

Classification of broadband target spectra in the mesopelagic using physics-informed machine learning

Emma Cotter, Christopher Bassett, and Andone Lavery

Citation: *The Journal of the Acoustical Society of America* **149**, 3889 (2021); doi: 10.1121/10.0005114

View online: <https://doi.org/10.1121/10.0005114>

View Table of Contents: <https://asa.scitation.org/toc/jas/149/6>

Published by the *Acoustical Society of America*

ARTICLES YOU MAY BE INTERESTED IN

[Seabed type and source parameters predictions using ship spectrograms in convolutional neural networks](#)

The Journal of the Acoustical Society of America **149**, 1198 (2021); <https://doi.org/10.1121/10.0003502>

[Comparison of mesopelagic organism abundance estimates using in situ target strength measurements and echo-counting techniques](#)

JASA Express Letters **1**, 040801 (2021); <https://doi.org/10.1121/10.0003940>

[Frequency- and depth-dependent target strength measurements of individual mesopelagic scatterers](#)

The Journal of the Acoustical Society of America **148**, EL153 (2020); <https://doi.org/10.1121/10.0001745>

[Estimating target strength and physical characteristics of gas-bearing mesopelagic fish from wideband in situ echoes using a viscous-elastic scattering model](#)

The Journal of the Acoustical Society of America **149**, 673 (2021); <https://doi.org/10.1121/10.0003341>

[Machine learning in acoustics: Theory and applications](#)

The Journal of the Acoustical Society of America **146**, 3590 (2019); <https://doi.org/10.1121/1.5133944>

[Nonlinear crosstalk in broadband multi-channel echosounders](#)

The Journal of the Acoustical Society of America **149**, 87 (2021); <https://doi.org/10.1121/10.0002943>



**Advance your science and career
as a member of the**

ACOUSTICAL SOCIETY OF AMERICA

LEARN MORE



Classification of broadband target spectra in the mesopelagic using physics-informed machine learning^{a)}

Emma Cotter,^{1,b)} Christopher Bassett,^{2,c)} and Andone Lavery^{1,d)}

¹Department of Applied Ocean Physics and Engineering, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA

²Applied Physics Laboratory, University of Washington, Seattle, Washington 98195, USA

ABSTRACT:

Broadband echosounders measure the scattering response of an organism over a range of frequencies. When compared with acoustic scattering models, this response can provide insight into the type of organism measured. Here, we train the k-Nearest Neighbors algorithm using scattering models and use it to group target spectra (25–40 kHz) measured in the mesopelagic near the New England continental shelf break. Compared to an unsupervised approach, this creates groupings defined by their scattering physics and does not require significant tuning. The model classifies human-annotated target spectra as gas-bearing organisms (at, below, or above resonance) or fluid-like organisms with a weighted *F1*-score of 0.90. Class-specific *F1*-scores varied—the *F1*-score exceeded 0.89 for all gas-bearing organisms, while fluid-like organisms were classified with an *F1*-score of 0.73. Analysis of classified target spectra provides insight into the size and distribution of organisms in the mesopelagic and allows for the assessment of assumptions used to calculate organism abundance. Organisms with resonance peaks between 25 and 40 kHz account for 43% of detections, but a disproportionately high fraction of volume backscatter. Results suggest gas bearing organisms account for 98.9% of volume backscattering concurrently measured using a 38 kHz shipboard echosounder between 200 and 800 m depth. © 2021 Acoustical Society of America.

<https://doi.org/10.1121/10.0005114>

(Received 6 November 2020; revised 7 May 2021; accepted 10 May 2021; published online 4 June 2021)

[Editor: Zoi-Heleni Michalopoulou]

Pages: 3889–3901

I. INTRODUCTION

Broadband echosounders are increasingly prevalent tools for fisheries research (Bassett *et al.*, 2018; Blanluet *et al.*, 2019; Demer *et al.*, 2017; Lavery *et al.*, 2017; Verma *et al.*, 2017) because they can measure scattered sound over a continuous band of frequencies, in contrast to narrowband echosounders that operate at a single frequency. In combination with physics-based acoustic scattering models, the frequency-dependent scattering response of an organism can provide insight into the type, and sometimes size, of an organism that is observed (Bassett *et al.*, 2020; Kubilius *et al.*, 2020; Reeder *et al.*, 2004; Stanton *et al.*, 2010). In this work, we analyze broadband target spectra of organisms in the ocean mesopelagic zone (approximately 200–1000 m deep), with the aim of better understanding their distribution and contribution to volume backscatter.

Broadband echosounders are particularly well suited to study the mesopelagic, which forms one of the largest habitats on earth and is home to a diverse community of organisms. However, there is significant uncertainty surrounding the abundance and distribution of organisms and biodiversity in this region of the ocean (Kaartvedt *et al.*, 2012;

Proud *et al.*, 2019). Volume backscatter measured by shipboard narrowband echosounders operating at 38 kHz, a frequency commonly used in fisheries acoustics to study epipelagic species (Simmonds and MacLennan, 2005), has recently been used to estimate global mesopelagic biomass (Irgoien *et al.*, 2014; Proud *et al.*, 2019). Interpretation of these measurements is complicated by the fact that 38 kHz is near the resonance frequency of many gas bladder-bearing mesopelagic fishes, where target strength (TS) is highly sensitive to frequency (Khodabandeloo *et al.*, 2021; Stanton *et al.*, 2010). Traditional narrowband acoustic-trawl survey techniques that rely on TS-weight relationships to estimate biomass are not effective when the scattering responses of organisms are not flat near the measurement frequency (Simmonds and MacLennan, 2005). Further, the resonance frequencies of organisms with gas inclusions (e.g., fish with swimbladders) change with ambient pressure, meaning that the TS of the same organism will vary more as a function of depth and physiological responses to changes in pressure. Multi-frequency measurements (i.e., multiple narrowband echosounders) can provide some insight into these complicating factors (Kloser *et al.*, 2016), but typically do not resolve the resonance frequency of a target nor have sufficient resolution to reliably estimate the resonance frequency and therefore estimate its impact on measured TS.

Our recent work analyzed target spectra from mesopelagic organisms measured *in situ* with broadband echosounders (20–40 and 50–90 kHz) (Bassett *et al.*, 2020).

^{a)}Note: This paper is part of a special issue on Machine Learning in Acoustics.

^{b)}Electronic mail: ecotter@whoi.edu, ORCID: 0000-0001-9162-9063.

^{c)}ORCID: 0000-0003-0534-0664.

^{d)}ORCID: 0000-0002-0713-1623.

Target spectra in a 3.5 h window of data were manually selected and grouped into five classes based on their shape and scattering physics. This preliminary analysis afforded a comparison with the TS values used in the calculation of existing biomass estimates and insight into the complexity of acoustic measurements in the mesopelagic. However, classification of target spectra relied on time-consuming human annotation, and broader application of this type of analysis requires automatic approaches to classify and analyze measured target spectra.

Automatic classification of target spectra is not straightforward. Unsupervised machine learning (clustering) has previously been employed to categorize broadband and multi-frequency echosounder data (Benoit-Bird and Waluk, 2020; Blanluet *et al.*, 2019; Peña, 2018; Ross *et al.*, 2013). There are several advantages to unsupervised approaches over supervised approaches (classification): they do not require annotated training data and may identify trends or groupings in data that a human would not. However, a human must choose the appropriate number of clusters and later interpret the machine-generated groupings. For acoustic scattering applications, these groupings may not be driven by the underlying physics of interest. Conversely, supervised machine learning approaches use annotated training data to build a classification model, and have also recently been applied to multi-frequency echosounder data (Brautaset *et al.*, 2020). A supervised approach ensures that data are categorized into the desired classes, but data annotation will replicate human biases and may not perform well for classes of targets that are relatively rare in the data set and therefore not well represented in the training data.

Here, we “bridge the gap” between unsupervised and supervised classification of broadband target spectra by using acoustic scattering models to develop a training data set. A supervised machine learning model is trained using the k-Nearest Neighbors algorithm (Fix and Hodges, 1951) with these data and used to classify target spectra of mesopelagic organisms measured by an *in situ* echosounder deployed at depth (25–40 kHz) based on their scattering physics. The results provide insight into the distribution of organisms within the mesopelagic as well as the size distribution of organisms that contribute most significantly to volume backscatter measured by a shipboard echosounder.

II. SCATTERING MODELS

Most mesopelagic organisms detected within the 25–40 kHz frequency band will fit into two broad groups: organisms with gas inclusions (e.g., swimbladder-bearing fishes or siphonophores with pneumatophores) and fluid-like organisms (e.g., fish without swimbladders or squid) (Stanton *et al.*, 2010). In this section, we describe the scattering models used to generate training data for each organism type. We note that these two scattering models do not describe all marine organisms. For example, some organisms, such as pteropods, have an elastic scattering response. However, due to their small size, we do not expect to detect

these organisms in the 25–40 kHz frequency band analyzed here (Lavery *et al.*, 2007; Stanton *et al.*, 1994).

