NOVEMBER 20 2024

Verifying models of the underwater soundscape from wind and ships with benchmark scenarios[a)]

S. Bruce Martin ; Martin Siderius ; Michael A. Ainslie ; Michele B. Halvorsen ; Leila Hatch ; Mark K. Prior ; Daniel Brooker ; James Caplinger ; Christine Erbe ; John Gebbie ; Kevin D. Heaney ; Alexander O. MacGillivray ; Marie-Noel Matthews ; Victor O. Oppeneer ; Alexandra Schäfke; Renée P. Schoeman; H. Özkan Sertlek

J. Acoust. Soc. Am. 156, 3422–3438 (2024)
https://doi.org/10.1121/10.0026597

JASA ARTICLE

Verifying models of the underwater soundscape from wind and ships with benchmark scenarios[a)]

S. Bruce Martin,[1,b)] Martin Siderius,[2] Michael A. Ainslie,[3] Michele B. Halvorsen,[4] Leila Hatch,[5] Mark K. Prior,[6] Daniel Brooker,[7] James Caplinger,[8] Christine Erbe,[9] John Gebbie,[10] Kevin D. Heaney,[11] Alexander O. MacGillivray,[12] Marie-Noel Matthews,[3] Victor O. Oppeneer,[13] Alexandra Schäfke,[14] Renée P. Schoeman,[9] and H. Özkan Sertlek[3]

[1]*JASCO Applied Sciences, 20 Mount Hope Avenue, Dartmouth, Nova Scotia B2Y 4S3, Canada*

[2]*Portland State University, 1825 Southwest Broadway, Portland, Oregon 97201, USA*

[3]*JASCO Applied Sciences, Lise-Meitner-Straße 9, 24223 Schwentinental, Germany*

[4]*JASCO Applied Sciences, 8701 Georgia Avenue, Suite 410, Silver Spring, Maryland 20910, USA*

[5]*Office of National Marine Sanctuaries, United States National Oceanic and Atmospheric Administration, 1305 East-West Highway, 11th Floor, Silver Spring, Maryland 20910, USA*

[6]*Comprehensive Nuclear-Test-Ban Treaty Organization, Vienna International Centre, P.O. Box 1200, 1400 Vienna, Austria*

[7]*Naval Research Lab, 4555 Overlook Avenue, Washington, DC 20375, USA*

[8]*Office of Protected Resources, United States National Oceanic and Atmospheric Administration, 1315 East-West Highway, 13th Floor, Silver Spring, Maryland 20910, USA*

[9]*Centre for Marine Science and Technology, Curtin University, B301 Hayman Road, Bentley Western Australia 6102, Australia*

[10]*Metron, Incorporated, 1818 Library Street #600, Reston, Virginia 20190, USA*

[11]*Applied Ocean Sciences, 5242 Port Royal Road #1032, Springfield, Virginia 22151, USA*

[12]*JASCO Applied Sciences, 2305–4464 Markham Street, Victoria, British Columbia V8Z 7X8, Canada*

[13]*Nederlandse Organisatie voor Toegepast Natuurwetenschappelijk Onderzoek, Motion Building, Radarweg 60, 1043 NT Amsterdam, the Netherlands*

[14]*Bundeswehr Technical Center for Ships and Naval Weapons, Maritime Technology and Research, Berliner Strasse 115, 24340 Eckernförde, Schleswig-Holstein, Germany*

ABSTRACT:

Models of the underwater acoustic soundscape are important for evaluating the effects of human generated sounds on marine life. The performance of models can be validated against measurements or verified against each other for consistency. A verification workshop was held to compare models that predict the soundscape from wind and vessels and estimate detection ranges for a submerged target. Eight modeling groups participated in the workshop which predicted sound levels with observation windows of 1 min and 1 km^2. Substantial differences were found in how modelers computed the propagation losses for decidecade bands and estimated the source level of wind. Further investigations resulted in recommendations on best practices. Choices of temporal and spatial modeling resolution affected the estimates of metrics proportional to total sound energy more than distributions of sound pressure level. Deeper receivers were less sensitive to these parameters than shallow ones. A temporal resolution of 1 min and spatial resolution of 100 m is recommended. Models that follow the recommendations will yield similar results. The detection range of underwater targets is highly variable when the ambient noise depends on moving noise sources. Future work to verify models against data and understand model uncertainty is recommended.

