Seabed classification from merchant ship-radiated noise using a physics-based ensemble of deep learning algorithms

Christian D. Escobar-Amado, Tracianne B. Neilsen, Jhon A. Castro-Correa, David F. Van Komen, Mohsen Badiey, David P. Knobles, and William S. Hodgkiss

Citation: The Journal of the Acoustical Society of America **150**, 1434 (2021); doi: 10.1121/10.0005936
View online: https://doi.org/10.1121/10.0005936
View Table of Contents: https://asa.scitation.org/toc/jas/150/2
Published by the Acoustical Society of America

JASA ARTICLE

Seabed classification from merchant ship-radiated noise using a physics-based ensemble of deep learning algorithms

Christian D. Escobar-Amado,[1,a)] Tracianne B. Neilsen,[2] Jhon A. Castro-Correa,[1] David F. Van Komen,[2,b)] Mohsen Badiey,[1] David P. Knobles,[3] and William S. Hodgkiss[4]

[1]*Department of Electrical Engineering, University of Delaware, Newark, Delaware 19716, USA*

[2]*Department of Physics and Astronomy, Brigham Young University, Provo, Utah 84602, USA*

[3]*Knobles Scientific and Analysis, LLC, Austin, Texas, USA*

[4]*Marine Physical Laboratory, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California 92093, USA*

ABSTRACT:

Merchant ship-radiated noise, recorded on a single receiver in the 360–1100 Hz frequency band over 20 min, is employed for seabed classification using an ensemble of deep learning (DL) algorithms. Five different convolutional neural network architectures and one residual neural network are trained on synthetic data generated using 34 seabed types, which span from soft-muddy to hard-sandy environments. The accuracy of all of the networks using fivefold cross-validation was above 97%. Furthermore, the impact of the sound speed and water depth mismatch on the predictions is evaluated using five simulated test cases, where the deeper and more complex architectures proved to be more robust against this variability. In addition, to assess the generalizability performance of the ensemble DL, the networks were tested on data measured on three vertical line arrays in the Seabed Characterization Experiment in 2017, where 94% of the predictions indicated that mud over sand environments inferred in previous geoacoustic inversions for the same area were the most likely sediments. This work presents evidence that the ensemble of DL algorithms has learned how the signature of the sediments is encoded in the ship-radiated noise, providing a unified classification result when tested on data collected at-sea. © *2021 Acoustical Society of America*.

https://doi.org/10.1121/10.0005936

(Received 27 April 2021; revised 30 July 2021; accepted 30 July 2021; published online 25 August 2021)

[Editor: Stan E. Dosso] Pages: 1434–1447

I. INTRODUCTION

Merchant ships are an active moving source of broadband noise in the ocean environment and provide frequency and spatial diversity, which can be recorded over great distances. For this reason, these types of signals are excellent candidates for passive geoacoustic inversions in shallow water.[1] These sources of opportunity, also known as ships of opportunity (SOOs), usually travel in a straight line at a constant speed, generating a characteristic signal that presents interference patterns, referred to as striations. The average source spectra of these broadband radiators have been modeled by Wales and Heitmeyer[2] over a 30–1200 Hz frequency band.

Recently, geoacoustic inversions using ship-radiated noise have been an active area of research in which seabed parameters, such as compressional sound speed and layer thickness, are inferred using SOO spectrograms.[3–7] Bayesian approaches[4,8,9] and search algorithms[5,6,10,11] are commonly used for finding the optimal model for these types of geoacoustic inversions. With the rich information content that the broadband SOO noise contains about the waveguide, researchers have used frequencies in the lower bands of 20–140 Hz, as well as in the higher bands of 1700–3300 Hz, for parameter estimations.[4,12]

Machine and deep learning (DL) techniques can also be used for inferring geoacoustic properties in the ocean environment.[13,14] Some approaches include neural and statistical classifiers,[15] artificial neural networks,[16] generalized additive models,[17] Bayesian learning,[18] and convolutional neural networks (CNNs).[19,20] Different types of sound sources have been used for this purpose; some examples are continuous wave signals,[16,19,21] linear frequency modulated signals,[17] and explosive charges.[20]

Another approach for inferring geoacoustic properties is by means of the signatures of the sediment encoded in the propagated sound for classifying the signals from several canonical seabed types. Efforts on using DL techniques for seabed classification have been performed with classes that span from soft-mud to hard-sand environments, using a few canonical sediments. Van Komen *et al.*[20,22] used CNNs for simultaneously predicting source localization parameters and classifying the sediment from four canonical seabed types (deep mud, mud over sand, sandy silt, and sand) using explosive charges in the 5–2500 Hz band and SOO spectrograms in the 300–1500 Hz band. Using the same four canonical environments, Neilsen *et al.*[19] implemented several CNNs for simultaneously performing source localization predictions and seabed classification with towed tonal sources in the mid frequency band. Frederick *et al.*[21] used several machine learning algorithms for classifying continuous

[a)]Electronic mail: aescobarc@udel.edu, ORCID: 0000-0003-2907-7311.
[b)]ORCID: 0000-0003-0610-0806.

 0001-4966/2021/150(2)/1434/14/$30.00

https://doi.org/10.1121/10.0005936

wave signals from clay, silt, sand, and gravel sediments and concluded that a broader spectrum of environments will help with the generalization performance of the models. A similar approach for seabed classification is followed in this work, but it expands the number of seabed types to 34 classes, which are found in the literature from different areas of the world that span from soft-muddy to hard-sandy sediments.

In this paper, an ensemble of DL algorithms is applied to SOO spectrograms in the 360–1100 Hz frequency band over 20 min to classify the sediment using 34 seabed types. For this study, five CNNs and one residual network (ResNet) are considered for the classification task. Due to the lack of measured SOO data, DL algorithms are trained using simulated data (from a physics-based model) with environment and source track variability that contain sufficient information content for the networks to learn. To test the ability of the networks to correctly classify samples not used during training, fivefold cross-validation (CV) results are reported and show validation accuracy above 97% for all of the DL architectures. In addition, several simulated test cases were created to evaluate the impact of the sound speed profile (SSP) mismatch on the generalizability of the networks.

To further assess the generalization performance of the DL algorithms, the trained networks were applied to 69 measured SOO spectrograms collected during the Seabed Characterization Experiment in 2017 (SBCEX 2017) in the New England Mud Patch Area. The merchant ships were identified using the automatic identification system (AIS) data retrieved from the Marine Cadastre.[23] The seabed classification across all of the CNNs and ResNet—the ensemble of DL algorithms—consistently point to the same seabed types that correspond to muddy environments, which were reported in previous geoacoustic inversions at the same area. These results show the ability of DL algorithms to distinguish among a wide variety of seabed types using broadband SOO spectrograms.

II. MEASURED DATA

A. Experiment

The data used in this paper were collected by three vertical line arrays (VLAs), which were deployed in the SBCEX 2017 in the New England Mud Patch area, as shown in Fig. 1. VLA1-MPL and VLA2-MPL were composed of 16 hydrophones each and were deployed by the Marine Physical Laboratory of the Scripps Institution of Oceanography (MPL). VLA-UD was deployed by the University of Delaware (UD) and had a total of 8 hydrophones. For this paper, only the hydrophone located approximately in the center of the water column at a depth of 33 m above the seafloor is used. The water depth in the area was about 75 m with some changes in bathymetry between 72 and 78 m. The details about the positions and recording times for each VLA are listed in Table I. VLA-UD was continuously recording from March 7, 2017—which corresponds to Julian day (JD) 66—until April 8, 2017 (JD 98) for a total of 32.08 days. VLA1-MPL and VLA2-MPL,

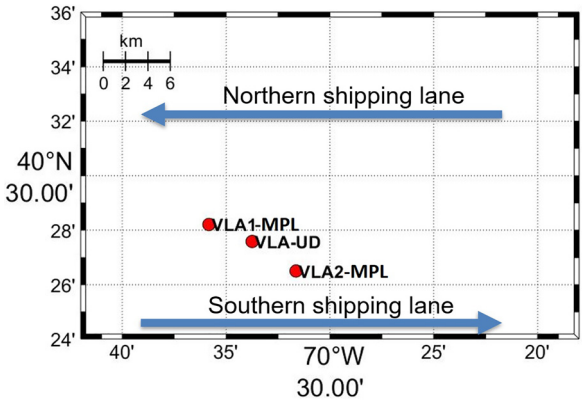

FIG. 1. (Color online) The positions of VLA1-MPL, VLA2-MPL, and VLA-UD during SBCEX 2017 along with arrows indicating the shipping lanes.

however, followed a more intermittent schedule for a total recording time of 9.1 and 7.5 days, respectively.

B. Identified ships

As shown in Fig. 1, the SBCEX 2017 area was near two shipping lanes. The merchant vessels in the area in which the VLAs were recording were identified using AIS data retrieved from the Marine Cadastre.[23] Northern and southern shipping lanes were active for the duration of the experiment and several merchant ships were identified passing by close to the three receivers. One example of a ship recorded simultaneously at the three VLAs is shown in Fig. 2(a), where the vessel Viking Bravery was moving from the southern to the northern shipping lane parallel to the arrays. Each blue marker on the map in Fig. 2 represents a data sample from the AIS, where parameters such as course over ground, ship speed, latitude, longitude, and draft were reported. The speed of this ship remained constant at 14.66 kn. The closest point of approach (CPA) range is computed based on the location of the vessel, i.e., its reported latitude and longitude. The CPAs for Viking Bravery were 3.30, 3.09, and 2.87 km for VLA1-MPL, VLA2-MPL, and VLA-UD, respectively. Figures 2(b)–2(d) show the spectrograms of the Viking Bravery received at three locations where a *nested* set of hyperbolas (also known as striations) with the vertices occurring at CPA are observed; the spectrograms exhibit well-defined striations, which span a broad frequency band. These spectrograms have a time spacing of 5 s with a frequency

TABLE I. Information about the receivers.

