159 (2018).
- ²⁸National Oceanic and Atmospheric Administration, "HICEAS Dataset" (2017), <http://www.soest.hawaii.edu/ore/dclde/dataset/> (Last viewed August 8, 2020).
- ²⁹S. Rankin, J. Barlow, Y. Barkley, and R. Valtierra, "A guide to constructing hydrophone arrays for passive acoustic data collection during NMFS shipboard cetacean surveys" U.S. Department of Commerce, NOAA Technical Memorandum NOAA-TM-NMFS-SWFSC-511 (Washington, DC, 2013).