A. Gas-bearing organism model

A depth-dependent prolate spheroid swimbladder resonance model (Love, 1978) is used to model organisms with gas inclusions, and the contribution to scattering from the organism body is considered to be negligible (Foote, 1980). We follow the implementation of the model in Nero *et al.* (2004). The scattering cross section, σ_{bs} , in m^2 , is given by

$$\sigma_{bs} = \frac{r_{es}^2}{\frac{f_0^2}{f^2 H^2} + \left(\frac{f_0^2}{f^2} - 1\right)^2}, \quad (1)$$

where f is the incident frequency, in Hz; r_{es} is the equivalent spherical radius of the prolate spheroid, in m; f_0 is the resonance frequency, in Hz; and H is a unitless damping factor. The equivalent spherical radius, r_{es} , is equivalent to $(ba^2)^{1/3}$, where a and b are the lengths of the minor and major axes of the prolate spheroid, respectively. The resonance frequency is calculated as follows:

$$f_0 = \sqrt{\zeta^2 \frac{3\gamma_a P}{4\pi^2 r_{es}^2 \rho_f}}, \quad (2)$$

where γ_a is the ratio of specific heats of air ($\gamma_a = 1.4$); P is the ambient pressure at depth z , in Pa; and ρ_f is the density of the flesh in kg/m^3 . The ambient pressure is given by $P = P_o + \rho_l g z$ where z is the depth, in m; P_o is the atmospheric pressure (101.3 kPa); ρ_l is the density of the surrounding water, in kg/m^3 ; and g is the acceleration due to gravity ($9.81 m/s^2$). A correction factor from Weston (1967) that accounts for the fact that most swimbladders are not spherical ($a \neq b$) is calculated according to

$$\zeta = \frac{\sqrt{2(1 - \epsilon^2)^{1/4}}}{\epsilon^{1/3}} \left(\ln \frac{1 + \sqrt{1 - \epsilon^2}}{1 - \sqrt{1 - \epsilon^2}} \right), \quad (3)$$

where $\epsilon = a/b$. Finally, the damping factor is given by

$$\frac{1}{H} = \frac{2\pi r f^2}{f_0 c_w} + \frac{\xi}{\pi r^2 f_0 \rho_f}, \quad (4)$$

where c_w is the speed of sound in the surrounding water, and ξ is the viscosity of the flesh. The damping factor accounts for radiation and viscous damping but neglects thermal damping as it has been shown to be negligible at depths greater than 100 m (Love, 1978).

The acoustic TS ($TS = 10 \log_{10} \sigma_{bs}$) for a representative gas-bearing organism using Eqs. (1)–(4) is shown in Fig. 1. The TS curve is characterized by three regimes: below resonance (B), where σ_{bs} increases with frequency with a slope greater than or equal to $\sigma_{bs} \propto f^4$, at resonance (Re), where a peak is observed at f_0 , and above resonance (A), where σ_{bs} is relatively

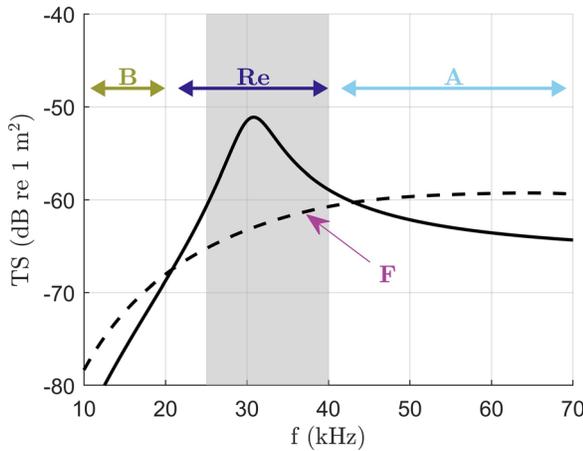


FIG. 1. (Color online) Representative modeled TS spectra for a gas-bearing organism (solid line) and a fluid-like organism (F, dashed line). For the gas-bearing curve, the below resonance (B), at resonance (Re), and above resonance (A) regimes are annotated. The gray shaded region indicates the frequency range analyzed in this work (25–40 kHz). For the gas-bearing curve, $a = 0.4$ mm and $z = 200$ m, and for the fluid-like curve, $L = 10$ cm and $a = 5$ mm. All other parameters are as defined in Table II.

flat, but gradually decreases with increasing frequency. Due to the inverse relationship between f_0 and r_{es} , in general, a smaller organism will have a higher resonance frequency.

B. Fluid-like organism model

The distorted wave Born approximation (DWBA) for a straight, finite-length cylinder (Jones *et al.*, 2009) was used to model fluid-like organisms. For a finite cylinder with length L and radius a , σ_{bs} is given by

$$\sigma_{bs} = \left(\frac{k_l k_v a^2 L}{2} (\gamma_\kappa - \gamma_\rho) e^{-iLk_v \sin\theta} \times \frac{J_1(k_v a \cos\theta) \sin k_v L \sin\theta}{k_v a \cos\theta \quad k_v L \sin\theta} \right)^2, \quad (5)$$

where k_l and k_v are the wave numbers in the surrounding water and within the volume, respectively ($k = 2\pi/\lambda$, where λ is the wavelength, in m), J_1 is the Bessel function of the first kind, and θ is the angle of incidence ($\theta = 180^\circ$ is broadside). γ_κ and γ_ρ are compressibility and density constants, respectively, defined as

$$\gamma_\kappa = \frac{1 - g_v h_v^2}{g_v h_v^2} \quad (6)$$

and

$$\gamma_\rho = \frac{g_v - 1}{g_v}, \quad (7)$$

where g_v is the ratio of the density of the fluid within the volume to the density of the surrounding water, and h_v is the ratio of the sound speed within the volume to the sound speed in the surrounding water.

Real organisms have more complex shapes and material properties than the finite-length cylinder modeled here. As a

result, the structure of σ_{bs} for real organisms will typically be less smooth than predicted by the model. To address this, we average σ_{bs} over ten angles of incidence normally distributed between -10 and 10 degrees (Stanton and Chu, 2000).

Figure 1 shows TS for a representative fluid-like organism. The TS curve is characterized by increasing TS with frequency (with a slope less than or equal to $\sigma_{bs} \propto f^4$) at lower frequencies, followed by a geometric scattering regime in which the frequency response is flat but may contain nulls.

III. METHODS

A. Instrumentation

Target spectra were measured using Deep-See, a towed instrumentation platform, depth rated to 1500 m, that combines active acoustic, optical, and environmental sensors for *in situ* measurement of the mesopelagic zone (Bassett *et al.*, 2020). Deep-See has seven broadband transducers spanning approximately 1–450 kHz, which were operated sequentially. While this approach mitigated the effects of acoustic crosstalk, it also meant that relatively few organisms were observed coincidentally between channels and a continuous spectrum across channels was not possible for most targets. Here, we focus on data from one channel, a custom Airmar transducer with an 18–48 kHz nominal frequency range that is operated by a custom Edgetech system. The receiver for this bi-static system is an eight-channel polyvinylidene fluoride (PVDF) array (a diagram of the PVDF array can be found in the supplemental information found in Bassett *et al.*, 2020). The Airmar transducer transmitted a 10 ms linear frequency-modulated chirp with a bandwidth of 18.5–47.5 kHz at a ping rate of 0.5 Hz. The transducer was aligned with the PVDF receive array in the athwartship (port-starboard) direction, and the centers of the transducer and PVDF array were separated by 65 cm in the alongship (forward-aft) direction. A Seabird 25+ conductivity, temperature, and depth (CTD) sensor was used to measure platform depth and *in situ* water properties.

B. Data collection

Data were collected from 24 July–7 August 2019 during a cruise on the FRV Henry Bigelow in the northwestern Atlantic off of the New England continental shelf break. Deep-See was deployed a total of ten times during this cruise, twice for calibration and testing and eight times for biological scattering measurements (four of these were during daylight hours). In this work, we present detailed results from one of the four daylight deployments (conducted on August 4), but all daytime data were analyzed using the same methods and results were generally consistent across deployments. For all deployments, Deep-See was towed at a speed of approximately 2 knots.

Concurrent volume backscattering measurements were made by a shipboard EK60 echosounder operating at 18, 38, 120, and 200 kHz. The EK60 was calibrated using a standard target (Demer *et al.*, 2015) following the cruise on

20 August 2019. Figure 2 shows the 38 kHz shipboard echosounder data collected during the 4 August deployment, the depth profile of Deep-See, and representative samples of the *in situ* echosounder data. In the shipboard data, a dominant scattering layer is seen between approximately 400 and 500 m depth, and a second, less intense scattering layer is observed between approximately 600 and 700 m depth.

C. Broadband data processing

1. Matched-filtering and split-beam processing

Matched filtering (Chu and Stanton, 1998; Turin, 1960) was performed in the Edgetech data acquisition software, and the filtered, decimated signal from each of the eight PVDF channels was recorded to a range of 300 m. In all further analysis, we only use data from the inner four channels of the PVDF array. These four quadrants form a circular array with a diameter of 58 cm, similar to commercially-available four-sector transducers. Split-beam processing using these four channels was implemented to determine the alongship and athwartship off-axis angles (ϕ_f and ϕ_p , respectively) of the target relative to the transducer (Bassett et al., 2020; Simmonds and MacLennan, 2005). The broadband off-axis angle was then calculated as $\phi = \sqrt{\phi_f^2 + \phi_p^2}$.

The broadband TS was determined from the coherent sum of the four channels, v ,

$$TS = 20 \log_{10}(|v|) + 40 \log_{10}r + 2\alpha r + 2G_c, \quad (8)$$

where r is the range from the transducer, in m, using the time difference of arrival; α is the attenuation coefficient at the center frequency, in dB/m (Francois and Garrison, 1982a,b); and G_c is the gain value from calibration at the center frequency (see Sec. III C 2). Range and $\alpha(f)$ were calculated using water properties measured by the CTD.