Parameter	VLA1-MPL	VLA2-MPL	VLA-UD
Latitude	40.470 N	40.442 N	40.460 N
Longitude	70.597 W	70.527 W	70.563 W
Starting date	March 22 (JD81)	March 23 (JD82)	March 7 (JD66)
Ending date	April 3 (JD93)	April 2 (JD92)	April 8 (JD98)
Recorded time	9.1 days	7.5 days	32.1 days

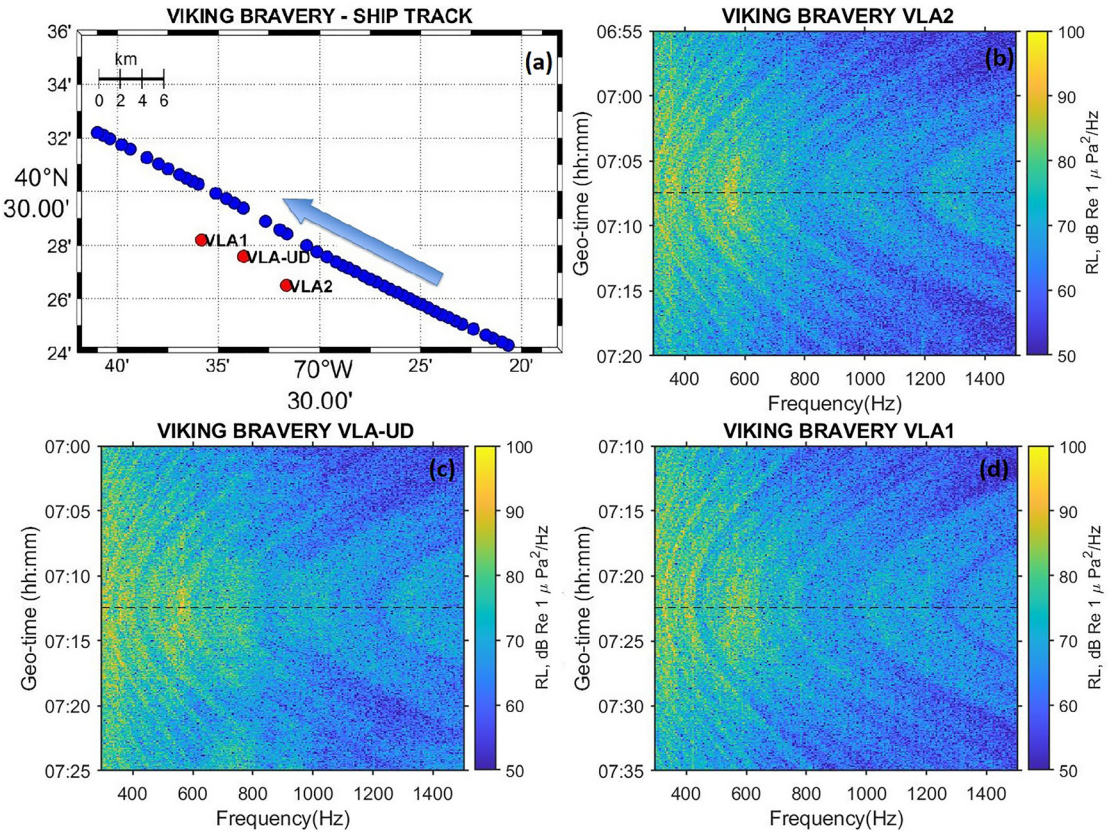

FIG. 2. (Color online) (a) The map indicating the position of the ship retrieved from the AIS data. The arrow indicates the direction of the vessel Viking Bravery, which was recorded at (b) VLA2-MPL, (c) VLA-UD, and (d) VLA1-MPL in the SBCEX 2017 on March 30, 2017 (JD 90). The spectrograms are shown for the three receivers as a function of the geo-time (UTC) and frequency.

step of 6.1 Hz. The black dotted lines on top of the spectrograms indicate the CPA positions, which occur at different geo-times for the three receivers.

All of the ships recorded on one or more of the VLAs are listed in Table II. Location parameters, inferred from the AIS data, such as CPA with respect to each VLA and mean ship speed (v_{sh}), are reported as well. The symbols next to the CPA parameter for each VLA indicate a qualitative description of the signal measured at that location. The symbol "+" represents a SOO spectrogram with clearly identifiable striation patterns, whereas a "−" indicates that the striations are visually present but with a low signal-to-noise ratio (SNR). In addition, an "*" signifies that the data sample contains a loud noise event that is different than the broadband noise generated by the ship in question. In total, 69 ship events were found at the 3 receivers from 51 identified vessels. The CPA ranges vary from 0.64 to 13.30 km, and the ship speeds vary from 8 to 20 kn.

III. METHODOLOGY

To train a supervised machine or DL model, a large labeled database is needed that has sufficient information content to capture the variety found in the testing scenarios. However, in ocean acoustics, at-sea labeled data are scarce because of the costliness of obtaining such measurements. In this work, 69 data samples have been extracted—an insufficient number for training a DL algorithm. Our approach is to simulate the training data with the range-independent, normal-mode propagation model, ORCA[24] by using 34 representative sediments found in the literature, which span from muddy to sandy environments. The simulated source-receiver configuration is similar to the one used for the measured data, where a merchant ship is moving with constant speed in a straight line relative to a receiver. This synthetic dataset is used for training and validating an ensemble of six DL algorithms, which are used as a composite classifier to determine the most likely seabed type for each spectrogram. Then, the trained networks are applied to five simulated test cases to evaluate the impact of the sound speed and water depth mismatch. Finally, the ensemble of trained networks is applied to the measured SOO spectrograms to assess the generalization ability of the DL algorithms for classifying the sediment using real-world data.

A. Data simulation

The scope of this work is to evaluate the ability of DL algorithms to distinguish among different types of seabeds when *a priori* information about the sediment is not available. To serve this purpose, the variability present in the measured samples is represented by including enough environmental variability in the simulated training dataset.

https://doi.org/10.1121/10.0005936

TABLE II. The list of ships recorded during the SBCEX 2017. The CPA ranges (in km) with respect to VLA1-MPL, VLA2-MPL, and VLA-UD are labeled as R-1, R-2, and R-UD, respectively. The speed of the ship, v_{sh} (in kn), is assumed to be constant. Symbols next to the CPA indicate a qualitative description of the signal. Superscripts "+" and "−" indicate that the SOO spectrogram has high or low SNR, respectively. Subscript "∗" represents that a loud noise event, different than the broadband noise generated by the ship, was present.

#	Ship name	R-1	R-2	R-UD	v_{sh}	JD	#	Ship name	R-1	R-2	R-UD	v_{sh}	JD
1	MATAQUITO	—	—	9.5^{+}_{*}	19.4	67	27	MSC GISELLE	—	—	6.5^{+}	18.3	85
2	ALICE OLDENDORFF	—	—	0.6^{+}	8.3	67	28	FEDOR	8.6^{-}	—	—	11.6	86
3	PAGNA	—	—	13.3^{+}_{*}	17.5	69	29	BRITISH TRANQUILLITY	8.5^{-}	12.0^{-}	9.8^{-}_{*}	13.5	86
4	JIA SHENG SHAN	—	—	9.1^{+}_{*}	11.2	69	30	CPO BALTIMORE	9.3^{+}_{*}	—	—	14.7	86
5	EVER LIVING	—	—	5.8^{+}	18.4	72	31	TORM SAONE	8.6^{-}	—	—	13	87
6	ZIM QINGDAO	—	—	8.8^{-}_{*}	11.3	73	32	ARDMORE SEAVANTAGE	—	—	5.0^{+}	15.4	87
7	STI CLAPHAM	—	—	4.6^{+}_{*}	11.8	73	33	LEOPARD	—	—	10.1^{+}	13.2	87
8	NYK RUMINA	—	—	5.0^{+}	19.8	73	34	CMA CGM MOLIERE	—	—	9.8^{-}	15.9	89
9	BARBARA	—	—	5.7^{+}_{*}	19.9	75	35	HAFNIA GREEN	—	2.8^{+}	4.8^{+}_{*}	10.9	89
10	MSC BREMEN	—	—	10.0^{+}_{*}	18.4	76	36	MSC NERISSA	8.7^{+}	—	9.8^{+}	15.6	90
11	OREGON HIGHWAY	—	—	0.7^{+}_{*}	15.3	77	37	VIKING BRAVERY	3.3^{+}	3.1^{+}	2.9^{+}	14.7	90
12	MSC LAUSANNE	—	—	5.5^{-}	12	77	38	MAERSK MATSUYAMA	7.2^{+}	4.6^{+}	6.6^{-}_{*}	11.6	90
13	NYK RIGEL	—	—	5.2^{+}_{*}	19.9	80	39	TOMBARRA	6.4^{+}	3.2^{+}	—	16.3	90
14	ZIM SHANGHAI	—	—	5.3^{+}_{*}	17.9	80	40	ATLANTIC SEA	9.3^{+}	12.7^{+}	10.6^{+}	17.6	90
15	TRANSPORT	—	—	1.5^{+}	8.4	80	41	KAZDANGA	—	1.9^{+}_{*}	—	9.2	92
16	BOW PIONEER	—	—	4.7^{+}_{*}	11.9	80	42	NYK DIANA	8.6^{+}_{*}	—	9.8^{+}_{*}	18.9	93
17	DISCOVERY BAY	—	—	5.9^{+}	13.3	81	43	CHEMICAL PIONEER	8.5^{+}	—	—	16.2	93
18	MSC ESTHI	—	—	11.1^{-}_{*}	17.8	81	44	DENAK VOYAGER	6.7^{+}_{*}	—	5.5^{+}_{*}	10.3	93
19	ATLANTIC CONVEYOR	9.0^{+}	12.2^{+}	10.2^{+}	16.1	83	45	ARCTIC BREEZE	8.2^{+}	—	—	14.3	93
20	MSC ANIELLO	—	3.6^{+}	—	14.3	83	46	PAGANELLA	—	—	4.8^{+}_{*}	14.9	93
21	MSC KALAMATA	5.9^{+}	3.1^{+}	4.9^{+}	16.7	83	47	MSC KOLKATA	—	—	4.7^{+}_{*}	9.6	93
22	CORRIDO	—	$4.0^{(+)}$	—	14.6	84	48	ALICE OLDENDORFF	—	—	10.3^{+}	10.7	94
23	YM UNANIMITY	—	3.8^{+}	—	9.1	84	49	MSC AMERICA	—	—	6.6^{+}_{*}	16	95
24	MINERVA ZOE	8.6^{-}	—	—	12.3	85	50	CSCL AMERICA	—	—	10^{+}_{*}	21.2	96
25	BBC TENNESSEE	7.4^{-}_{*}	4.2^{+}_{*}	—	8.3	85	51	STEALTH BERANA	—	—	5.2^{+}_{*}	13.7	96
26	CHEM VENUS	9.3^{-}	—	—	12.9	85							

A key question for seabed classification efforts is how to divide the wide variety of seafloors in the shallow oceans into different classes. In prior work,[15,19–22] four or five distinct seabed classes were used to represent the different, general types of sediments. However, a broader spectrum of environments will help with the generalization performance of the DL algorithms.[20–22] To expand the catalog of seabed classes, a measure of the acoustic similarity is needed to help select the seabed classes that are acoustically distinct enough to be learned by the classifier. The 34 seabed classes used in this work (Table III) were reported in peer-reviewed literature on shallow ocean studies. These 34 seabed classes are selected because the transmission losses modeled over the frequency (one-third octave band center frequencies 250–1250 Hz and ranges 1–15 km) have a Pearson correlation of 0.8 or less. (See Ref. 25 for details on the method.) This measure of acoustic similarity helps guide the choice of a seabed catalog that captures the acoustic variability likely to be found in the measured data.