 https://doi.org/10.1121/10.0026597

(Received 1 March 2024; revised 19 June 2024; accepted 20 June 2024; published online 20 November 2024)

[Editor: James F. Lynch] Pages: 3422–3438

I. INTRODUCTION

Sound from human activities is recognized as a worldwide stressor for marine life and a source of ambient noise (Duarte *et al.*, 2021). Predicting and mitigating acute effects (hearing injury, physical injury) of extremely intense human sounds like seismic airgun surveys (Gisiner, 2016) and impact pile driving (Dahl *et al.*, 2015) has been the focus of substantial research. Recently, interest has been growing into the effects of vessel sound on marine life (International Maritime Organization, 2014). While the sound from shipping is unlikely to cause acute injury, it has the potential to cause masking of biologically relevant sounds (Erbe *et al.*, 2016; Pine *et al.*, 2018), or changes in behavior. Behavioral

[a)]This paper is part of a special issue on Verification and Validation of Source and Propagation Models for Underwater Sound.

[b)]Email: bruce.martin@jasco.com

https://doi.org/10.1121/10.0026597

changes that have been documented range from impacted rest periods in seals (Mikkelsen *et al.*, 2019), to changes in vocal behavior in humpbacks (Fournet *et al.*, 2018), to changes in settlement behavior of reef fish larvae (Simpson *et al.*, 2016). The intense sound sources that cause acute injuries are generally contained in small areas for limited times. This contrasts with the large number of vessels in the world's oceans that often travel along shipping lanes and expose those areas to elevated sound levels many times per day.

A question that often arises during assessment of the potential impact of a proposed marine development is how the changed environment will affect animals. When considering behavioral changes or disturbance, recent approaches have employed dose response curves that predict the percentage of a population that will respond as the sound level changes. For example Wood *et al.* (2012) have sound pressure level thresholds at which 10%, 50%, and 90% of individuals are expected to be disturbed for three categories of marine mammals—sensitive odontocetes, migrating mysticetes, and all other marine mammals. Another example is the use of a dose-response curve that predicts the probability that harbor porpoise are disturbed using the single-strike sound exposure level from pile driving (Heinis *et al.*, 2022). Studies assessing changes in masking from changes in vessel patterns have overlapped auditory weighted sound pressure levels with vessel density and mammal densities to estimate hot spots (Erbe *et al.*, 2014), computed relative listening range reductions (Pine *et al.*, 2018), predicted the potential lost foraging time to vessels (Joy *et al.*, 2019), or computed a pressure index that integrates the time and area where human sound exceed natural sounds by some threshold (Kinneging and Tougaard, 2021). All of these approaches distill time and spatially varying sound fields from moving source on moving animals into single indexes that may be compared to assess the merits of different approaches to managing the effects of human sounds on marine life. As a concrete example, it is more insightful to know what the average change in potential lost foraging time is across a whale population's preferred habitat for a full season than to know the value at a single time and location.

Figure 1 provides a conceptual framework for the data and computing needed to estimate the daily average sound pressure level over a wide area from wind and vessels. While many of the operations are familiar to underwater sound propagation modelers, there are challenges within each of the operations and decisions that must be made for the modeling to be computationally feasible. First, the sound spectrum of the vessels and wind are broadband, which implies that the acoustic propagation modelling must be performed for a wide spectrum of frequencies, likely using different numerical models as a function of frequency. The source spectrum for vessels is generally provided as the sound level averaged over decidecade bands, which means the propagation losses must also be representative of the average loss over that decidecade band. The vessel Automated Identification System (AIS) tracks are likely obtained from a satellite AIS source, with irregular reports that must be interpolated to the time resolution of the modelling that will be performed—a choice the modeler must make. Similarly, the wind speed data are likely very coarse (1–3 h), and the modeler must choose how to select a wind speed for each of the times that vessel noise is estimated—typical options are selecting the nearest neighbor and using linear or cubic spline interpolations. The data output at each modeling time step of the vessel sound are three-dimensional matrices of sound pressure level at the modeling spatial resolution—likely on the order of 25 m in latitude and longitude, and at an irregular spacing in the depth dimension (finer at the surface, getting coarser with depth). The matrices likely cover a spatial extent of 10 to 100s of kilometers, and therefore must be immediately spatially down-sampled to avoid generating unmanageably large data sets—the modeler must choose this resolution to meet the requirements of the project while preserving the accuracy of the data. Conversely, the wind sound levels are likely at spatial resolutions of 0.125° or 0.25° and must be up-sampled to be on the same sampling grid as the vessel sound levels. The two data streams may be summed to obtain the instantaneous sound pressure level for a time step and all spatial grid cells. The grid cell closest to the location of an acoustic recorder could then be compared with the modeled time series to validate the modeling.