2. Calibration

Because of the offset between the transducer and receive array, the two-way beam pattern of the Airmar/PVDF system, b , is a function of alongship angle relative to the receiver array (ϕ_p), athwartship angle relative to the receiver array (ϕ_f), range (r), and frequency (f). The beam pattern was measured dockside in June 2020 for a target at a single range, and this measured beam pattern was compared to the theoretical formulation used for *in situ* TS measurements.

A 20 cm diameter aluminum sphere was suspended at a range of approximately 8.9 m below Deep-See in a test well that was approximately 15 m deep, depending on tidal elevation. At this range, the target was beyond the acoustic near-field (Medwin and Clay, 1998) and far enough from the seafloor to avoid interference due to the matched filter side-lobes (Lavery et al., 2017). The pitch and roll of Deep-See were gradually varied while data were collected at 4 Hz. The angular position of the target in each ping was calculated as the alongship and athwartship angles (ϕ_f and ϕ_p) at the peak in v associated with the target.

First, the split-beam processing was verified by ensuring that the relationships between vehicle pitch and ϕ_f and roll and ϕ_p were approximately 1:1. Then, the uncalibrated frequency response of the target, $TS_u(f)$, was calculated using a 0.65 m window with the target peak at 0.15 m, v_{targ} . A Tukey window with a cosine fraction of 0.05 was applied to v_{targ} before padding the signal to a length of 2048 points and calculating the frequency content of the target, V_{targ} , using a fast Fourier transform (FFT). $TS_u(f)$, in dB re 1 m^2 , was then calculated as follows:

$$TS_u(f) = 20 \log_{10}(|V_{targ}|) + 40 \log_{10}r_{targ} + 2\alpha(f)r_{targ}, \quad (9)$$

where r_{targ} is the range to the peak in the matched filter data, in m, using the time difference of arrival, and $\alpha(f)$ is the

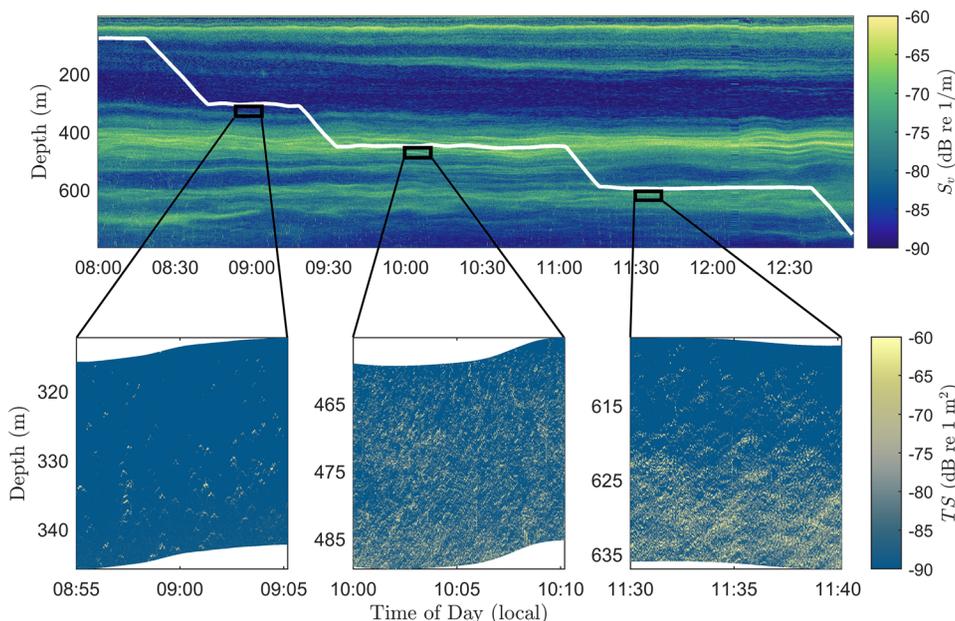


FIG. 2. (Color online) (Top) Shipboard 38 kHz echosounder data. The white line indicates the depth of Deep-See. (Bottom) Representative 15-min segments of *in situ* match-filtered echosounder data collected with the Deep-See Airmar channel (25–40 kHz). Note that Deep-See was towed behind the ship and data are not exactly concurrent. All data were collected on 4 August 2019.

frequency-dependent attenuation coefficient, in dB/m (Francois and Garrison, 1982a,b).

The average of the highest ten observed values of $TS_u(\phi_p, \phi_f)$ at each frequency, $TS_{u,max}$, were assumed to be the TS of the target when $b = 0$. For each frequency, the measured beam pattern at $r = 8.9$ m, b_{meas} , was then approximated as

$$b_{meas}(\phi_p, \phi_f) = TS_u(\phi_p, \phi_f) - TS_{u,max}. \tag{10}$$

The theoretical beam pattern was calculated by assuming that both the transmitting transducer (Airmar) and the receiving transducer (PVDF array) could be modeled as piston transducers with directivity patterns described by

$$D = \frac{2J_1(ka \sin \phi)}{ka \sin \phi}, \tag{11}$$

where J_1 is the Bessel function of the first kind, k is the wavenumber ($k = 2\pi/\lambda$, where λ is the wavelength, in m), a is the radius of the transducer, in m, and ϕ is the off-axis angle (Medwin and Clay, 1998). The Airmar transducer is an off-the-shelf system with a diameter of 13 cm, and the effective diameter of the PVDF array was inferred to be 50 cm by minimizing error between the theoretical beam pattern and the measured data. The beam pattern, $b(f, \phi)$, is then

$$b(f, \phi) = 10 \log_{10}(D_t^2(f, \phi_t)D_r^2(f, \phi_r)), \tag{12}$$

where D_t and D_r are the directivity patterns of the transmitting transducer (Airmar) and the receiving transducer (PVDF array), respectively, and ϕ_t and ϕ_r are the broadband phase angles relative to the two transducers. The phase angle relative to the PVDF array, ϕ_r , is the broadband phase angle at the target peak from split-beam processing. For a given target with off-axis angles relative to the PVDF array of ϕ_p and ϕ_f , the off-axis angles relative to the Airmar transducer, $\phi_{p,t}$ and $\phi_{f,t}$, can be calculated from the system geometry

$$\phi_{p,t} = \tan^{-1}\left(\frac{d - r \tan(\phi_p)}{r}\right), \tag{13}$$

$$\phi_{f,t} = \phi_f, \tag{14}$$

where d is the distance between the centers of the PVDF array and the Airmar transducer (0.65 m). The broadband

phase angle relative to the Airmar transducer, ϕ_t , is then $\phi_t = \sqrt{\phi_{p,t}^2 + \phi_{f,t}^2}$.

Figure 3 shows the measured beam pattern, the theoretical beam pattern, and the deviation between the two at the center frequency, 32.5 kHz. Within 2 degrees, the mean absolute deviation between the measurements and theory was 0.3 dB re 1 m^2 at 38 kHz. We note that this deviation was higher when we did not account for the bistatic angle ($\phi_t = \phi_r$), and the mean absolute deviation increased to 0.7 dB re 1 m^2 , though the effect of the bistatic correction decreases with range. Deviation was similar at higher and lower frequencies, with mean absolute deviations between the bistatic theoretical beampattern and the measured beampattern of 0.38 and 0.22 dB at 23.5 and 38 kHz, respectively.

The calibration curve, $G(f)$, was measured *in situ* on 28 July 2019. The same 20 cm diameter aluminum target that was used to measure the beam pattern was suspended below Deep-See at a range of approximately 20 m while Deep-See was towed throughout the water column. Due to the pitch of Deep-See, the target was never on-axis ($\phi < 0.25$ degrees), and the theoretical beam pattern derived above was applied to targets with $\phi < 2$ degrees to correct their TS based on their position within the beam. Both full-wave and partial-wave calibration techniques were used to achieve a relatively smooth calibration curve, $G(f)$, between 25 and 40 kHz (Bassett *et al.*, 2020; Stanton and Chu, 2008; Stanton *et al.*, 2010). Calibration of this transducer/receiver array during a previous cruise found minimal variation in the calibration curve as a function of depth (see supplemental material found in Bassett *et al.*, 2020).

D. Target detection and processing

Targets in the *in situ* data were automatically detected and tracked using custom software implemented in MATLAB R2019B. Target detection was implemented on the broadband TS signal [Eq. (8)] following the wideband target detection algorithm used in the Echoview software.¹ A 30 cm window of v centered around the target peak, v_{targ} , was used to calculate the frequency-dependent TS, $TS(f)$, of each detected target. To do this, a Tukey window with a cosine fraction of 0.05 was applied to v_{targ} before using a fast Fourier transform (FFT) to calculate the spectrum of the target, V_{targ} . The TS spectrum, $TS(f)$, in dB re 1 m^2 was calculated as

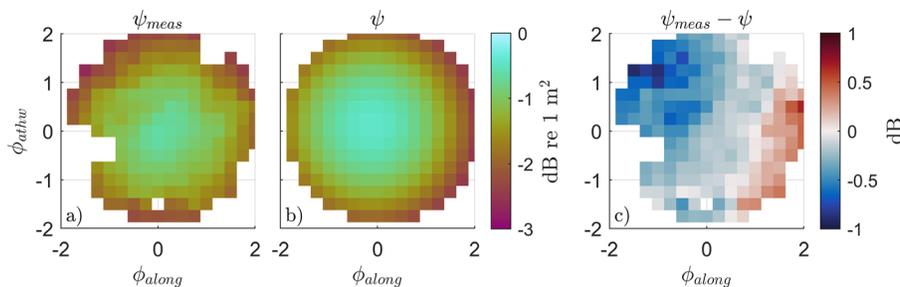


FIG. 3. (Color online) Measured and theoretical two-way beampattern at the center frequency, 32.5 kHz, at a range of 8.9 m. The leftmost panel shows the measured beampattern, the middle panel shows the theoretical beampattern used for TS calculations, and the rightmost panel shows the deviation between the two.

$$TS(f) = 20 \log_{10}(|V_{targ}|) + 40 \log_{10}(r_{pk}) + 2\alpha(f)r_{pk} + 2G(f) + b(\phi, f), \tag{15}$$

where r_{pk} is the range from the transducer to the target peak using the time difference of arrival, in m; $\alpha(f)$ is the frequency-dependent attenuation coefficient, in dB/m (Francois and Garrison, 1982a,b); $G(f)$ is the gain curve from calibration, in dB re 1 m^2 ; b is the two-way beampattern of the bistatic system, in dB re 1 m^2 ; and ϕ is the off-axis angle from split beam processing at the target peak.