The 34 seabeds are listed in Table III along with the sound speed at the top of the sediment and reference from which the seabed parameterization was obtained. The seabed label names with the suffix "_sbc" were inferred in previous studies for the SBCEX 2017 in the New England Mud Patch Area; a detailed summary about these geoacoustic inversions can be found in Wilson *et al.*[26] A deep mud environment studied in the Gulf of Mexico[27] is also included (#7) among the possible classes. The seabed names with the suffix "_35m" were modeled using the nominal sound speed, density, and attenuation parameters found in

TABLE III. The list of the 34 sediments used for simulation. Column "#" corresponds to the seabed number used for identification in the results. c_1^{top} is the value of the sound speed at the top of the first layer reported in the paper specified in the reference (Ref) column.

#	Seabed label	c_1^{top}	Ref	#	Seabed label	c_1^{top}	Ref
0	Lin_sbc	1387	39	17	Dahl2020_sbc	1479	4
1	Knobles_sbc	1436	40	18	Brown_sbc	1479	41
2	Dahl_sbc	1445	42	19	Michalopoulou_sbc	1491	43
3	Belcourt_2_sbc	1445	44	20	Clay_35m	1500	28
4	Potty_sbc	1452	45	21	Tollefsen_sbc	1508	46
5	Wan_2_sbc	1452	47	22	Malta_a	1510	30
6	Belcourt_1_sbc	1453	44	23	West_Florida	1540	29
7	Deep_mud	1454	27	24	Korean_c	1553	31
8	Barclay_2_sbc	1455	48	25	Korean_a	1558	32
9	Barclay_3_sbc	1459	48	26	Korean_k	1567	33
10	Bonnel_1_sbc	1464	49	27	Silt_35m	1575	28
11	Wan_1_sbc	1467	50	28	Sand	1592	38
12	Mud_over_sand_sbc	1469	40	29	Sandy_silt	1592	34
13	Barclay_1_sbc	1470	48	30	Knobles14j	1650	35
14	Mud_35m	1470	28	31	NewJersey_a	1704	36
15	Bonnel_2_sbc	1471	51	32	Korean_h	1757	37
16	Bonnel_3_sbc	1474	51	33	Gravel_35m	1800	28

Jensen *et al.*[28] for mud (#14), clay (#20), silt (#27), and gravel (#33) environments to define a 35 m-thick layer over a hard-sand half-space. Harder sediments from different areas of the world are also considered, such as the ones studied in the West Florida sand shelf[29] (#23), Malta Plateau in the Mediterranean Sea[30] (#22), Mid-eastern Yellow Sea[31] (#24), Northern East China Sea[32] (#25), Jinhae Bay—off the coast of Southeast Korea[33] (#26), New England in the Middle Atlantic Bight[34] (#28), New Jersey Continental Shelf (#30, #31),[35,36] and Western Continental Margin of the East Sea[37] (#32). The conclusion of a study on several sandy sea bottoms by Zhou *et al.*[38] is also included as seabed type #28.

These 34 seabed classes have been sorted based on the sound speed at the top of the first layer (c_1^{top}). The values of c_1^{top} are the values reported in the papers and have not been modified based on the water sound speed. The sediments numbered 0–15 have a sound speed ratio less than one, whereas the sediments numbered 16–33 have a ratio greater than one. This ordering allows us to have a representative sorted list that goes from softer (mud) to harder (sand) sediments.

Each seabed type has a different layer structure that goes from a simple one-layer sediment over a half space to a more complicated environment, which is composed of several thin layers. Eight examples of the sediment layering are shown in Fig. 3. The parameters involved in the modeling are the layer thickness and density (g/cm^3), attenuation (db/m-kHz), and sound speed (m/s) at the top and bottom of each layer. Seabeds #7, #12, #29, and #28 correspond to the four canonical environments used in Refs. 19, 20, and 22 for deep mud, mud over sand, sandy silt, and sandy sediments, respectively. The layering structure for these four seabed classes can be found in Ref. 19.

The environmental variability in the ocean includes not only the seafloor but also the water column sound speed and depth. For the training dataset, ten SSPs measured during SBCEX 2017, shown in Fig. 4, are used with each of the 34 seabeds for a total of 340 different environments. The SSPs measured *in situ* are shown in Fig. 4 as solid black lines with bottom sound speeds from 1470.5 to 1472.4 m/s, which are similar to the profiles used in the previous geo-acoustic inversions in the SBCEX 2017 that were on the order of 1469.6 ± 2.24 m/s. The water depth used for simulating the training data was 75 m with the receiver position at 33 m above the seafloor.

The range of source parameters used to generate the synthetic training dataset was selected to account for the full range of values covered by the measured data samples (Table II). To make sure that the source parameter space is completely sampled, some labels for the CPA range and ship speed (v_{sh}) are held fixed for all of the environments. For each of the 340 different environments, the CPA range has 5 fixed values of 2, 5, 8, 11, and 14 km, whereas v_{sh} has 3 fixed values of 10, 14, and 18 kn. Additional values of the source parameters are randomly drawn from a uniform distribution between the minimum and maximum values in Table IV for each of the 340 environments. All of the

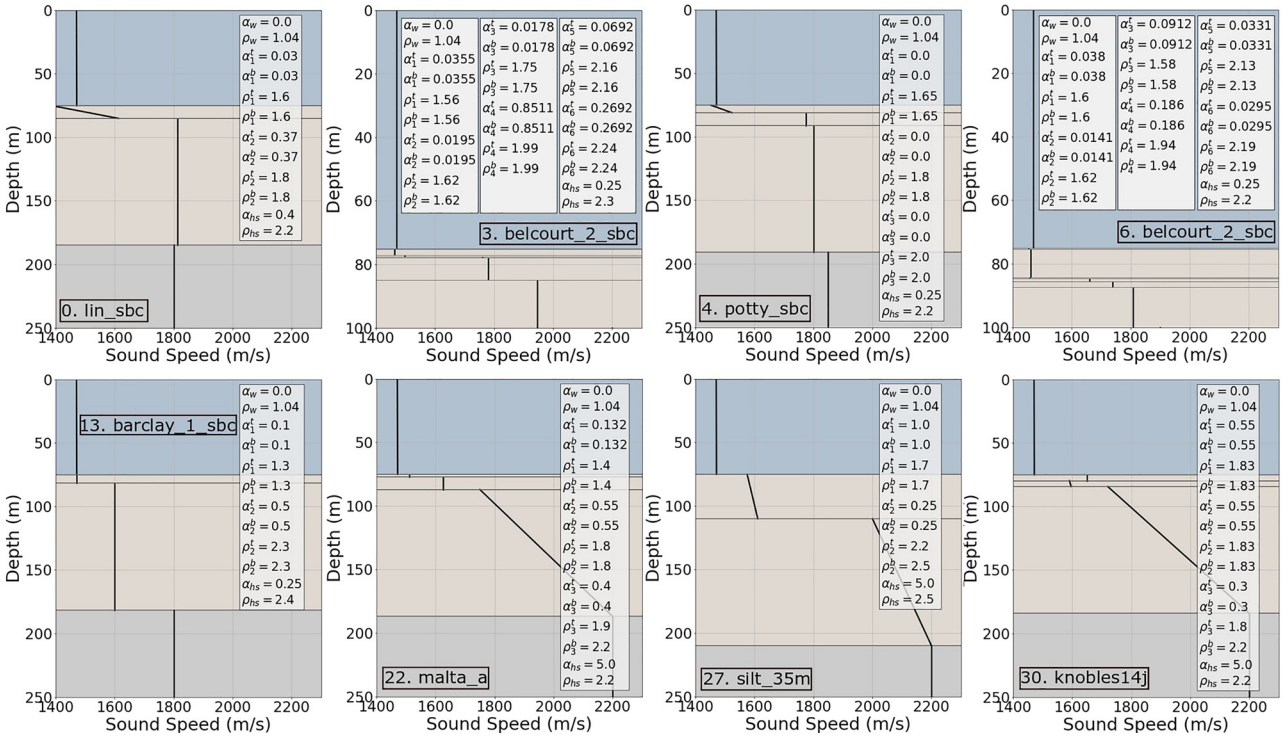

FIG. 3. (Color online) Examples of the sediment profiles used for simulating the training datasets. The compressional SSPs are shown as thick black lines, whereas the legend contains the compressional attenuation (α) and density (ρ) for the water (subscript w), the top (superscript t) and bottom (superscript b) of each layer (subscript $i = 1 - n$, where n is the number of layers), and the half-space (subscript hs).

https://doi.org/10.1121/10.0005936

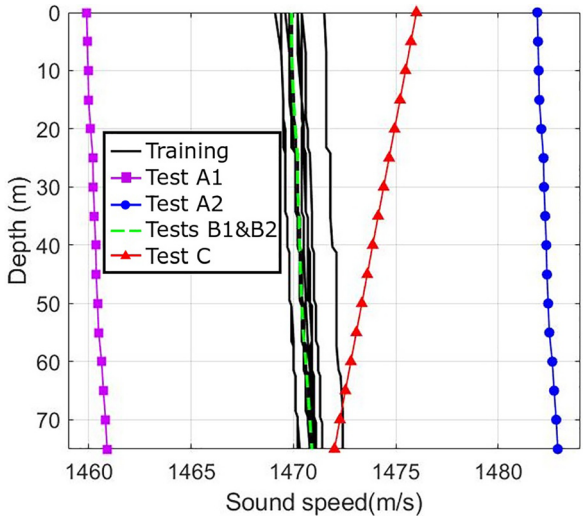

FIG. 4. (Color online) The SSPs used for simulating the training and testing datasets.

combinations of this sampling result in 405 different sets of source parameters per environment; a total of 137 700 data samples are generated for training the DL algorithms.