From the "instantaneous" levels, the modelers must make further decisions to distill this time-frequency-spatial data into the metrics employed for comparisons between options for project developments, or to compare an altered soundscape to the baseline. The first decision is how to compress the data in time, for instance, to sum and obtain a sound exposure level or weighted sound exposure level, to select the maximum level, to average and obtain the mean level, to compute the statistical distribution of sound levels over a period of time, or to derive a different metric of masking or disturbance based on the sound pressure or sound exposure levels and time compress that metric. Similarly, decisions must be made on how to compress the data spatially.

Important decisions on authorizing proposed human developments are informed by these model predictions. Regulators, project proponents, and those cautioning against development use soundscape models to assess possible effects of sound on marine life, and need to know how accurate these models are. There are two means of evaluating their accuracy—by comparison to measurements and by comparisons between models. While many would jump immediately to comparisons with measurements (model validation), there is much to be learned by first comparing models to understand differences between model predictions for identical inputs (model verification). To simplify the model verification, best practice from previous experience is to determine what issues one wants to address in the comparison and develop scenarios that isolate for those issues (Ainslie *et al.*, 2016; Ainslie *et al.*, 2019; Ainslie *et al.*, 2024).

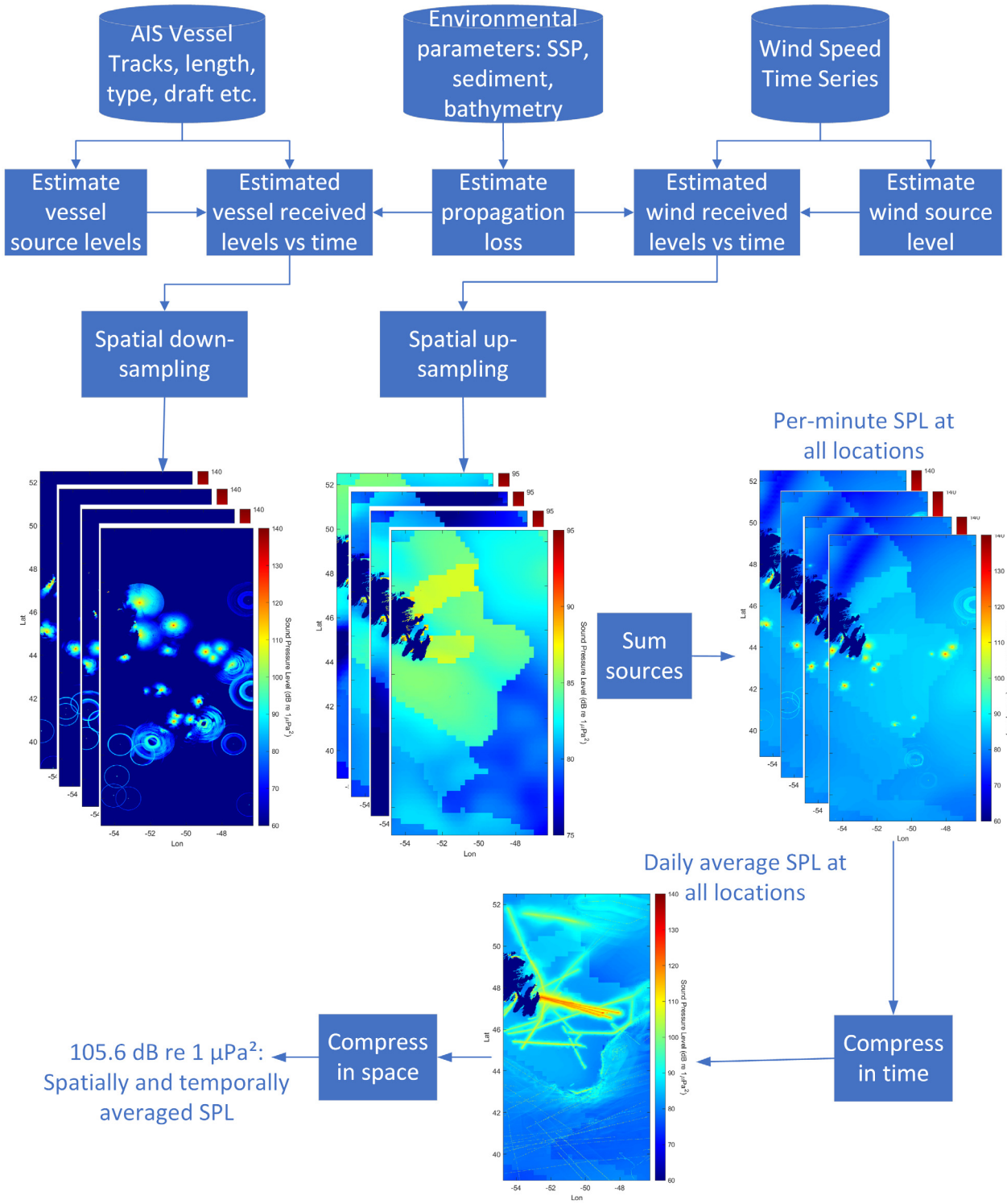

FIG. 1. (Color online) Flow chart for a soundscape model that estimates the average daily sound pressure level at the seabed over a wide area from wind and vessels.

To provide a forum for comparing soundscape modeling, a group of academic, industrial, and government researchers, supported by the United States Office of Naval Research and the Netherlands Ministry of Infrastructure and Water Management (Rijkswaterstaat) conducted an ambient sound modeling workshop. The workshop was held in conjunction with the *Effects of Noise on Aquatic Life* 2022 