The minimum range for target detection was 10 m, and a threshold was used to dynamically set the maximum range to maximize the number of targets detected while avoiding regions where individual targets were not detectable (e.g., volume backscattering). Targets were considered to be detectable when the median broadband TS in 2-meter, 20-s bins did not exceed $-85 \text{ dB re } 1 \text{ m}^2$ (i.e., 50% of samples in the bin did not contain a target with TS above this threshold).

To avoid double-counting the same target in multiple pings, detected targets were tracked through the beam using an alpha-beta tracking algorithm (Blackman, 1986), again following the implementation used in the Echoview software.² Two changes were made to adapt this algorithm for broadband target detection from a depth-varying platform: the mean absolute difference between $TS(f)$ for two targets was used in lieu of ΔTS at a single frequency, and the z position for each target was defined as its depth (sum of platform depth and r_{pk}) rather than the range from the transducer. For any target detected in multiple pings, $TS(f)$ for the target was defined as the linear average of all detections. Target detection and tracking parameters are provided in Table I.

E. System noise level

Due to spreading and attenuation, the signal-to-noise ratio of a given target decreases as a function of range. The minimum detectable TS will therefore increase as a function of range. To investigate this effect and determine the lowest TS organisms that could be detected, the system noise level was calculated using a 30 min segment of data collected in a region with relatively low organism density (8:45–9:15 AM in Fig. 2). First, target detection was performed with lowered thresholds ($\phi_{max} = 360^\circ$, $TS_{min} = -90 \text{ dB re } 1 \text{ m}^2$,

$\sigma_{\phi, max} = 1^\circ$) to identify regions of the data where targets were present (i.e., not noise) and an 0.5 m window around each detected target was removed from the data set. The noise level as a function of frequency $N(f)$ was then calculated in 0.5 m windows with 50% overlap for all windows that did not contain a detected target by applying Eq. (15) with $b = 0$. Finally, the median noise level as a function of range, $N(r, f)$, was calculated as the median of all $N(f)$ curves in each range bin.

F. Target spectra annotation

A human reviewer (first author) annotated 100 target spectra randomly selected from each of the four daytime Deep-See dives (400 total target spectra). The reviewer classified spectra into the four classes described in Fig. 1, plus an additional class for target spectra that were not well-described by the scattering models:

- *at resonance* (R): resonance peak observed between 25 and 40 kHz
- *above resonance* (A): decrease in TS with frequency, and TS exceeds $-54 \text{ dB re } 1 \text{ m}^2$ at 25 kHz
- *below resonance* (B): increase in TS with frequency with a slope greater than $\sigma_{bs} \propto f^4$
- *fluid-like* (F): increase in TS with frequency with a slope less than $\sigma_{bs} \propto f^4$ or flat, negative slope, or complex shape with TS that does not exceed $-54 \text{ dB re } 1 \text{ m}^2$
- *complex/unclassified* (C): any target spectra that does not fit into the aforementioned categories

The $-54 \text{ dB re } 1 \text{ m}^2$ threshold was determined through analysis of the modeled spectra using the parameters in Table II. Gas bearing targets observed well below resonance ($f_0 \gg 40 \text{ kHz}$) will have slopes approaching or equal to $\sigma_{bs} \propto f^4$, and would likely be difficult or impossible to discern from fluid-like targets. These targets, however, would have relatively low TS and would likely not be detected (see Sec. III E). Similarly, fluid-like targets larger than 11 cm (the upper bound in Table II) may have relatively flat TS spectra that, following the criteria described here, would be categorized as above resonance. While there are little data published on size distributions in the sampled region, non-swimbladdered fishes are expected to be smaller than this threshold (Davison et al., 2015; Dornan et al., 2019; Irigoien et al., 2014; Underwood et al., 2020), and studies in

TABLE I. Parameters used for single target detection and tracking. All weights for target tracking were 50%.

Detection Parameters			Tracking Parameters		
Parameter	Units	Value	Parameter	Units	Value
Pulse length determination level	dB re 1 m^2	6	α	—	0.1
Maximum ϕ	Degrees	2	β	—	0.1
Minimum pulse length	m	0.04	Exclusion distance	m	1
Maximum pulse length	m	0.3	Missed ping expansion	%	50
Maximum std. of ϕ	Degrees	0.6	Max ping gap	Count	2
Minimum TS	dB re 1 m^2	-80	Maximum ΔTS	dB re 1 m^2	2
Minimum separation	m	15			

TABLE II. Parameters used to generate training data from scattering models. **Bold** parameters were randomly selected within the specified range, while all other parameters were held constant. To produce more realistic spectra for fluid-like organisms, σ_{bs} was averaged over ten angles of incidence, θ , normally distributed between -10 and 10 degrees.

Gas-bearing organism model			Fluid-like organism model		
a	Minor axis length	0.2–4 mm	L	Length	3–11 cm
z	Depth	200–800 m	g	Density ratio	1.04
γ_a	Ratio of specific heats of air	1.4	h	Sound speed ratio	1.04
P_o	Atmospheric pressure	101.3 kPa	ϵ	Aspect ratio (radius/length)	15
c_w	Sound speed in water	1500 m s ⁻¹	c_w	Sound speed in water	1500 m s ⁻¹
ϵ	Axis ratio	0.5			
ρ_f	Fish flesh density	1050 kg m ⁻³			
ξ_f	Fish flesh viscosity	5 kg m ⁻¹ s ⁻¹			

other regions suggest that other organisms larger than 11 cm are expected to be relatively rare (Benoit-Bird and Au, 2001).

To aid in annotation, a curve with slope proportional to f^4 was shown to the reviewer on the same axis as each annotated target spectra. This allowed the reviewer to discern between target spectra classified here as fluid-like organism and gas-bearing organisms below resonance based on their slope. The reviewer also indicated whether each annotation was high or low certainty. Examples of targets that would be assigned low certainty include a target that appears to have a resonance peak at the upper edge of the frequency band, and the distinction between “at” and “below” resonance is ambiguous, or a target spectrum that increases with frequency with a relatively steep slope ($>f^4$), but is not smooth or has nulls. Figure 4 shows an example of each class of target spectra (high certainty levels) from the annotated data set, and all annotated target spectra are shown in the supplementary material, Fig. S1.³

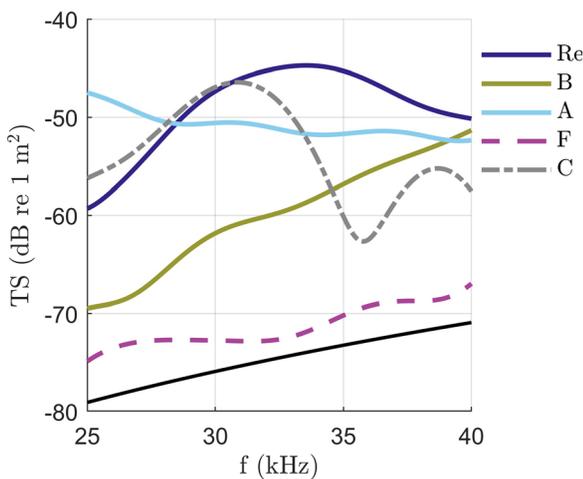


FIG. 4. (Color online) examples of measured target spectra associated with each class: at resonance (R), below resonance (B), above resonance (A), fluid-like (F), and complex/unknown (C). The solid black curve at the bottom of the figure shows a spectrum where $\sigma_{bs} \propto f^4$ ($TS \propto 40 \log_{10} f$). This steepness threshold was used to distinguish between fluid-like scatterers and gas-bearing scatterers observed below resonance.

G. Model-generated training data

Training data were generated for the non-complex/unclassified target classes using the scattering models described in Sec. II. Table II lists the parameters used for each model, including which parameters were varied and which were held constant. Where available, parameter values representative of mesopelagic fishes from previous studies were used (Davison *et al.*, 2015; Love, 1978; Scoulding *et al.*, 2015; Yasuma *et al.*, 2003). The selected value of the swimbladder aspect ratio, $\epsilon = 0.5$, is the average aspect ratio of mesopelagic lanternfishes (*Dhiapus theta*) measured in a laboratory in Yasuma *et al.* (2003).

Parameters were randomly generated within the ranges specified in Table II by drawing samples from a uniform distribution to generate 3000 gas-bearing and 2000 fluid-like target spectra. The modeled gas-bearing target spectra were classified as below, at, or above resonance based on where the resonance frequency fell relative to the echosounder frequency band: below resonance for $f_0 > 39.5$ kHz, at resonance for $25.5 \text{ kHz} \leq f_0 \leq 39.5 \text{ kHz}$, and above resonance for $f_0 < 25.5 \text{ kHz}$. The modeled data supported our approach to human annotation - for the modeled gas-bearing target with the highest resonance frequency ($f_0 = 130 \text{ kHz}$), σ_{bs} had a slope proportional to $f^{4.3}$, and the fluid-like target with the steepest frequency response in the echosounder frequency band had a slope proportional to $f^{3.8}$.