To analyze the performance of the networks in the face of environmental mismatch, five additional datasets have been simulated with environments not used for training. Tests A1 and A2 were created to evaluate the impact on the SSP mismatch using the same water depth of 75 m but with the SSPs shifted by ± 10 m/s with respect to the range of SSPs used for the training data, as depicted in Fig. 4 by the lines with the square and circle markers. For tests B1 and B2, the water depth was shifted by $+3$ m (78 m) and -3 m (72 m), respectively, using one of the SSPs from the training data. Test C is the most challenging case and was simulated using a single downward refracting SSP at 75 m water depth, as shown in Fig. 4 by the line with triangle markers. The source location parameters used for the five test cases are as shown in Table IV, following the same approach as for the training data. Thus, each synthetic test case is composed of 405 samples for each of the 34 seabed types listed in Table III, yielding a total of 13 770 data samples.

The synthetic SOO spectrograms are generated using the simulated ocean response and a modeled source spectrum. The simulated acoustic channel response is obtained from ORCA, a range-independent, elastic, normal-mode model.[24] The broadband source ship noise spectral density, $S(f)$ in dB, re 1 μPa^2/Hz, is modeled by the Wales and Heitmeyer[2] equation as

$$S(f) = S_0 - 10\log\left(f^{3.594}\right) + 10\log\left(\left(1+(f/340)^2\right)^{0.917}\right), \quad (1)$$

where S_0 represents the baseline source level of the ship in decibels and f is the frequency. The empirical value found was $S_0 = 230$ dB, and they applied this model over the 30–1200 Hz band. However, during the SBCEX 2017, several loud noise events frequently happened below 350 Hz. For this reason, the frequency band of 360–1100 Hz was selected to make sure that the measured and modeled signals have a common well-studied band with a clearly identifiable signal. Examples of the simulated SOO spectrograms for four seabed types (#0, #7, #20, and #28) are shown in Fig. 5 for three CPA ranges of 1, 3, and 8 km with a constant ship speed of 10 kn and source depth of 10 m. The 20-min spectrograms were generated using a 5 s time interval with a frequency spacing of 6.1 Hz, yielding a matrix with 243×123 spectral density levels. The selected time and frequency resolutions allowed us to capture the details in the striations of the SOO samples and keep a well-defined aspect ratio.

TABLE IV. The source parameter selections for the training and testing datasets. Some fixed values for the CPA range and ship speed are used for each environment. Additional source parameters are drawn randomly for each environment. All combinations of these 3 source parameters produces 405 data samples for each environment.

	Minimum value	Maximum value	Number fixed	Number random
CPA (km)	0.5	15	5	10
v_{sh} (kn)	8	20	3	6
Zs (m)	6	12	0	3

B. DL algorithms

To obtain robust predictions, an ensemble of DL models is trained for seabed classification. Having a set of diverse DL architectures will increase the reliability of the predictions and guard against incorrect responses. When a new sample is presented, if all of the networks agree to a unified answer, then the selected class is likely to be correct because each network might have learned different representative patterns during training.

For this study, five different CNN architectures and one residual neural network (ResNet) are implemented. CNNs are a DL tool used for temporally or spatially correlated signals.[13] CNNs are useful for learning representative features in the convolutional layers by sliding filters (or kernels) across the data. It is common to use maximum pooling at the output of such filters to downsample the feature maps by keeping only the largest value in a given moving window. More information on CNNs can be found in Goodfellow *et al.*[52] A ResNet is, in essence, a CNN with skip connections that improve the gradient flow through the network and enables the training of deeper architectures.[53] In the area of ocean acoustics, ResNet algorithms have been recently used by Niu *et al.*[54] for source localization using multi-frequency magnitude-only data. The topologies of the networks, along with data pre-processing and augmentation techniques used for training, are presented in this section.

1. Network architectures

The architectures of the CNNs implemented for this work are summarized in Table V. Each convolutional layer is composed of convolutions that can be followed by

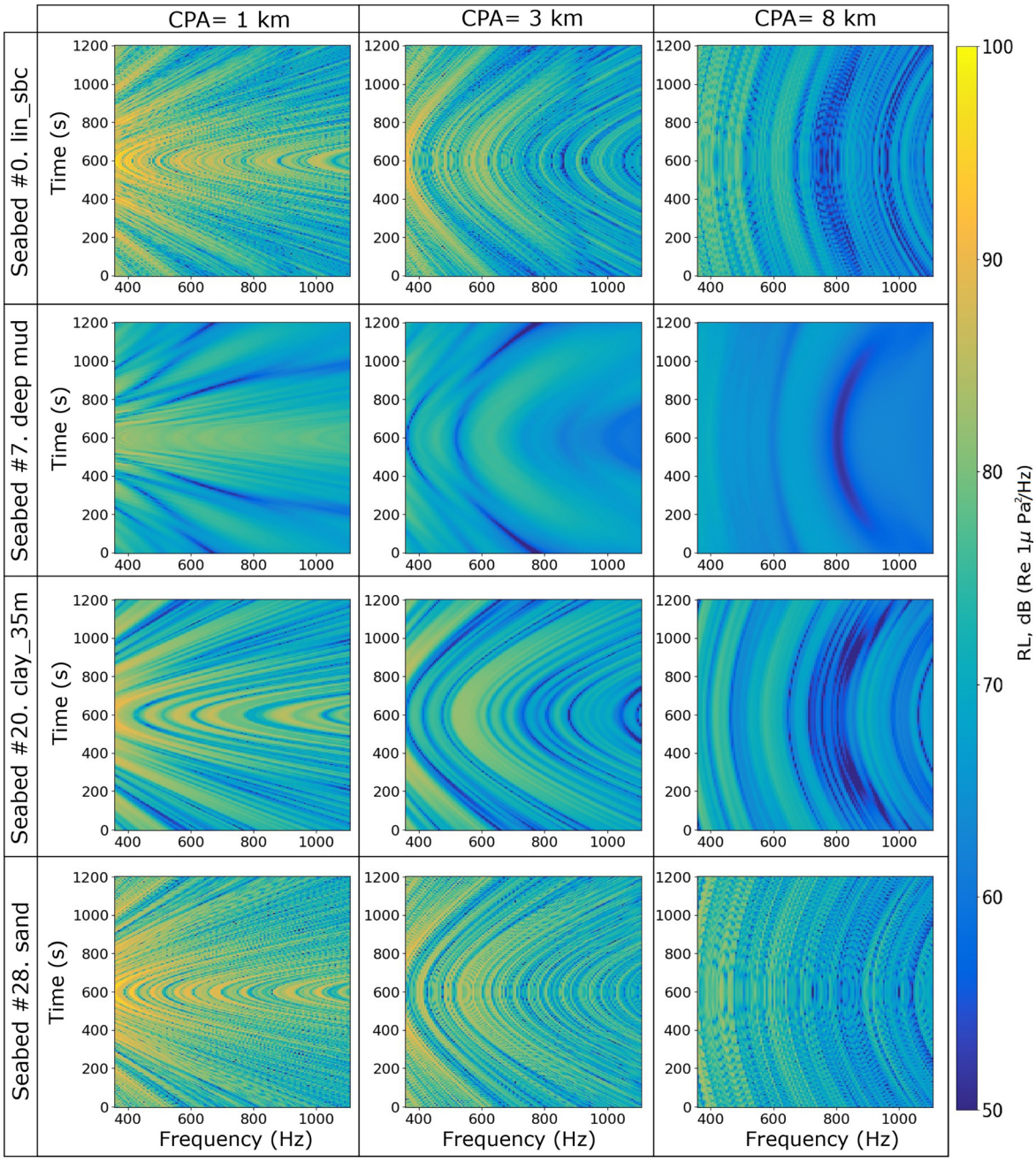

FIG. 5. (Color online) Examples of the synthetic SOO spectrograms for CPA ranges = 1, 3, and 8 km with constant ship speed =10 kn and source depth =10 m. Seabeds 0, 7, 20, and 28 correspond to mud over sand, deep mud, clay, and sandy environments, respectively. The color represents the received level (RL) in dB Re 1 μPa^2/Hz.

maximum pooling to decrease the dimensionality of the network. The output of these convolutional layers is sent to the fully connected (FC) layers, which then assign a score to each of the 34 outputs—one associated with each seabed type. Several approaches for the design of the networks are considered. First, a shallow CNN with only three convolutional layers, with maximum pooling in all of them, and two FC layers were implemented. For this network, named Selkie3 (presented in Ref. 22), data are input in two formats, one-dimensional (1D) and two-dimensional (2D). The 1D input format is structured such that each frequency corresponds to a channel, having a total of 123 channels with 243 points each. The 2D input format is a single channel spectrogram with size 243 × 123.

Two deeper CNNs, Selkie5 and CNN5-UD, with five convolutional layers were implemented using only the 2D input format of the spectrograms. Selkie5, also presented in Ref. 22, has two FC layers, and pooling was not employed. CNN5-UD was inspired by the network implemented in Neilsen *et al.*,[19] where max pooling is applied in the first two layers, and one single FC layer is used at the end. In addition, the well-known CNN topology, AlexNet,[55] is also

https://doi.org/10.1121/10.0005936

TABLE V. The CNN architectures. The parameters for each convolutional layer are presented as (kernel size)(stride)(padding)(number of channel outputs). The value that follows the FC nomenclature corresponds to the number of neurons in that fully connected (FC) layer.