Conference, on 10 July 2022 in Berlin, Germany. The workshop asked acoustic modeling practitioners to model the same prescribed scenarios with the objective of arriving at a consensus on how best to model the soundscape from ambient sources (wind) and vessels. The goal of the scenarios was to specify benchmark test cases that may be used in the future to demonstrate agreement in modeled predictions

https://doi.org/10.1121/10.0026597

between long-term ambient sound models. The scenarios involve the movement of vessels in a straight shipping lane over the period of one week. Wind speed is also specified for the same 1-week period. These scenarios simplify the acoustic propagation environment and vessel tracks so that differences between models should be caused by how the models include wind speed data, aggregate the vessel distributions, and perform temporal and spatial averaging.

The time series of sound pressure levels from the scenario were also employed as a spatially and temporally variable noise field as part of a submerged vessel detection scenario. This problem, method, and the results obtained are described in Oppeneer *et al.* (2024). Their method of determining if a submerged vessel can be detected along a track could be extended to the detection of marine mammals and their passive monitoring.

This manuscript provides an overview of the ambient soundscape modeling scenario, terminology associated with wide-area, long-term modeling, the metrics used to compare the results, introduces some of the issues the modelers encountered, and provides a summary of the initial results presented at the workshop. Other manuscripts in this Special Issue investigate how to model sound generated by wind, how to estimate propagation losses for decidecade bands, and how ambient sound variability integrates with the detection of sounds within the soundscape. Future manuscripts are encouraged to further understand and improve upon the results presented at the workshop.

II. TERMINOLOGY

The acoustic terminology in this paper follows ISO 18405 (International Organization for Standardization, 2017) where the necessary terms exist. During development of the scenarios several additional terms were needed and are defined here (ref: ADEON Project Dictionary).