H. Classification model

Because our objective was to group targets based on both the magnitude and the shape of the TS spectra, both $TS(f)$ and its slope, dTS/df , were used for automatic classification. The measured and modeled target spectra were both smoothed and preprocessed for classification in the same way. Because the real target spectra were less smooth than the model-generated spectra, $\sigma_{bs}(f)$ was averaged in 2 kHz bins with 50% overlap. The smoothed spectra ($TS = 10 \log_{10} \sigma_{bs}$) and its first difference (δTS) were then assembled into an $n \times 1$ vector, x , where n is the sum of the number of points in the smoothed spectra (31) and the number of points in dTS (30). The vectors x associated with the model-generated training data were assembled into an $m \times n$ matrix $\mathbf{X}_{\text{train}}$, where m is the total number of model-generated targets (400). The maximum and

minimum values of each row of \mathbf{X} were used to normalize the data such that each value of TS and δTS ranged between 0 and 1. The same values (maximum and minimum values of the training data) were used to normalize the measured target spectra. This prevented values with smaller magnitudes (i.e., dTS) from dominating the classification model.

The k -nearest neighbor algorithm was used to predict the class of each measured target spectrum. This algorithm was selected for classification because it has been shown to be robust to the noisy data characteristic of acoustic measurements. We note that other supervised, multi-class classification algorithms, such as the random forest algorithm (Breiman, 2001), may provide similar performance.

The Euclidean distance, d , between its associated vector x and each column of $\mathbf{X}_{\text{train}}$ was calculated, and the five model-generated targets with the smallest values of d were selected (nearest neighbors). For each target class, c , represented in the nearest neighbors, a weighted distance was calculated,

$$W_c = \frac{\sum \frac{1}{d_c}}{\sum \frac{1}{d_n}}, \tag{16}$$

where d_c are the distances between x and the nearest neighbors belonging to class c , and d_n are the distances between the unclassified target and all five nearest neighbors. The unclassified target was then tentatively predicted to belong to the class with the highest value of W_c . The average distance between the unclassified target and the nearest neighbors belonging to the predicted class, \bar{d}_c , was calculated. If \bar{d}_c exceeded a specified threshold, d_{max} , the target was determined to not fit well into any of the five modeled classes, and assigned to the complex/unknown target class. This distance threshold was empirically tuned by optimizing the weighted F_1 score for the validation data set (see Sec. III). We note that this classification approach is equivalent to a distance-weighted k -nearest neighbor algorithm (Dudani, 1976) with $k = 5$ and a distance threshold.

I. Model tuning and validation

To tune the complex target threshold, d_{max} , and assess classification performance, the model described in the previous section was used to predict the classes of the 400 human-annotated targets with values of d_{max} ranging from 0 (no complex targets) to the maximum observed value of \bar{d}_c (all targets are complex). For each value of d_{max} tested, the F_1 score was calculated for each class of target spectra. The F_1 score for an individual class is defined as

$$F_1 = \frac{TP}{TP + 0.5(FP + FN)}, \tag{17}$$

where TP is the number true positives (target spectra correctly classified as belonging to the class), FP is the number of false positives (target spectra incorrectly classified as

belonging to the class), and FN is the number of false negatives (target spectra belonging to the class that were incorrectly classified). The F_1 score is equivalent to a weighted average of precision and recall. The weighted average of the F_1 score for all classes was then computed, where the F_1 score was weighted by the number of annotated target spectra in each class. The value of d_{max} that yielded the highest weighted-average F_1 score was retained in all further analyses. We note that we use the same data for model tuning and validation due to the relative scarcity of complex target spectra (nine complex target spectra out of 400 annotations). Complex target spectra were rare in the data set because the selected scattering models were representative of the observed biology. While this is not best practice, we feel that it is preferable to the alternative of removing complex target spectra from the validation data set. However, in applications where more training data is available or there are a significant number of target spectra that are not well-described by scattering models, it would be recommended to split the annotated data into two training sets, one for testing, and one for validation.

J. Analysis of classified targets

After validation and tuning, the classification model was used to predict the class of all measured target spectra from the 4 August 2019 dive, and the following section describes the analysis of the classified target spectra. All depth-varying metrics are calculated in 25 m depth bins except where noted.

1. Distribution of target classes

The fraction of targets belonging to each class as a function of depth, $p_c(z)$, was calculated

$$p_c(z) = \frac{n_c(z)}{n(z)}, \tag{18}$$

where $n_c(z)$ is the number of targets belonging to the class, c , in the depth bin, and $n(z)$ is the total number of detected targets in each depth bin. The depth of each target is defined as the sum of the range to the peak in the matched filtered data associated with the target (r_{pk}) and the depth of Deep-See.

2. Contribution to volume backscatter

The contribution to volume backscatter from each class was estimated for comparison with traditional shipboard measurements. First, the average backscattering cross section in the echosounder frequency band was calculated as a function of depth for all targets [$\sigma_f(z)$] and for targets associated with each class [$\sigma_{f,c}(z)$]. The backscattering cross section of a target at frequency f was calculated as the average of $\sigma_{bs}(f)$ in a 1 kHz window around f .

The volume backscattering coefficient, s_v , in m^{-1} , is defined as

$$s_v = \frac{\sum n\sigma_{bs,n}}{V}, \tag{19}$$

where n is the number of targets with a given backscattering cross section, $\sigma_{bs,n}$, and V is the volume sampled by the beam, in m^3 . Therefore, the volume backscattering coefficient that would be measured by a shipboard echosounder operating at frequency f is predicted by

$$s_v(z) = \frac{n(z)\overline{\sigma_f(z)}}{V(z)}, \tag{20}$$

where $n(z)$ is the total number of targets in each depth bin, $\overline{\sigma_f(z)}$ is the average backscattering cross section of all organisms at frequency f , and $V(z)$ is the volume of water sampled by the echosounder in each depth bin. This requires the assumption that the detected organisms are representative of the distribution of organisms in each depth bin. The volume backscatter at a given frequency from each individual class, $s_{v,c}(z)$ can be approximated by

$$s_{v,c}(z) = \frac{n_c(z)\overline{\sigma_{f,c}(z)}}{V(z)}, \tag{21}$$

where the subscript c indicates organisms associated with class c . The fractional contribution to the total volume backscattering by each class, $p_{sv,c}$, can then therefore be estimated as

$$p_{sv,c} = \frac{n_c(z)\overline{\sigma_{f,c}(z)}}{n(z)\overline{\sigma_f(z)}}, \tag{22}$$

where the measured volume terms cancel.

3. Size of targets observed at resonance

For gas-bearing targets at resonance, the resonance frequency (f_0) was defined as the frequency at the peak in the measured target spectrum. The resonance frequency was then used to approximate the size of the gas inclusion of the detected organism. To do this, Eq. (2) was solved for r_{es} , using the values of ρ_f , ξ_f and ϵ in Table II and ambient pressure calculated using the depth of the measured organism. The average value of r_{es} as a function of depth, $r_{es}(z)$, was calculated in 50m depth bins. To aid in interpretation, the largest ($f_0 = 25$ kHz) and smallest ($f_0 = 40$ kHz) values of r_{es} that could be resolved in the echosounder frequency band were also calculated at the center of each depth bin. Finally, to ensure that the resonance model was a good fit for the measured data, σ_{bs} was modeled using the estimated value of r_{es} and the parameters in Table II [Eq. (1)], and the mean absolute deviation between the measured and modeled spectrum was calculated.

IV. RESULTS

A. System noise level

Figure 5 shows the system noise level, $N(r, f)$, processed for comparison with measured TS's, at the center frequency,

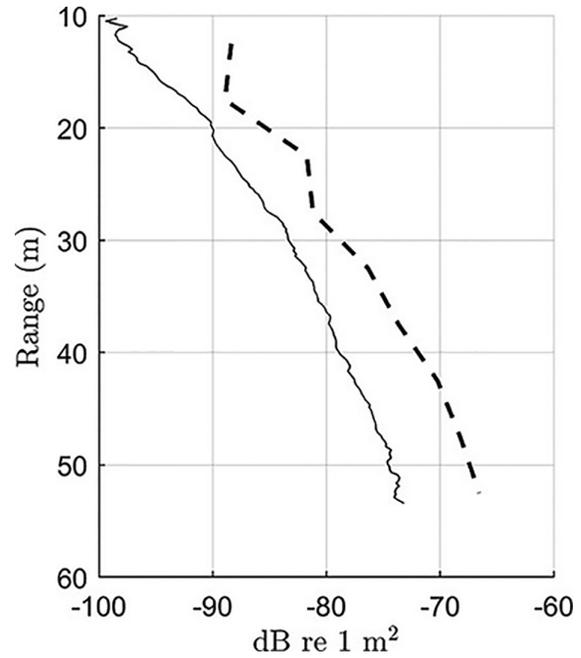


FIG. 5. System noise level, $N(r, f)$, at the center frequency, $f = 32.5$ kHz, (solid line) and the minimum TS value at 32.5 kHz of any target detected as a function of range from the transducer (dashed line). The system noise level was relatively insensitive to frequency.