Layer	Selkie3 (1D)	Selkie3 (2D)	Selkie5 (2D)	CNN5-UD (2D)
Convolutional	$(5 \times 1)(1 \times 0)(2 \times 0)(\mathbf{100})$	$(5 \times 5)(1 \times 1)(0 \times 2)(\mathbf{8})$	$(9 \times 9)(2 \times 2)(4 \times 4)(\mathbf{10})$	$(9 \times 5)(1 \times 1)(4 \times 2)(\mathbf{32})$
Maximum pool	$(2 \times 1)(2 \times 0)(1 \times 0)$	$(2 \times 2)(2 \times 2)(1 \times 1)$	—	$(5 \times 3)(3 \times 2)(2 \times 1)$
Convolutional	$(5 \times 1)(1 \times 0)(0 \times 0)(\mathbf{50})$	$(5 \times 5)(1 \times 1)(0 \times 0)(\mathbf{16})$	$(7 \times 7)(2 \times 2)(3 \times 3)(\mathbf{20})$	$(5 \times 3)(1 \times 1)(2 \times 1)(\mathbf{64})$
Maximum pool	$(2 \times 1)(2 \times 0)(1 \times 0)$	$(2 \times 2)(2 \times 2)(1 \times 1)$	—	$(5 \times 3)(3 \times 2)(2 \times 1)$
Convolutional	$(5 \times 1)(1 \times 0)(0 \times 0)(\mathbf{20})$	$(5 \times 5)(2 \times 1)(0 \times 0)(\mathbf{4})$	$(5 \times 5)(2 \times 2)(2 \times 2)(\mathbf{30})$	$(3 \times 3)(2 \times 2)(1 \times 1)(\mathbf{128})$
Maximum pool	$(2 \times 1)(2 \times 0)(1 \times 0)$	$(2 \times 2)(2 \times 2)(1 \times 1)$	—	—
Convolutional	—	—	$(5 \times 5)(2 \times 2)(2 \times 2)(\mathbf{40})$	$(3 \times 3)(2 \times 1)(1 \times 1)(\mathbf{64})$
Convolutional	—	—	$(5 \times 5)(1 \times 1)(0 \times 0)(\mathbf{20})$	$(3 \times 3)(1 \times 1)(1 \times 1)(\mathbf{32})$
	FC - 2000	FC - 2000	FC - 3000	FC - 2000
	FC - 500	FC - 500	FC - 2000	—
		Output (34 classes)		

implemented using only half of the network architecture corresponding to five convolutional layers and is referred to as HalfAlexNet throughout this study. A ResNet with 18 layers, named ResNet18, was implemented following the configuration suggested in He *et al.*[56] Table VI presents the architecture of ResNet18, composed of eight basic blocks presented as *conv**n**_x*. Kernels of size 3×3 were used in all of the layers with an average pooling layer of size 4×4 after the convolutions. Finally, a FC layer is used to perform the classification from the 34 seabed types. The key difference between ResNet and the other CNNs is that "skip connections" connect the beginning of each block such that information from earlier in the network is available at later stages.

The DL algorithms were written in the open-source PyTorch framework.[57] Rectified linear unit (ReLu) activation functions,[58] followed by batch normalization, were employed for training the networks. The AdamW optimizer,[59] a variant of the Adam optimizer[60] with decouple weight decay regularization, is implemented using the cosine annealing learning rate procedure described in Van Komen *et al.*[20]

2. Data pre-processing

Whereas the spectral shape of the broadband noise from the transiting ships is captured by Eq. (1), the exact spectral levels depend on the detailed characteristics of each ship. This difference is represented by the S_0 in Eq. (1). To mitigate the effects of unknown S_0 values for the ships listed in Table II, the data samples are normalized individually based on their standard deviation for both measured and simulated data. For each spectrogram, X, the standard deviation (σ) is computed as

$$\sigma = \sqrt{\frac{1}{KL}\sum_{i=1}^{K}\sum_{j=1}^{L}\left(X(t_i,f_j) - \bar{X}\right)^2}, \tag{2}$$

where $X(t_i,f_j)$ is the spectral density re 1 μPa^2/Hz in linear units of the spectrogram at time t_i and frequency f_j; K and L are the numbers of times and frequencies, respectively; and $\bar{X}$ is the mean value across the spectrogram. Then, the normalized spectral densities in dB (X_L) are calculated as

$$X_L = 20 \log\left(\frac{X}{\sigma}\right). \tag{3}$$

3. Data augmentation

Data augmentation is often used for improving the generalization performance of the DL models.[61] A common approach is to implement data augmentation into a dataset by applying a small amount of random noise to each sample.[62] Because the background noise level in the ocean varies with the frequency, we use the realistic wind noise spectrum found in Fig. 1 of Hildebrand.[63] Specifically the wind noise spectrum denoted as level #2 is used to define the maximum amount of noise applied to the spectrograms during the training process. For each epoch, when a data sample is drawn, a different random noise signal is added; each random noise signal is drawn from a uniform distribution with maximum values denoted by the wind noise spectrum and minimum values of 0 dB, forming a matrix that matches the dimensions of the input spectrograms. This additive noise matrix is then applied to each spectrogram X

TABLE VI. The architecture of ResNet-18. The description shows the number of blocks in parenthesis, kernel size, number of kernels, and the stride applied to the input.

Layer name	Description
Conv1	$(\times 1)$ 3×3, 12, Stride 1
Conv2_x	$(\times 2)$ $\begin{bmatrix} 3 \times 3, 12 \\ 3 \times 3, 12 \end{bmatrix}$, Stride 1
Conv3_x	$(\times 2)$ $\begin{bmatrix} 3 \times 3, 24 \\ 3 \times 3, 24 \end{bmatrix}$, Stride 2
Conv4_x	$(\times 2)$ $\begin{bmatrix} 3 \times 3, 48 \\ 3 \times 3, 48 \end{bmatrix}$, Stride 2
Conv5_x	$(\times 2)$ $\begin{bmatrix} 3 \times 3, 96 \\ 3 \times 3, 96 \end{bmatrix}$, Stride 2
Pool1	$(\times 1)$ Average pooling, 4×4
FC	$(\times 1)$ FC layer, 34-d, softmax

when it is drawn during the training, as $X = X + \text{Noise}_{\text{rand}}$, before computing the normalization factor σ in Eq. (2). With this processing, the networks will learn from data samples with different levels of noise at each iteration.

IV. RESULTS AND DISCUSSION

Seabed classification results for the training, validation, and generalization stages are presented in this section for six DL architectures. The training and validation of the networks was done using synthetic data to evaluate whether the algorithms have learned to distinguish among the 34 seabed types. To assess the performance of the DL algorithms against environmental mismatch, the networks were tested on several simulated test cases with SSPs and water depths different from the ones used for training. The metric used for evaluating the performance of the networks is classification accuracy, which is calculated by counting the number of times the network predicted the correct seabed class. For further assessing the generalizability of the networks, the trained models are applied to the 69 data samples measured in the New England Mud Patch Area.

A. Training and validation

The six networks were tested with fivefold CV, where the training data are randomly partitioned into five equally sized segments. Then, five iterations of training and validation are performed such that for each iteration, a different segment of the data is used for validation and the remaining four segments are used for training.[64] Results for the validation stage of the fivefold CV are shown in the summary Table VII, where the average accuracy across the five trained models per network is reported. The DL algorithms exhibit an accuracy above 97% when tested on samples not seen during training. This validation accuracy provides evidence that the trained networks have learned how the signatures of the 34 seabed classes are encoded in the spectrograms.

B. Generalization results: Tests on simulated data

Five simulated test cases are used for evaluating the ability of the networks to classify the sediment type when the sound speed mismatch occurs in the water column. For each test case, the classification accuracy is computed independently for each seabed type based on predictions from the five networks trained in the CV step. The mean and median accuracies across the predictions are reported in Table VII for the six DL architectures. Tests A1 and A2, used for evaluating the impact of the SSP mismatch, exhibit the best performances across all of the test cases with a median accuracy in the range of 75.36%–98.17%. Tests B1 and B2, which have a water depth mismatch of ± 3 m, have a lower median accuracy in the 56.12%–92.12% range. Test C was the most challenging case—when a downward refracting SSP was used for simulating the SOO spectrograms—and the median accuracy significantly dropped to 10.3%–54.12%. A pictorial representation of the networks' median accuracy for the five test cases is shown in Fig. 6(a). It is observed that networks with a deeper and more complicated topology exhibit a higher accuracy for all of the test cases, whereas Selkie3 (1D), with the simplest and shallowest architecture, has the lowest performance. In particular, CNN5-UD, with a topology inspired by a previous study for seabed classification using towed tonals,[19] outperformed the rest of the networks in every test. These results illustrate the impact that environmental mismatch can have on seabed classification efforts.

To further analyze the impact of environmental mismatch on the performance of the DL algorithms, confusion matrices for the five test cases are shown in Fig. 6. These matrices show the percentage of times that the testing data samples generated with a particular seabed are classified from the 34 possible classes. These matrices give insights into where the possible errors among classes are happening for each test case. These results are presented for the five trained DL models of CNN5-UD and lead to several observations about the classification performance for independent seabed types. First, for all of the cases, the networks confuse seabed *korean_h* (#32) with *gravel_35m* (#33) and vice versa, having a low accuracy of less than 50%. It must be noted that these two sediments have the highest sound speed ratio used in this study and share similar sediment profiles. Similarly, but to a lesser extend, seabeds #4, #5, and #10 are often confused with each other. Another observation is that in the face of the sound speed mismatch in the water column (tests A1, A2, and C), seabed *belcourt_2* (#6) is frequently classified as *belcourt_1* (#3) but, interestingly, the opposite does not hold. One important finding is that the most robust classes in this analysis correspond to some of the sediments inferred in previous geacoustic inversions for the SBCEX 2017, indexed as seabed types #0, #1, #2, #3, #8, #18, and

TABLE VII. The validation results on the simulated data for the six trained networks. The mean/median values of accuracy in % are reported for the five trained models per network for the synthetic test cases results.