Temporal observation window (TOW): interval of time within which a statistic of the sound field is calculated or estimated. For example, averaging the squared sound pressure over 1-min to produce a single sample of sound pressure level.

Temporal analysis window (TAW): interval of time during which statistics are calculated over multiple TOWs. For instance, computing the probability distribution function of 1-min SPLs over a TAW of 1 day.

Spatial observation window (SOW): region of space within which a spatially averaged power quantity is calculated or estimated, for a specified duration of the TOW. For example, a 1 km × 1 km area, at a specified receiver depth, for instance 1 m above the seabed.

Spatial analysis window (SAW): region of space within which statistics are calculated over multiple SOWs. For instance, one could compute the distribution of 1-min SPLs averaged over 1 km × 1 km SOWs within a 50 km × 50 km SAW.

The analysis metric employed in the scenarios is the spatially averaged mean square sound pressure. The mean square sound pressure for the TOW δt is

$$\overline{p}^2 = \frac{1}{\delta t}\int_0^{\delta t} p^2 dt \tag{1}$$

and its associated level is

$$L_p = 10\log_{10}\frac{\overline{p}^2}{p_0^2}\text{dB}, \tag{2}$$

where δt is the TOW duration of 1 min. The spatially averaged mean square sound pressure is then

$$\langle \overline{p}^2 \rangle = \frac{1}{\delta A}\int_{-\delta A/2}^{\delta A/2} \overline{p}^2 dA \tag{3}$$

with a level of

$$L_{p,av} = 10\log_{10}\frac{\langle \overline{p}^2 \rangle}{p_0^2}\text{dB}, \tag{4}$$

where δA is the area of the SOW (1 km × 1 km grid cell).

In some cases, the quantity of interest is for a decidecade band, for example the propagation loss for a decidecade band, $N_{\overline{F}}$ (defined to be the band from f_1 to f_2). $N_{\overline{F}}$ is defined as

$$N_{\overline{F}} = 10\log_{10}\frac{1/\overline{F}}{1\,\text{m}^2}\text{dB}, \tag{5}$$

where

$$\overline{F} = \frac{1}{f_2 - f_1}\int_{f_1}^{f_2} F(f) df, \tag{6}$$

$$F(f) = 10^{-N_{\text{PL}}(f)/(10\,\text{dB})}\text{m}^{-2}, \tag{7}$$

and N_{PL} is the propagation loss for frequency f.

III. MODELING SCENARIOS

Model results were compared using a scenario that estimated the wind and vessel soundscape. The environment was radially symmetric and temporally constant, that is, the propagation loss depended only on distance, and not on direction or time. The scenario generated a variety of metrics for comparison, including time series and histograms of 1-min SPLs at a single location, maps and histograms of the total SPL over the period of one day, and a comparison of the average SPL over a wide area as a function of maximum ship speed.

A. Environmental and modeling parameters

1. Modeling frequencies

In the scenarios the sound pressure levels were computed for the decidecade bands with nominal center frequencies of 63, 630, and 6300 Hz (IEC 61260-1:2014 band indexes –12, –2, and +8, precise center frequencies of 63.096, 630.96, and 6309.6 Hz). The 63 Hz band was