$f = 32.5$ kHz, and the minimum TS detected at 32.5 kHz as a function of range from the transducer. On average, the minimum detected TS is 6 dB above the calculated noise floor. While the noise floor is constant with range, spreading and attenuation cause the apparent noise floor in TS to increase with range. This indicates that weakly scattering (low TS) organisms are likely under-counted to a greater extent as range increases. However, due to the logarithmic nature of scattering, this does not have a significant impact on the average TS, which was relatively insensitive to range and should not impact detection of strongly scattering organisms with gas inclusions.

B. Model validation

Figure 6 shows classification performance for the 400 human-annotated target spectra after tuning the complex target threshold ($d_{max} = 1.83$). The weighted-average F_1 score was 0.90 and the predicted class agreed with human annotation for 91% of annotated target spectra. The reviewer had low certainty for 46% of classifications that did not agree with human annotation. As shown in Fig. 6, 20 targets annotated as below or above resonance were classified as at resonance. Figures 6(a) and 6(b) show that these were target spectra with resonance peaks on the edge of the echosounder frequency band, where distinction between classes was ambiguous. Seven “complex/unknown” target spectra were also classified as resonant targets [Fig. 6(c)]. These were relatively noisy target spectra that the human reviewer did not feel confident classifying, but still have a notable peak in TS. These results indicate that the machine learning model

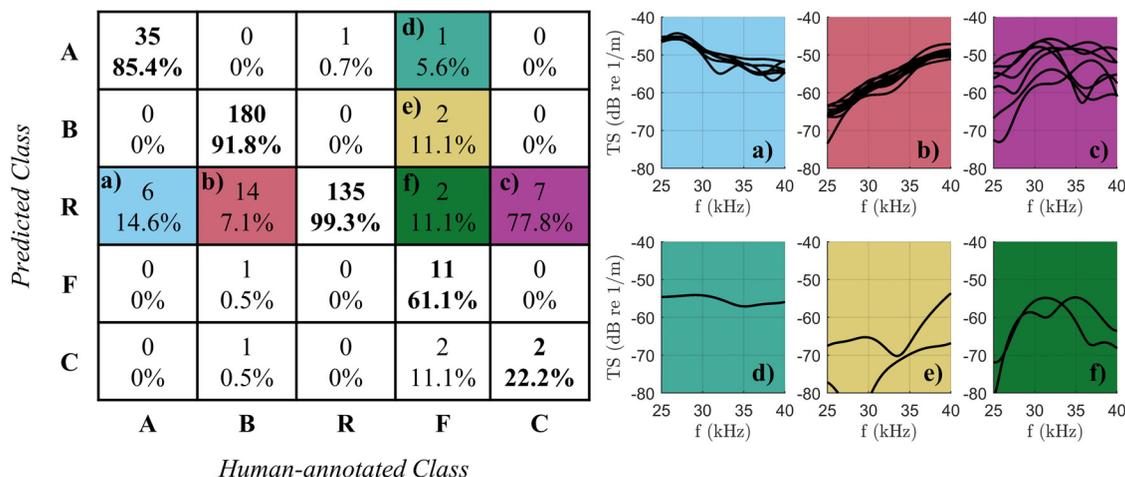


FIG. 6. (Color online) (left) Confusion matrix for classification of human-annotated target spectra. [right; (a)–(f)] Target spectra in representative matrix cells where human annotation did not match the predicted class. The color of each figure corresponds to the cell in the confusion matrix of the same color. See Sec. IV B for a detailed description of (a)–(f).

was likely more consistent in classifying targets at resonance than the human reviewer.

Only 61.1% of targets annotated as fluid-like were classified as such. One was classified as above resonance [Fig. 6(d)], and was relatively flat with TS near the -54 dB re 1 m² threshold. In this case, distinction between a large, fluid-like organism and a gas-bearing organism observed well above resonance is ambiguous. Two were classified as below resonance [Fig. 6(e)] and had slopes that indicated they were associated with gas-bearing organisms but were relatively noisy detections with bumps in the target spectra. In these cases, it is not clear whether the human annotations or model predictions are correct. Finally, two target spectra with relatively low TS, but sharp peaks in TS, were annotated as fluid-like but classified as at resonance. Again, classification here is ambiguous, and these could be gas-bearing organisms or fluid-like organisms in the geometric scattering regime with nulls observed at the edges of the frequency band. We note that relatively few fluid-like organisms were detected (see Sec. IV C) and classification performance for this class is largely inconsequential for the analysis presented here.

C. Analysis of classified targets

Figure 7 shows the number of targets detected in each depth bin, and Fig. 8 shows the fraction of targets belonging to each class as a function of depth, $p(z)$, and their fractional contribution to volume backscatter, $p_{sv}(z)$, at three representative frequencies: 26, 32, and 38 kHz. All classified target spectra are shown in Fig. S2³ in the supplementary materials. Throughout the water column, the majority of target spectra were observed at or below resonance, except at around 300 m depth. While only four targets were detected in the 300 m depth bin due to relatively low organism density in this region (see Fig. 2), the increased presence of targets observed above resonance at shallower depths is consistent with compression of the gas inclusion with depth

(i.e., gas inclusions are larger near the surface). Fluid-like and complex/unknown target spectra comprised less than 3.4% and 0.3%, respectively, of all measured organisms. These values are consistent with the preliminary analysis in Bassett *et al.* (2020) of data obtained in 2018 in the same region. We note that the relatively low number of fluid-like organisms detected does not necessarily indicate that there is a low number of fluid-like organisms present. Rather, it indicates that they are weakly scattering organisms that do not contribute significantly to scattering in the frequency band used (25–40 kHz) and smaller organisms of this type may not be detected, especially at longer ranges from the sonar (as discussed in Sec. IV A). Therefore, we cannot draw conclusions regarding the abundance of fluid-like organisms.

The fractional contribution to volume backscatter from each class of target spectra (p_{sv}) is not equivalent to fractional distribution of targets (p). Due to their relatively high TS, targets observed at resonance account for the majority of volume backscatter at nearly all depths and frequencies, even where they do not constitute the majority of observed targets. For example, at 450 m depth (within the scattering layer observed by the 38 kHz shipboard echosounder), target spectra observed below and at resonance account for 67% and 27% of measured organisms, but contribute to 33% and 66% of volume backscatter at 38 kHz, respectively. The fractional contribution to volume backscatter from gas-bearing organisms observed below resonance increases with frequency due to the steep slope of the TS spectra; at 26 kHz, there is near-zero contribution from organisms observed below resonance. Conversely, the fractional contribution from gas-bearing organisms observed above resonance decreases with increasing frequency due to the slightly negative slope of the TS spectra. Across all measured frequencies, at all depths, target spectra associated with gas-bearing targets (at, above, and below resonance) account for at least 88% of volume backscatter.

For target spectra classified as at resonance, the resonance model agreed well with measurements; the average

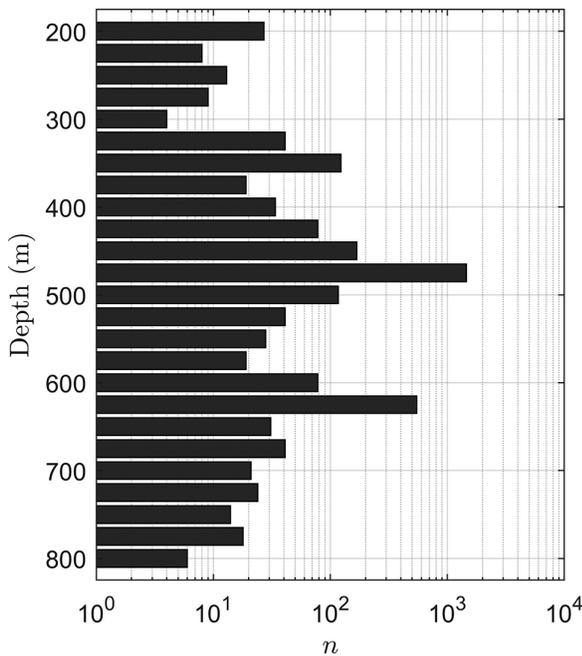


FIG. 7. Number of targets detected in each depth bin. Note that the number of targets detected is a function of both time spent sampling at that depth and organism abundance.

mean absolute deviation between the measured and modeled TS spectra was 1.7 dB re 1 m^2 with a standard deviation of 1 dB re 1 m^2 . Figure 9 shows inferred $\overline{r_{es}(z)}$, as well as the range of inferred r_{es} values that could be resolved at each depth in the echosounder frequency band. Both the size of the smallest gas inclusion and the range of gas inclusion sizes that can be resolved increase with depth. Throughout most of the water column, the average inferred radius from target spectra observed at resonance, $\overline{r_{es}(z)}$, follows a

similar trend and increases with depth as is expected due to the increasing ambient pressure. However, a notable decrease in $\overline{r_{es}(z)}$ is observed around 500 m depth, towards the bottom of the scattering layer observed in the 38 kHz volume backscatter measured by the shipboard echosounder. This, in conjunction with the relatively low number of target spectra observed above resonance and high fraction of target spectra observed below resonance, suggests that organisms likely had smaller gas bladders at this depth than elsewhere in the water column. This trend could also be related to other factors, including the material properties of the organisms (e.g., higher ρ_f) or the inflation of gas bladders (e.g., non-neutrally buoyant inflation). While the swimbladder aspect ratio used in our resonance model ($\epsilon = 0.5$) is based on measurements made at the surface, the estimated equivalent spherical radii were relatively insensitive to the value of ϵ [Eq. (3)]. On average, $\overline{r_{es}(z)}$ increased by 0.04 mm when ϵ was decreased to 0.2. We note that the same trends in $\overline{r_{es}(z)}$ were observed in the data from the other daytime Deep-See dives during the same cruise.