	Selkie3 (1D)	Selkie3 (2D)	Selkie5	CNN5-UD	HalfAlexNet	ResNet-18
Fivefold CV	99.15	98.51	98.81	99.99	97.15	99.96
Test A1	75.08/82.10	80.47/83.83	82.55/86.86	91.97/97.78	84.96/90.15	88.76/95.58
Test A2	72.34/75.36	76.8/83.83	80.81/87.98	90.4/98.17	82.55/89.58	87.89/94.84
Test B1	58.17/57.68	65.45/69.26	68.54/76.54	78.44/82.44	72.44/81.51	73.83/82.96
Test B2	57.39/56.12	64.73/64.62	71.49/77.83	81.25/92.12	73.76/79.36	74.95/84.37
Test C	17.76/10.3	41.46/39.88	35.37/30.72	53.11/54.12	37.2/30.12	48.29/48.64

https://doi.org/10.1121/10.0005936

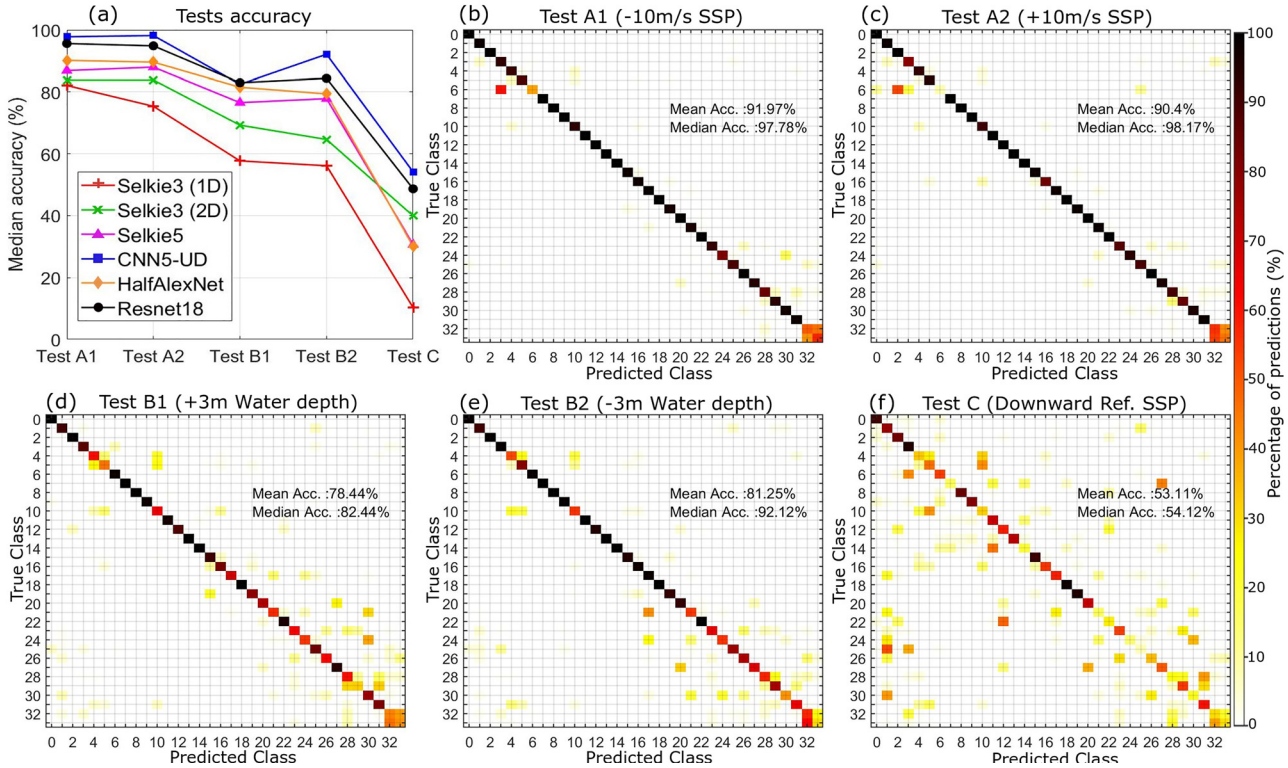

FIG. 6. (Color online) (a) The median classification performance of six DL architectures on the synthetic test cases. The confusion matrices for [(b),(c)] tests A1 and A2 with the SSPs shifted by ±10 m/s, [(d),(e)] tests B1 and B2 with ±3 m water depth mismatch, and (f) test C with a downward refracting SSP. The confusion matrices are presented for the five DL models trained for CNN5-UD. The correct predictions are located along the diagonal cells. The color indicates the percentage of the number of predictions per seabed type in which the values in each row sum up to 100.

#19; these results indicate that although these classes represent a mud over sand environment from the same area, they are unique enough for CNN5-UD to distinguish among them with high accuracy even in the face of the downward refracting SSP used for test C. These confusion matrices show that incorrect classifications generally occur between similar seabeds.

C. Generalization results: Test on measured data

The trained DL models were applied to the measured SOO spectrograms listed in Table II to assess the generalizability of the networks on data collected at-sea. Classification of the 69 samples collected from the merchant ships moving next to 3 receivers located at different locations but similar depths in the SBCEX 2017 are shown in Fig. 7. Classification results obtained from the five trained models (from the CV step) for each of the six DL topologies (described in Sec. III B 1) are presented in the form of a stacked bar plot for the seabed type classifications using the 69 measured data samples, where the colors and patterned symbols of the bars differentiate the network architectures. Results indicate that over 94% of the time, all of the networks classify the SOO spectrograms as a mud over sand sediment inferred in the SBCEX 2017, which is the same area in which the samples were collected. Seabed types belcourt_2_sbc (#3) and belcourt_1_sbc (#6) got 47.83% of the overall predictions, making them the most likely sediments in this DL analysis. Interestingly, these two seabeds were inferred in the same study performed by Belcourt *et al.*,[44] using a trans-dimensional Bayesian inversion approach on the reflection coefficient data from two locations in the SBCEX 2017 close to the VLAs and shipping lanes used for this work. It should be noted that in the synthetic test cases, seabed #6 was mistakenly classified as #3 in the face of a SSP mismatch in the water column. Also, seabeds #4, #5, and #10 were often confused with each other in the simulated tests, and these three sediments combined got 25.12% of the predictions when tested on the measured data samples.

To further assess the classification results, a data-model comparison for three of the recorded ships at different positions is shown in Fig. 8. The spectrograms have been normalized, based on their standard deviation as explained in Sec. III B 2, to mitigate the effects of the unknown source level intrinsic to each ship. For each VLA, one example of a measured ship spectrogram and a modeled ship spectrogram (using seabed #3) is presented. The source parameters used for simulating the spectrograms were inferred based on the speeds and positions of the ships reported in the AIS database; each ship's Global Positioning System (GPS) data from the database are shown in the right column of Fig. 8. The vessels used for this comparison are as follows. With a CPA range of 8.96 km with respect to VLA1-MPL on JD 83 at 03:26 UTC, the vessel Atlantic Conveyor was transiting in the northern shipping lane at a constant speed of 16.1 kn.

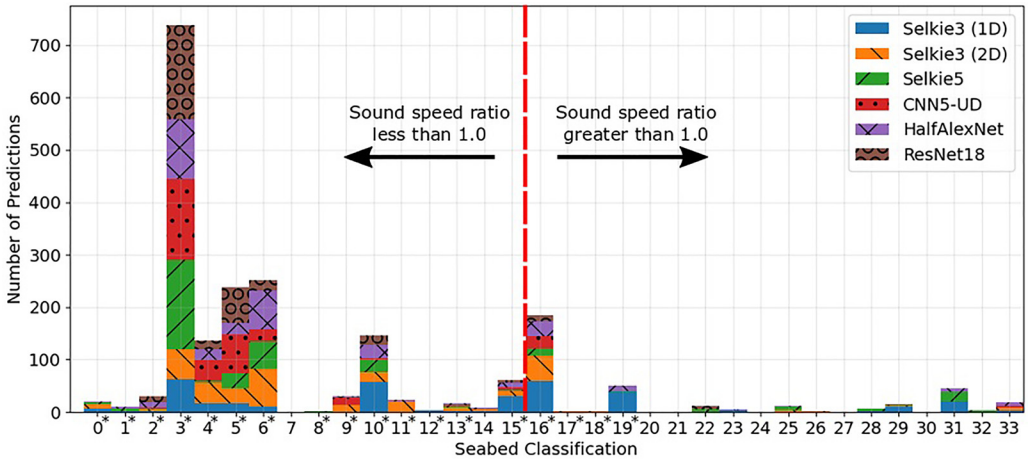

FIG. 7. (Color online) The seabed classification results for 6 DL architectures tested on 69 measured SOO spectrograms. In total, 2070 predictions are reported, corresponding to the 5 trained models per network type (indicated by the color). Seabed classes with an "*" were inferred in previous geoacoustic inversions for the SBCEX 2017 area. The red dotted line in the center of the figure separates the sediments with sound speed ratios less than and greater than one.