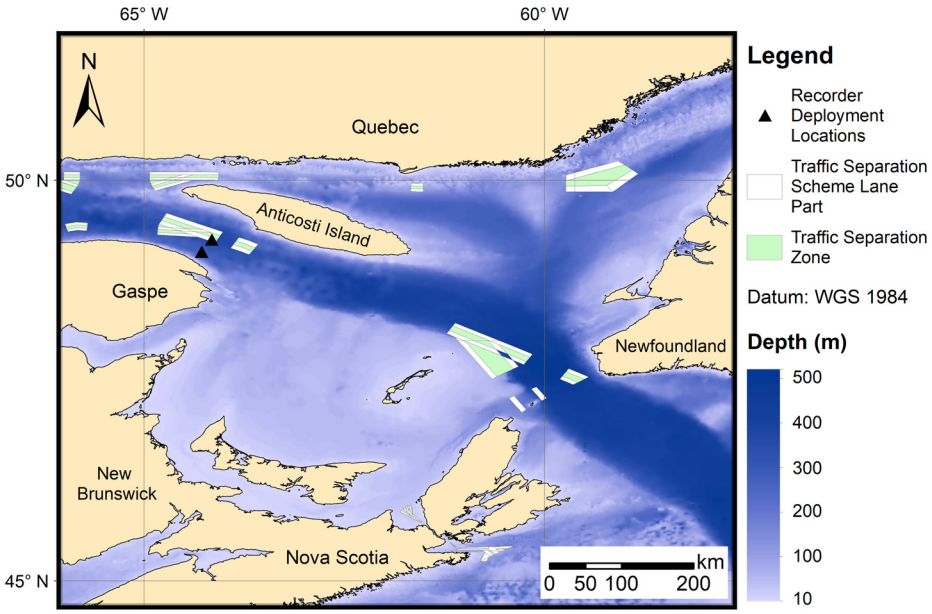

FIG. 2. (Color online) The selected location for ambient sound modeling benchmark scenarios were the shipping lanes between Gaspe and Anticosti Island. Two months of soundscape recordings are available at sites A and B which can be employed to assess the accuracy of soundscape models in future work (courtesy of Transport Canada).

selected as it is near the peak frequency in the spectrum of large vessels. 630 Hz is near the peak frequency of wind driven sound, where vessels still generate substantial sound levels (Wenz, 1962). The 6300 Hz band is dominated by wind and is expected to require a different computational approach than 63 Hz.

2. Location

The selected area for the scenario is encompassed the traffic lanes between the Gaspe Peninsula and Anticosti Island, in the Gulf of St. Lawrence, Canada (Fig. 2). This site was selected because (1) it has a mostly flat seabed with a depth ~320 m, (2) there are two high quality soundscape recordings available for two months in the fall of 2019 that may be used in the future to assess the accuracy of soundscape models, and (3) it has a reasonably high vessel density of ~20 commercial vessels per day passing through the area.

The modeling results were generated for 1 km × 1 km SOWs within a 50 km × 50 km SAW, with detailed analysis results required for the SOWs at site B (Table I). Results were required at depths of 10, 70, and 319 m. The shallow depth of 10 m is relevant for many forms of marine life, 70 m is the sound speed minimum and 319 m is relevant for benthic animals as well as for recorders located near the seabed.

3. Environment

To simplify propagation modeling, the scenario employed a flat bathymetry with a depth of 320 m and geoacoustic properties of a sand-like fluid half space given in Table II. The water column sound speed profile is given in Fig. 3 and provided as a ".csv" file in the supplementary material. Wind generated sound was modeled to establish the background level for ambient sound. Wind speeds every three hours are shown in Fig. 3 and are provided as a ".csv" file in the supplementary material. Modelers were requested to obtain values of the water column sound speed and wind speeds between provided data points by linear interpolation. Volume absorption of sound by seawater was requested to be computed using the approach of Horton (1959) and Thorp (1965), as stated by Fisher and Simmons (1977).

4. Vessel modeling

One goal of this soundscape modeling exercise was to examine how vessel speed affects the sound pressure field. As a starting point for analysis, it is noted that each class of vessels (e.g., tankers, container ships, cruise ships) has a maximum cruising speed (MacGillivray and de Jong, 2021). We assume that if regulators apply a speed limit to vessels that causes them to travel slower than their preferred speeds, the vessel density increases to maintain a constant delivery rate of goods. Assuming unloading and loading are "instant," then the number of ships that arrive at the port

TABLE I. Coordinates for the Spatial Observation and Analysis Windows. Coordinates in UTM zone 20 U. For the SAW, the SOWs were each 1 km^2 inside the SAW. The actual locations of measurements at sites A and B were 49.1042 N, 64.2841 W and 49.2586 N, 64.1588 W respectively.