V. DISCUSSION

We have demonstrated that acoustic scattering models can be used to train a machine learning model for classification of broadband target spectra based on their scattering physics. While this approach does not solve the “grand challenge” of species identification (MacLennan and Holliday, 1996), it does provide insight into organism size distribution, and we anticipate that a similar approach could be used for broadband echosounder data from transducers with different operating frequencies. Further, our approach of using model-generated training data could likely be adapted for different machine learning algorithms or

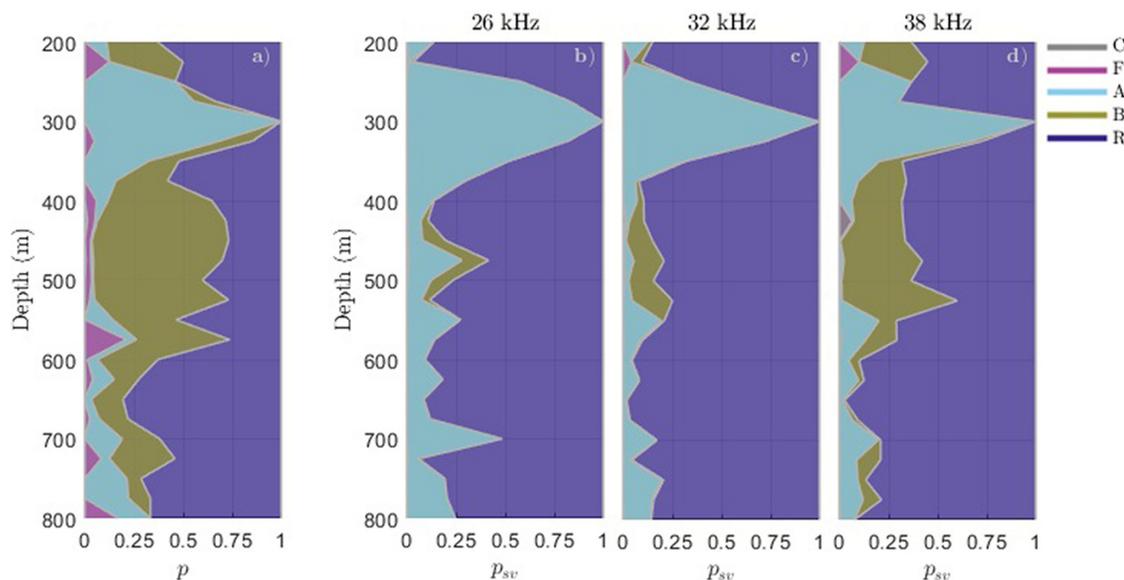


FIG. 8. (Color online) (a) Fractional distribution of detected targets belonging to each class, $p(z)$, and (b)–(d) their contribution to volume backscatter, $p_{sv}(z)$, at 26, 32, and 38 kHz, respectively. The cumulative fraction of targets is shown, such that the colored area corresponding to each target class represents the fractional contribution from that class. All target classes are plotted, but the contribution from complex and fluid-like targets is too small to see throughout most of the water column.

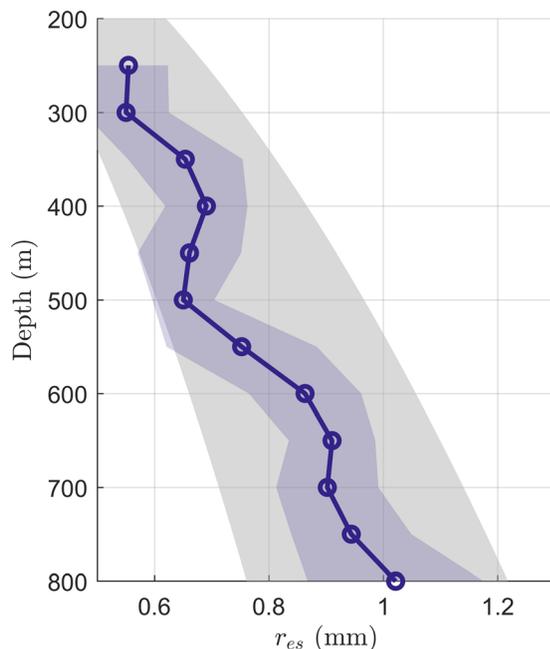


FIG. 9. (Color online) Average, inferred equivalent spherical radius of targets detected at resonance for each deployment. The colored, shaded region indicates the standard deviation, and the gray, shaded region indicates the range of inferred gas inclusion sizes that could be resolved at each depth ($25 \text{ kHz} > f_0 > 40 \text{ kHz}$).

scientific objectives. However, application of this approach to data collected at different frequencies or in a different environment will require careful selection of scattering models and parameters.

The results of target spectra classification allow us to evaluate assumptions that have been made in previous studies to convert shipboard volume backscattering to organism abundance. First, the application of TS-length regressions at 38 kHz (as used in Irigoien *et al.*, 2014) requires the assumption that scattering from gas-bearing organisms is in the geometric scattering regime (e.g., above resonance). The *in situ* target spectra measurements here show that most organisms are observed at or below resonance, and TS-length relationships are not sufficient to model the TS of these organisms at mesopelagic depths. While the effect of resonance has been hypothesized or modeled in previous studies (Davison *et al.*, 2015; Escobar-Flores *et al.*, 2020; Proud *et al.*, 2019), it has not previously been quantified using *in situ* measurements.

We can gain further insight by applying our results to the concurrent shipboard measurements made at 38 kHz. To do this, for every ping from the *in situ* echosounder, we calculated a one-minute average of s_v measured by the shipboard 38 kHz echosounder in the depth range used for *in situ* target detection. These “local” s_v values were then averaged in 25 m depth bins and multiplied by the values of p_{sv} in Fig. 8. In sum, 98.9% of volume backscatter at 38 kHz from 200 to 800 m deep is likely attributable to targets with gas inclusions (targets observed below, at, or above resonance). This supports assumptions made in previous studies that 100% of volume backscatter at 38 kHz can be attributed

to organisms with gas inclusions (Irigoien *et al.*, 2014; Proud *et al.*, 2019). We can further constrain this assumption at our test site: 70.4% of volume backscatter can be attributed to targets with resonance frequencies between 25 and 40 kHz (gas inclusions in the size range described in Fig. 9). However, these organisms only comprise 43% of all detections according to the classifications described here. While our measurements do not facilitate a direct comparison, these findings are in accordance with Kloser *et al.* (2016), who found that 88% of volume backscattering at 38 kHz from 25 to 975 m deep in the Tasman Sea could be attributed to gas inclusions with r_{es} between 0.5 and 2.1 mm.

VI. CONCLUSIONS

In this work, we have demonstrated a physics-based approach to broadband target spectra classification. Using model-generated training data, we are able to coarsely classify target spectra from gas-bearing and fluid-like organisms based on their scattering physics with an *F1* score of 0.90. Analysis of the classified target spectra provides insight into the size and distribution of organisms in the mesopelagic zone and how they contribute to volume backscatter measured by shipboard echosounders. For the deployment analyzed in this work, gas-bearing organisms account for 98.9% of 38 kHz volume backscatter between 200 and 800 m depth, and 70.4% can be attributed to organisms with inferred equivalent spherical radii between 0.5 and 1.2 mm.

ACKNOWLEDGMENTS

The authors would like to thank Bob Pettit, Kaitlyn Tradd, Peter Weibe, Joel Llopiz, and Noa Randall from the Woods Hole Oceanographic Institution (WHOI) and Michael Jech from the National Oceanic and Atmospheric Administration (NOAA) for their contributions to Deep-See development and deployment. Development of Deep-See was funded by the National Science Foundation’s Major Research Initiative Program, and field work and testing were funded by NOAA. This project was also supported by the WHOI Audacious/TED project, and E.M. was supported by a WHOI postdoctoral scholarship.

¹For more information, see https://support.echoview.com/WebHelp/Reference/Algorithms/Operators/Single_target_detection/Single_target_detection_-_wideband.htm.

²For more information, see https://support.echoview.com/WebHelp/Reference/Algorithms/Fish_tracking_module/Fish_tracking_algorithms.htm.

³See supplementary material at <https://www.scitation.org/doi/suppl/10.1121/10.0005114> for supplementary figures and captions.

Bassett, C., Cotter, E., Stanton, T., and Lavery, A. (2020). “Frequency- and depth-dependent target strength measurements of individual mesopelagic scatterers,” *J. Acoust. Soc. Am.* **148**(2), EL153–EL158.

Bassett, C., De Robertis, A., and Wilson, C. (2018). “Broadband echosounder measurements of the frequency response of fishes and euphausiids in the Gulf of Alaska,” *ICES J. Mar. Sci.* **75**(3), 1131–1142.

Benoit-Bird, K., and Au, W. (2001). “Target strength measurements of Hawaiian mesopelagic boundary community animals,” *J. Acoust. Soc. Am.* **110**(2), 812–819.