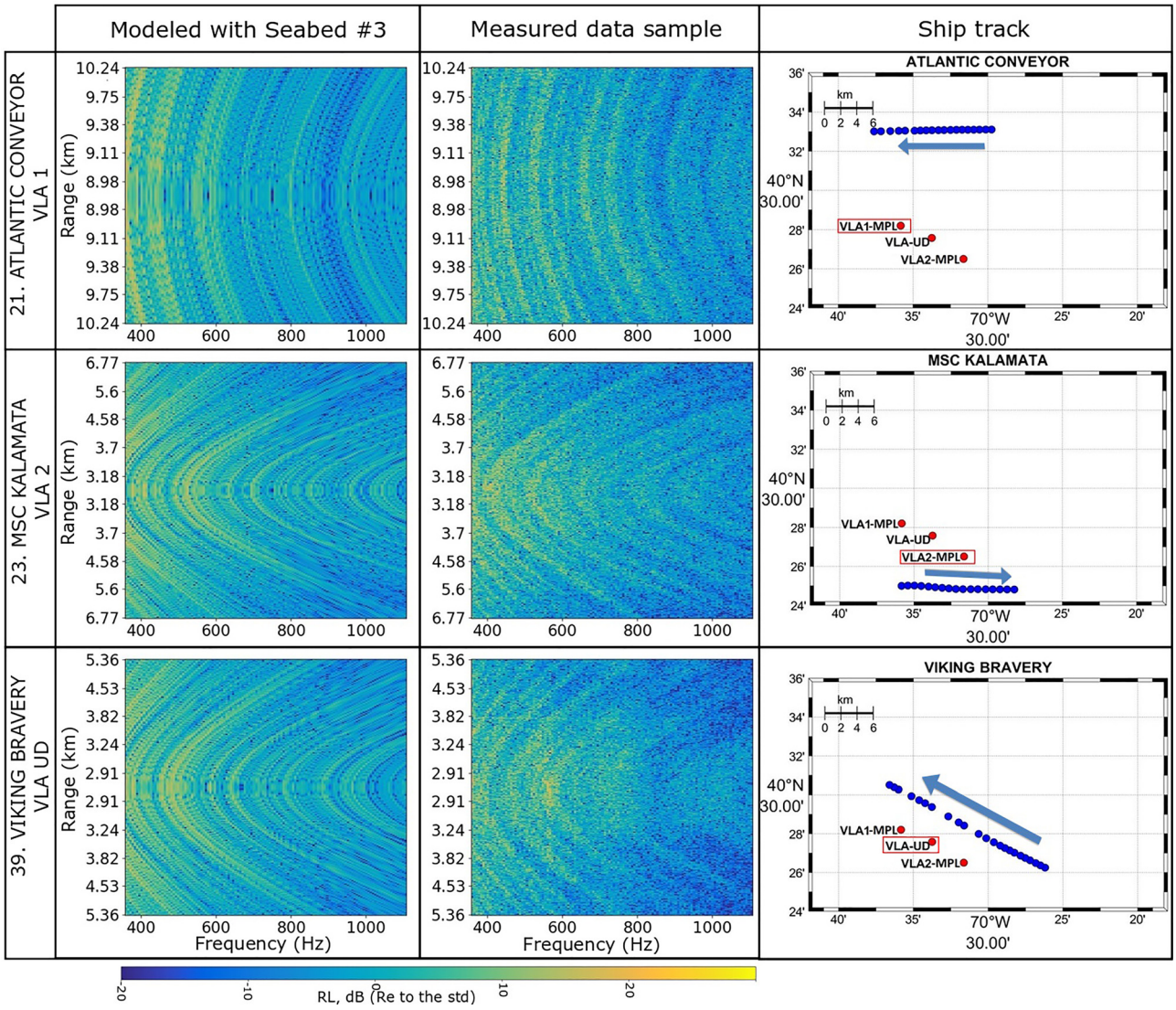

FIG. 8. (Color online) The data-model comparison of the normalized spectrograms with respect to their standard deviation. The modeled spectrograms using seabed #3 (left column) and measured spectrograms (middle column) for ship 21. The ATLANTIC CONVEYOR was measured at VLA1-MPL with CPA = 8.96 km and $v_{sh} = 16.065$ kn (top row), ship 23. The MSC KALAMATA was measured at VLA2-MPL with CPA = 3.6 km and $v_{sh} = 19.47$ kn (middle row), and ship 39. The VIKING BRAVERY was measured at VLA UD with CPA = 2.87 km and $v_{sh} = 14.665$ kn (bottom row). The location of the VLAs and the ship locations from the AIS database over the time period displayed are shown in the right column, where the arrows indicate the direction of the vessel trajectory.

The MSC Kalamata, on the other hand, was moving in the southern shipping lane at 19.9 kn at a closer CPA range of 3.11 km with respect to VLA2-MPL on JD 83 at 18:30 UTC. The Viking Bravery was transiting across the shipping lanes (from southern to northern shipping lane) with a similar CPA range (2.87 km with respect to VLA-UD) on JD 90 at 07:15 UTC but was moving slower with a constant speed of 14.67 kn.

Whereas these merchant ships were moving through different shipping lanes at different combinations of CPA ranges and speeds and recorded at three different locations, the spectrograms show the qualitative agreement between the measured signals collected during the SBCEX 2017 and SOO spectrograms modeled using the most likely seabed type #3, belcourt_2_sbc. This agreement provides support for the ensemble approach to DL for seabed classification. Although the predictions for individual data samples with one network might not always be accurate, the most likely prediction from the ensemble gives a reasonable and more reliable estimate of the seabed class.

V. CONCLUSIONS

An ensemble of DL algorithms was implemented for seabed classification using 34 seabed types that spanned from soft-muddy to hard-sandy environments. The models were trained on 20-min normalized spectrograms over the 360–1100 Hz frequency band generated using the Wales-Heitmeyer source spectrum for a merchant ship and range-independent normal-mode model ORCA. Several approaches were considered for the design of the networks by using different complexity levels in the architectures. The DL algorithms implemented in this study are as simple as a shallow 3-layer CNN (Selkie3) and as complex as a deep 18-layer ResNet (ResNet18). The validation results show accuracy above 97% in the fivefold CV test for all of the architectures. This performance indicates that the SOO spectrograms have enough information content about the waveguide for all of the networks to learn representative patterns during training for distinguishing the 34 seabed classes.

The generalization performance of the networks in the face of an environmental mismatch in the water column was evaluated using five simulated test cases. The more complex and deeper architectures proved to be more robust against SSP and water depth mismatch, whereas the shallowest network had a lower performance. Selkie3, in both 1D and 2D input formats, exhibited the lowest accuracy among all of the networks. However, by using the 1D input format, the median accuracy of the network decreased by about 10% (onn average across the five test cases) with respect to the 2D format. This result shows the importance of taking into account the relationship among adjacent frequency points in the convolutional stage by keeping the 2D structure of the data for these types of signals.

For further assessing the generalization performance of the networks, the trained models were applied to 69 samples collected in the SBCEX 2017. The results showed that 94% of the time, the SOO spectrograms were classified as a mud over sand sediments inferred from geoacoustic inversions of data obtained in the same area. Moreover, the classifications across all of the networks consistently indicated that the most likely seabed type, with 35.65% of the overall predictions, is belcourt_2_sbc (#3), obtained in Ref. 44, followed by belcourt_1_sbc (#6) from the same study with 12.17% of the predictions. The data model comparisons for ships transiting different routes and moving at various speeds show agreement between the simulated and measured ship-radiated noises using belcourt_2_sbc (#3).

This work provides evidence that DL algorithms are able to find representative patterns from SOOs for seabed classification using a wide variety of seabed types. There are many ways to build a DL architecture by tuning, for instance, the depth, number of channels, kernel size, pooling, and input format. These modifications change the patterns that are learned during training from network to network. By using an ensemble approach on the trained DL models, when a unified prediction is obtained, then that answer is more likely to be correct than relying on the prediction from one single architecture. This behavior was observed in this work, where the measured data samples yielded a unified answer toward seabed #3 using six networks with diverse architectures. Furthermore, this study demonstrated the importance of using simulated test cases to evaluate the robustness of the networks independently before applying them to data collected at-sea. This generalization assessment is useful for designing DL architectures suitable for the desired input data.

As a final remark, it is important to emphasize that the selection of SSPs used for simulating the synthetic training dataset is a key aspect to consider when using DL for classifying the most likely effective seabed type. In this work, the behavior of the water column in the SBCEX 2017 was well studied and presented to the networks during training. However, for waveguides with substantially different water column depths and SSPs, new networks will need to be trained using a synthetic dataset that encompasses the information about the possible new water column.

ACKNOWLEDGMENTS

This work was funded by the Office of Naval Research Contract No. N00014-19-C-2001. The SBCEX 2017 was funded by the U.S. Navy Office of Naval Research. The authors thank the captains and crews of RV Endeavor.

[1]C. Gervaise, B. G. Kinda, J. Bonnel, Y. Stéphan, and S. Vallez, "Passive geoacoustic inversion with a single hydrophone using broadband ship noise," J. Acoust. Soc. Am. **131**(3), 1999–2010 (2012).
[2]S. C. Wales and R. M. Heitmeyer, "An ensemble source spectra model for merchant ship-radiated noise," J. Acoust. Soc. Am. **111**(3), 1211–1231 (2002).
[3]K. D. Heaney, "Rapid geoacoustic characterization using a surface ship of opportunity," IEEE J. Ocean. Eng. **29**(1), 88–99 (2004).
[4]P. H. Dahl and D. R. Dall'Osto, "Estimation of seabed properties and range from vector acoustic observations of underwater ship noise," J. Acoust. Soc. Am. **147**(4), EL345–EL350 (2020).