Area	Easting (m)	Northing (m)
Site A SOW	406 000	5 439 000
	406 000	5 440 000
	407 000	5 439 000
	407 000	5 440 000
Site B SOW	415 000	5 457 000
	415 000	5 456 000
	416 000	5 457 000
	416 000	5 456 000
SAW	395 000	5 433 000
	395 000	5 483 000
	445 000	5 433 000
	445 000	5 483 000

https://doi.org/10.1121/10.0026597

TABLE II. Geoacoustic properties of the seabed.

Material	Density (g/cm^3)	P-wave speed (m/s)	P-wave attenuation (dB/λ)	S-wave speed (m/s)	S-wave attenuation (dB/λ)
Sand	2.216	1890	0.848	0	0

over longer periods of time (e.g., per day or week) must be the same regardless of the speed limit. This implies that the number of ships at sea, i.e., the density, must go up, and is inversely proportional to the speed, scaled by the percentage of the voyage where speed restrictions apply (X). Then the number of ships at sea becomes

$$N_{new} = N_0\left(1 - \left(1 - \frac{V_c}{V_{new}}\right)X\right). \tag{8}$$

The baseline value of vessels, N_0, depends on the class of vessels that use a particular waterway. For simulation purposes, we will "start" vessels of each class every $24/N_{new}$ hours. For the target simulation area in Honguedo Strait we can further assume that (a) cruise vessel counts are not proportional to speed, they will adjust schedules instead and (b) tugs, dredgers, recreational, fishing, naval, research, passenger (ferries) are outside of scope for this exercise. It was assumed that a total of 24 vessels pass when there are no speed restrictions, then the speed limits were applied to assess their effects. The maximum speed of each class was the reference speed of each class as defined in the MacGillivray and de Jong (2021) model. The number of vessels per class was divided arbitrarily for simulation purposes as shown in Table III. Each vessel was then randomly assigned to the inbound or outbound shipping lanes. The simulation lasted for 1 week in September 2019 during which the wind speeds changed substantially (Fig. 3).

Modelers were provided with .csv files that mimic AIS data and contain the tracks to be simulated (see supplementary material). There was one .csv file per speed limit to be simulated (6, 9, 12, 15, and 18 kn,[1] where 18 kn is effectively no speed limit since it is the maximum speed of any vessel class). Vessels with reference speeds above the speed limit were slowed to the current speed limit. For example, at the 15-kn limit, containerships, cruise ships and vehicle carriers were slowed to 15 kn. The .csv files contain a track point every minute; linear interpolation of the positions were used if finer vessel position data were required by a modeling group's approach.

The vessel source levels model proposed in MacGillivray and de Jong (2021) was employed. This model