Benoit-Bird, K., and Waluk, C. (2020). “Exploring the promise of broadband fisheries echosounders for species discrimination with quantitative

- assessment of data processing effects," *J. Acoust. Soc. Am.* **147**(1), 411–427.
- Blackman, S. (1986). *Multiple-Target Tracking with Radar Applications* (Artech House, Norwood, MA).
- Blanluet, A., Doray, M., Berger, L., Romagnan, J., Le Bouffant, N., Lehuta, S., and Petitgas, P. (2019). "Characterization of sound scattering layers in the Bay of Biscay using broadband acoustics, nets and video," *PLoS One* **14**(10), e0223618.
- Brautaset, O., Waldeland, A. U., Johnsen, E., Malde, K., Eikvil, L., Salberg, A.-B., and Handegard, N. O. (2020). "Acoustic classification in multifrequency echosounder data using deep convolutional neural networks," *ICES J. Mar. Sci.* **77**(4), 1391–1400.
- Breiman, L. (2001). "Random forests," *Mach. Learn.* **45**, 5–32.
- Chu, D., and Stanton, T. (1998). "Application of pulse compression techniques to broadband acoustic scattering by live individual zooplankton," *J. Acoust. Soc. Am.* **104**(1), 39–55.
- Davison, P., Koslow, A., and Kloser, R. (2015). "Acoustic biomass estimation of mesopelagic fish: Backscattering from individuals, populations, and communities," *ICES J. Mar. Sci.* **72**(5), 1413–1424.
- Demer, D., Andersen, L., Bassett, C., Berger, L., Chu, D., Condiotty, J., Cutter, G., Hutton, B., Korneliusson, R., Le Bouffant, N., Macaulay, G., Michaels, W., Murfin, D., Pobitzer, A., Renfree, J., Sessions, T., Stierhoff, K., and Thompson, C. (2017). "Evaluation of a wideband echosounder for fisheries and marine ecosystem science," ICES Cooperative Research Report 336 (ICES, Copenhagen, Denmark).
- Demer, D. L. B., Bernasconi, M., Bethke, E., Boswell, L., Chu, D., Domokos, R., Dunford, A., Fässler, S., Gauthier, S., Hufnagle, L., Jech, J., Bouffant, N., Lebourges-Dhaussy, A., Lurton, X., Macaulay, G., Perrot, Y., Ryan, T., Parker-Stetter, S., Stienessen, S., Weber, T., and Williamson, N. (2015). "Calibration of acoustic instruments," ICES Cooperative Research Report 326 (ICES, Copenhagen, Denmark).
- Dornan, T., Fielding, S., Saunders, R., and Genner, M. (2019). "Swimbladder morphology masks southern ocean mesopelagic fish biomass," *Proc. R. Soc. B: Biol. Sci.* **286**(1903), 20190353.
- Dudani, S. (1976). "The distance-weighted k-nearest-neighbor rule," *IEEE Trans. Syst. Manuf. Cybernet. SMC*-6(4), 325–327.
- Escobar-Flores, P., O'Driscoll, R., Montgomery, J., Ladroit, Y., and Jendersie, S. (2020). "Estimates of density of mesopelagic fish in the Southern Ocean derived from bulk acoustic data collected by ships of opportunity," *Polar Biol.* **43**(1), 43–61.
- Fix, E., and Hodges, J. (1951). "Discriminatory analysis, nonparametric discrimination: Consistency properties," Technical report 4 (USAF School of Aviation Medicine, Randolph Field, TX); available at <https://www.jstor.org/stable/1403797?seq=1>.
- Foote, K. (1980). "Importance of the swimbladder in acoustic scattering by fish: A comparison of gadoid and mackerel target strengths," *J. Acoust. Soc. Am.* **67**, 2084–2089.
- Francois, R., and Garrison, G. (1982a). "Sound absorption based on ocean measurements: Part I: Pure water and magnesium sulfate contributions," *J. Acoust. Soc. Am.* **72**(3), 896–907.
- Francois, R., and Garrison, G. (1982b). "Sound absorption based on ocean measurements. Part II: Boric acid contribution and equation for total absorption," *J. Acoust. Soc. Am.* **72**(6), 1879–1890.
- Irigoin, X., Klevjer, T., Røstad, A., Martinez, U., Boyra, G., Acuña, J., Bode, A., Echevarria, F., Gonzalez-Gordillo, J., Hernandez-Leon, S., Agusti, S., Aksnes, D., Duarte, C., and Kaartvedt, S. (2014). "Large mesopelagic fishes biomass and trophic efficiency in the open ocean," *Nat. Commun.* **5**, 3271.
- Jones, B., Lavery, A., and Stanton, T. (2009). "Use of the distorted wave Born approximation to predict scattering by inhomogeneous objects: Application to squid," *J. Acoust. Soc. Am.* **125**(1), 73–88.
- Kaartvedt, S., Staby, A., and Aksnes, D. (2012). "Efficient trawl avoidance by mesopelagic fishes causes large underestimation of their biomass," *Mar. Ecol. Prog. Ser.* **456**, 1–6.
- Khodabandeloo, B., Agersted, M., Klevjer, T., Macaulay, G., and Melle, W. (2021). "Estimating target strength and physical characteristics of gas-bearing mesopelagic fish from wideband *in situ* echoes using a viscous-elastic scattering model," *J. Acoust. Soc. Am.* **149**(1), 673–691.
- Kloser, R., Ryan, R., Keith, G., and Gershwin, L. (2016). "Deep-scattering layer, gas-bladder density, and size estimates using a two-frequency acoustic and optical probe," *ICES J. Mar. Sci.* **73**(8), 2037–2048.
- Kubilius, R., Macaulay, G., and Ona, E. (2020). "Remote sizing of fish-like targets using broadband acoustics," *Fisheries Res.* **228**, 105568.
- Lavery, A., Bassett, C., Lawson, G., and Jech, J. (2017). "Exploiting signal processing approaches for broadband echosounders," *ICES J. Mar. Sci.* **74**(8), 2262–2275.
- Lavery, A., Wiebe, P., Stanton, T., Lawson, G., Benfield, M., and Copley, N. (2007). "Determining dominant scatterers of sound in mixed zooplankton populations," *J. Acoust. Soc. Am.* **122**(6), 3304–3326.
- Love, R. (1978). "Resonant acoustic scattering by swimbladder-bearing fish," *J. Acoust. Soc. Am.* **64**(2), 571–580.
- MacLennan, D. N., and Holliday, D. V. (1996). "Fisheries and plankton acoustics: Past, present, and future," *ICES J. Mar. Sci.* **53**(2), 513–516.
- Medwin, H., and Clay, C. (1998). *Fundamental of Acoustical Oceanography* (Academic Press, Boston, MA), pp. 138–141.
- Nero, R. W., Thompson, C. H., and Jech, J. M. (2004). "In situ acoustic estimates of the swimbladder volume of Atlantic herring (*Clupea harengus*)," *ICES J. Mar. Sci.* **61**(3), 323–337.
- Peña, M. (2018). "Robust clustering methodology for multi-frequency acoustic data: A review of standardization, initialization and cluster geometry," *Fisheries Res.* **200**, 49–60.
- Proud, R., Handegard, N., Kloser, R., Cox, M., and Brierley, A. (2019). "From siphonophores to deep scattering layers: Uncertainty ranges for the estimation of global mesopelagic fish biomass," *ICES J. Mar. Sci.* **76**(3), 718–733.
- Reeder, D., Jech, J., and Stanton, T. (2004). "Broadband acoustic backscatter and high-resolution morphology of fish: Measurement and modeling," *J. Acoust. Soc. Am.* **116**(2), 747–761.
- Ross, T., Keister, J. E., and Lara-Lopez, A. (2013). "On the use of high-frequency broadband sonar to classify biological scattering layers from a cabled observatory in Saanich Inlet, British Columbia," *Methods Oceanogr.* **5**, 19–38.
- Scoulding, B., Chu, D., Ona, E., and Fernandes, P. (2015). "Target strengths of two abundant mesopelagic fish species," *J. Acoust. Soc. Am.* **137**(2), 989–1000.
- Simmonds, E., and MacLennan, D. (2005). *Fisheries Acoustics*, 2nd ed. (Blackwell Science LTD, Oxford, UK).
- Stanton, T. K., and Chu, D. (2000). "Review and recommendations for the modelling of acoustic scattering by fluid-like elongated zooplankton: Euphausiids and copepods," *ICES J. Mar. Sci.* **57**(4), 793–807.
- Stanton, T. K., and Chu, D. (2008). "Calibration of broadband active acoustic systems using a single standard spherical target," *J. Acoust. Soc. Am.* **124**(1), 128–136.
- Stanton, T. K., Chu, D., Jech, J., and Irish, J. (2010). "New broadband methods for resonance classification and high-resolution imagery of fish with swimbladders using a modified commercial broadband echosounder," *ICES J. Mar. Sci.* **67**(2), 365–378.
- Stanton, T. K., Wiebe, P. H., Chu, D., Benfield, M. C., Scanlon, L., Martin, L., and Eastwood, R. L. (1994). "On acoustic estimates of zooplankton biomass," *ICES J. Mar. Sci.* **51**(4), 505–512.
- Turin, G. (1960). "An introduction to matched filters," *IRE Trans. Inf. Theory* **6**(3), 311–329.
- Underwood, M. J., García-Seoane, E., Klevjer, T. A., Macaulay, G., and Melle, W. (2020). "An acoustic method to observe the distribution and behaviour of mesopelagic organisms in front of a trawl," *Deep Sea Res. Part II* **180**, 104873.
- Verma, A., Kloser, R., and Duncan, A. (2017). "Potential use of broadband acoustic methods for micronekton classification," *Acoust. Australia* **45**(2), 353–361.
- Weston, D. (1967). "Sound propagation in the presence of a bladder fish," in *Underwater Acoustics* (Plenum Press, New York, NY), pp. 55–88.
- Yasuma, H., Sawada, K., Ohshima, T., Miyashita, K., and Aoki, I. (2003). "Target strength of mesopelagic lanternfishes (family Myctophidae) based on swimbladder morphology," *ICES J. Mar. Sci.* **60**, 584–591.