[5]R. A. Koch and D. P. Knobles, "Geoacoustic inversion with ships as sources," J. Acoust. Soc. Am. **117**(2), 626–637 (2005).
[6]S. A. Stotts and R. A. Koch, "Geoacoustic inversions and localizations with adaptively beamformed data from a surface ship of opportunity source," J. Acoust. Soc. Am. **127**, 84–95 (2010).
[7]M. Nicholas, J. S. Perkins, G. J. Orris, L. T. Fialkowski, and G. J. Heard, "Environmental inversion and matched-field tracking with a surface ship and an *L*-shaped receiver array," J. Acoust. Soc. Am. **116**(5), 2891–2901 (2004).
[8]D. Tollefsen and S. E. Dosso, "Bayesian geoacoustic inversion of ship noise on a horizontal array," J. Acoust. Soc. Am. **124**(2), 788–795 (2008).
[9]D. P. Knobles, "Maximum entropy inference of seabed attenuation parameters using ship radiated broadband noise," J. Acoust. Soc. Am. **138**, 3563–3575 (2015).
[10]C. Park, W. Seong, and P. Gerstoft, "Geoacoustic inversion in time domain using ship of opportunity noise recorded on a horizontal towed array," J. Acoust. Soc. Am. **117**(4), 1933–1941 (2005).
[11]D. J. Battle, P. Gerstoft, W. A. Kuperman, W. S. Hodgkiss, and M. Siderius, "Geoacoustic inversion of tow-ship noise via near-field-matched-field processing," IEEE J. Ocean. Eng. **28**(3), 454–467 (2003).
[12]X. Zhang, N. C. Durofchalk, H. Niu, L. Wu, R. Zhang, K. G. Sabra, X. Zhang, N. C. Durofchalk, H. Niu, L. Wu, and R. Zhang, "Geoacoustic inversion using ray-based blind deconvolution of shipping sources," J. Acoust. Soc. Am. **147**, 285–299 (2020).
[13]M. J. Bianco, P. Gerstoft, J. Traer, E. Ozanich, M. A. Roch, S. Gannot, C. A. Deledalle, and W. Li, "Machine learning in acoustics: Theory and applications," arXiv:3590(2019) (2019).
[14]Y. Stephan, X. Demoulin, and O. Sarzeaud, "Neural direct approaches for geoacoustic inversion," J. Comput. Acoust. **06**(01n02), 151–166 (1998).
[15]Z. H. Michalopoulou, D. Alexandrou, and C. de Moustier, "Application of neural and statistical classifiers to the problem of seafloor characterization," IEEE J. Ocean. Eng. **20**(3), 190–197 (1995).
[16]J. Benson, N. R. Chapman, and A. Antoniou, "Geoacoustic model inversion using artificial neural networks," Inverse Probl. **16**, 1627–1639 (2000).
[17]J. Piccolo, G. Haramuniz, and Z.-H. Michalopoulou, "Geoacoustic inversion with generalized additive models," J. Acoust. Soc. Am. **145**(6), EL463–EL468 (2019).
[18]H. Niu, P. Gerstoft, E. Ozanich, Z. Li, R. Zhang, Z. Gong, and H. Wang, "Block sparse Bayesian learning for broadband mode extraction in shallow water from a vertical array," J. Acoust. Soc. Am. **147**(6), 3729–3739 (2020).
[19]T. B. Neilsen, C. D. Escobar-Amado, M. C. Acree, W. S. Hodgkiss, D. F. Van Komen, D. P. Knobles, M. Badiey, and J. Castro-Correa, "Learning location and seabed type from a moving mid-frequency source," J. Acoust. Soc. Am. **149**(1), 692–705 (2021).
[20]D. F. Van Komen, T. B. Neilsen, K. Howarth, D. P. Knobles, and P. H. Dahl, "Seabed and range estimation of impulsive time series using a convolutional neural network," J. Acoust. Soc. Am. **147**(5), EL403–EL408 (2020).
[21]C. Frederick, S. Villar, and Z. H. Michalopoulou, "Seabed classification using physics-based modeling and machine learning," arXiv:859 (2020).
[22]D. F. Van Komen, T. B. Neilsen, D. B. Mortenson, M. C. Acree, D. P. Knobles, M. Badiey, and W. S. Hodgkiss, "Seabed type and source parameters predictions using ship spectrograms in convolutional neural networks," J. Acoust. Soc. Am. **149**(2), 1198–1210 (2021).
[23]Bureau of Ocean Energy Management (BOEM) and National Oceanic and Atmospheric Administration (NOAA), available at https://marinecadastre.gov/ais/ (AIS_2017_03_Zone19 and AIS_2017_04_Zone19) (Last viewed August 7, 2020).
[24]E. K. Westwood, C. T. Tindle, and N. R. Chapman, "A normal mode model for acousto-elastic ocean environments," J. Acoust. Soc. Am. **100**(6), 3631–3645 (1996).
[25]D. J. Forman, T. B. Neilsen, D. F. Van Komen, and D. P. Knobles, "Validating deep learning seabed classification via acoustic similarity," JASA Express Lett. **1**(4), 040802 (2021).
[26]P. S. Wilson, D. P. Knobles, and T. B. Neilsen, "Guest editorial an overview of the seabed characterization experiment," IEEE J. Ocean. Eng. **45**(1), 1–13 (2020).
[27]D. P. Knobles, R. A. Koch, L. A. Thompson, K. C. Focke, and P. E. Eisman, "Broadband sound propagation in shallow water and geoacoustic inversion," J. Acoust. Soc. Am. **113**(1), 205–222 (2003).
[28]F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics*, 2nd ed. (Springer, New York, 2011), p. 39.
[29]D. P. Knobles, R. A. Koch, and M. S. Haire, "Geoacoustic inversion studies," Oceans Conf. Rec. **2**, 737–744 (2001).
[30]S. E. Dosso, P. M. Giles, G. H. Brooke, D. F. McCammon, S. Pecknold, and P. C. Hines, "Linear and nonlinear measures of ocean acoustic environmental sensitivity," J. Acoust. Soc. Am. **121**(1), 42–45 (2007).
[31]W.-H. Ryang, J.-H. Jin, and J. Hahn, "Geoacoustic model at the YSDP-105 Long-core Site in the Mid-eastern Yellow Sea," J. Korean Earth Sci. Soc. **40**(1), 24–36 (2019).
[32]W. H. Ryang, S. P. Kim, S. Kim, and D. C. Kim, "Geoacoustic model of the transverse acoustic variability experiment area in the northern East China Sea," Geoscie. J. **17**(3), 267–278 (2013).
[33]H. Kwon, J. W. Choi, W.-H. Ryang, S.-U. Son, and S.-k. Jung, "Measurements of mid-frequency bottom-interacting signals and geoacoustic inversion in Jinhae Bay, Southeast Korea," J. Acoust. Soc. Am. **145**(3), 1205–1211 (2019).
[34]G. R. Potty, J. H. Miller, and J. F. Lynch, "Inversion for sediment geoacoustic properties at the New England Bight," J. Acoust. Soc. Am. **114**(4), 1874–1887 (2003).
[35]D. P. Knobles, J. A. Goff, R. A. Koch, P. S. Wilson, and J. A. Shooter, "Effect of inhomogeneous sub-bottom layering on broadband acoustic propagation," IEEE J. Ocean. Eng. **35**(4), 732–743 (2010).
[36]Y.-M. Jiang, N. R. Chapman, and M. Badiey, "Quantifying the uncertainty of geoacoustic parameter estimates for the New Jersey shelf by inverting air gun data," J. Acoust. Soc. Am. **121**(4), 1879–1894 (2007).
[37]W.-H. Ryang, S.-P. Kim, D.-C. Kim, and J. Hahn, "Geoacoustic model of coastal bottom strata at Jeongdongjin in the Korean Continental Margin of the East Sea," J. Korean Earth Sci. Soc. **37**, 200–210 (2016).
[38]J.-X. Zhou, X.-Z. Zhang, and D. P. Knobles, "Low-frequency geoacoustic model for the effective properties of sandy seabottoms," J. Acoust. Soc. Am. **125**(5), 2847–2866 (2009).
[39]Y. T. Lin, J. Bonnel, D. P. Knobles, and P. S. Wilson, "Broadband waveform geoacoustic inversions with absolute travel time," IEEE J. Ocean. Eng. **45**(1), 174–188 (2020).
[40]D. P. Knobles, P. S. Wilson, J. A. Goff, L. Wan, M. J. Buckingham, J. D. Chaytor, and M. Badiey, "Maximum entropy derived statistics of sound-speed structure in a fine-grained sediment inferred from sparse broadband acoustic measurements on the New England continental shelf," IEEE J. Ocean. Eng. **45**(1), 161–173 (2020).
[41]E. M. Brown, Y. T. Lin, J. D. Chaytor, and W. L. Siegmann, "Geoacoustic inversion for a New England Mud Patch sediment using the silt-suspension theory of marine mud," IEEE J. Ocean. Eng. **45**(1), 144–160 (2020).
[42]P. H. Dahl and D. R. Dall'Osto, "Vector acoustic analysis of time-separated modal arrivals from explosive sound sources during the 2017 seabed characterization experiment," IEEE J. Ocean. Eng. **45**(1), 131–143 (2020).
[43]Z. H. Michalopoulou and P. Gerstoft, "Multipath broadband localization, bathymetry, and sediment inversion," IEEE J. Ocean. Eng. **45**(1), 92–102 (2020).
[44]J. Belcourt, C. W. Holland, S. E. Dosso, J. Dettmer, and J. A. Goff, "Depth-dependent geoacoustic inferences with dispersion at the New England Mud Patch via reflection coefficient inversion," IEEE J. Ocean. Eng. **45**(1), 69–91 (2020).
[45]G. R. Potty and J. H. Miller, "Effect of shear on modal arrival times," IEEE J. Ocean. Eng. **45**(1), 103–115 (2020).
[46]D. Tollefsen, S. E. Dosso, and D. P. Knobles, "Ship-of-opportunity noise inversions for geoacoustic profiles of a layered mud-sand seabed," IEEE J. Ocean. Eng. **45**(1), 189–200 (2020).
[47]L. Wan, M. Badiey, D. P. Knobles, and P. S. Wilson, "The Airy phase of explosive sounds in shallow water," J. Acoust. Soc. Am. **143**(3), EL199–EL205 (2018).
[48]D. R. Barclay, D. A. Bevans, and M. J. Buckingham, "Estimation of the geoacoustic properties of the New England Mud Patch from the vertical coherence of the ambient noise in the water column," IEEE J. Ocean. Eng. **45**(1), 51–59 (2020).
[49]J. Bonnel, Y.-T. Lin, D. Eleftherakis, J. A. Goff, S. Dosso, R. Chapman, J. H. Miller, and G. R. Potty, "Geoacoustic inversion on the New England Mud Patch using warping and dispersion curves of high-order modes," J. Acoust. Soc. Am. **143**(5), EL405–EL411 (2018).

[50]L. Wan, M. Badiey, D. P. Knobles, P. S. Wilson, and J. A. Goff, "Estimates of low-frequency sound speed and attenuation in a surface mud layer using low-order modes," IEEE J. Ocean. Eng. **45**(1), 201–211 (2020).
[51]J. Bonnel, S. E. Dosso, D. Eleftherakis, and N. R. Chapman, "Trans-dimensional inversion of modal dispersion data on the New England Mud Patch," IEEE J. Ocean. Eng. **45**(1), 116–130 (2020).
[52]I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (MIT Press, Cambridge, MA, 2016).
[53]K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2016).
[54]H. Niu, Z. Gong, E. Ozanich, P. Gerstoft, H. Wang, and Z. Li, "Deep-learning source localization using multi-frequency magnitude-only data," J. Acoust. Soc. Am. **146**(1), 211–222 (2019).
[55]A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proceedings of the 25th International Conference on Neural Information Processing Systems* (NIPS'12) (Curran Associates Inc., Red Hook, NY, 2012), Vol. 1, pp. 1097–1105.
[56]K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2016), pp. 770–778.
[57]A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, and L. Antiga, "PyTorch: An imperative style, high-performance deep learning library," in *Advances in Neural Information Processing Systems* (2019), pp. 8024–8035.
[58]V. Nair and G. E. Hinton, "Rectified linear units improve restricted Boltzmann machines," in *Proceedings of the 27th International Conference on Machine Learning (ICML-10)* (Omnipress, Madison, WI, 2010), pp. 801–814, available at https://icml.cc/Conferences/2010/papers/432.pdf (Last viewed: August 17, 2021).
[59]I. Loshchilov and F. Hutter, "Decoupled weight decay regularization," arXiv:1711.05101 (2017).
[60]D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," in *3rd International Conference on Learning Representations, ICLR* (2015), Conference Track Proceedings 1–15.
[61]K. Imai and D. A. Van Dyk, "A Bayesian analysis of the multinomial probit model using marginal data augmentation," J. Econometr. **124**(2), 311–334 (2005).
[62]L. Tóth, G. Kovács, and D. Van Compernolle, "A perceptually inspired data augmentation method for noise robust CNN acoustic models," in *International Conference on Speech and Computer* (Springer, Cham, 2018), pp. 697–706.
[63]J. Hildebrand, "Anthropogenic and natural sources of ambient noise in the ocean," Mar. Ecol. Prog. Ser. **395**, 5–20 (2009).
[64]R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *Proceedings of the Twelfth International Conference on Machine Learning (ICML'95)* (Morgan Kaufmann, San Francisco, CA, 1995), pp. 1137–1143.