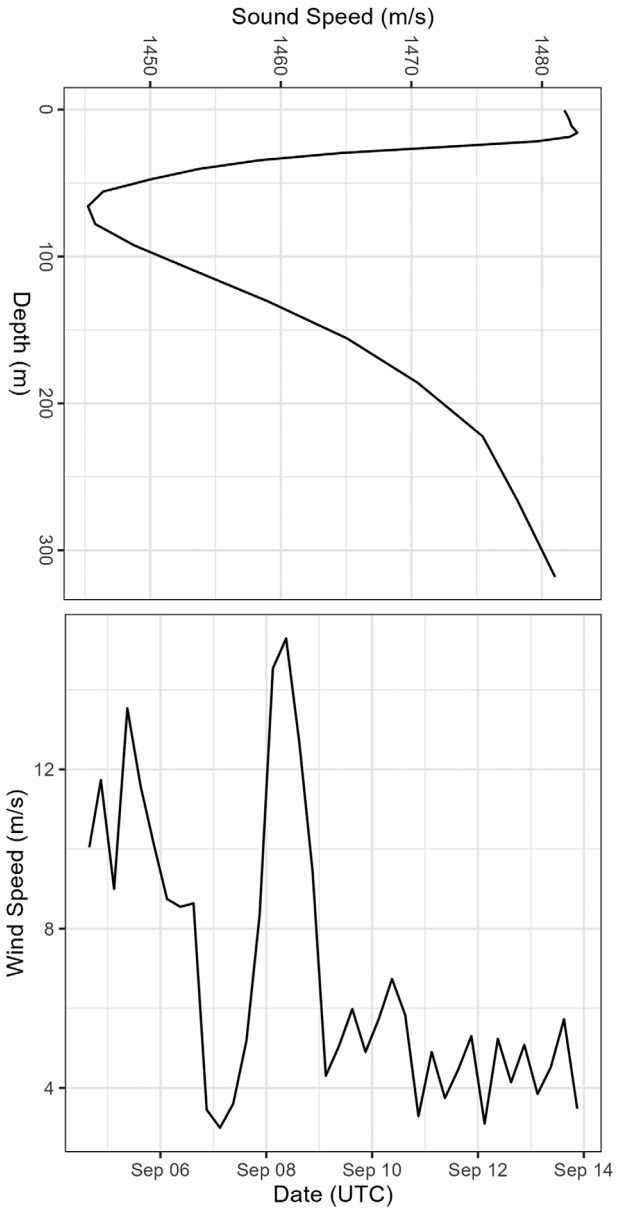

FIG. 3. Environmental data: (Top) Sound speed profile for benchmark scenario. The sound speed profile is interpolated linearly between the points provided. The SSP is assumed to be constant throughout the 1-week TAW. (Bottom) Wind speed history at 10 m height above the sea surface to be used for modeling wind driven sound. The time resolution of the wind data is 3 h. The wind speed for intermediate times should be obtained by linear interpolation between the measured values. The wind speed was assumed to vary in time only—i.e., it had the same value at all geographical positions.

TABLE III. The number of vessels per class per day along with their standard speeds. The reference speeds are from MacGillivray and de Jong (2021) model.

Vessel class	AIS SHIPTYPE ID	Reference speed (kn)	Class N_0 per day
Bulker	70, 75–79 (speed $\leq$ 16 kn)	13.9	6
Containership	71–74 (all speeds); 70, 75–79 (speed $>$ 16 kn)	18.0	11
Cruise	60–69 (length $>$ 100 m)	17.1	1
Tanker	80–89	12.4	4
Vehicle Carrier	n/a	15.8	2

TABLE IV. Summary of the Soundscape Scenario. The requested analysis frequencies were for $N_{\bar{F}}$ at the center frequencies for the decidecade bands listed, i.e., 63.096, 630.96, and 6309.6 Hz. As noted in Sec. III A 4, for vessels whose cruising speed is less than 18 kn, their maximum speed is used instead.

Step	Output	TOW	TAW	SOW	SAW	Analysis Freq. (Hz)	Analysis Depths (m)	Maximum Vessel Speeds (kn)	Outputs
1	Propagation Loss	n/a	n/a	n/a	n/a	63 630, 6300	10, 70, 319	n/a	PL vs distance
2	SPL vs time at site B for three frequencies	60 s	n/a	1 km × 1 km	n/a	63 630, 6300	10	18	Wind Only SPL, Ships Only SPL, Total SPL
3	Daily SPL for all SOWs	60 s	1 d	1 km × 1 km	n/a	63 630, 6300	10, 70, 319	18	Total SPL and Excess SPL, 8 and 10 Sep
4	SPL vs speed limit	1 d	n/a	50 km × 50 km	n/a	63 630, 6300	10, 70, 319	6 -18 knots in 3 knot steps	Total SPL, 8 and 10 Sep

generates an average source level that depends on the class, length, and speed of the vessel, and indicates that a standard deviation of 6 dB can be expected about those means. The source depth for all vessels in the model is 6 m. In the simulations, the length of the vessels was randomly generated within the bounds of 10% of the mean length provided in MacGillivray and de Jong (2021). The vessel track files include the decidecade band source level for each vessel which was randomly generated within two standard deviations of the class means. The speed-to-the-power-of-6 relationship assumed in the J-E source model is valid when vessels are cavitating, which means that the model may over-predict sound level reductions for speeds of 6 and 9 knots. In reality, radiated sound from vessels is dominated by engine noise below cavitation inception, which (approximately) follows a speed-cubed relationship (Ross, 1987).

5. Modeling of underwater sound from wind

Modelers were requested to use the ambient sound model of their choice for wind generated sound. There are two approaches to modeling the ambient sound that are commonly used: either multiple point sources that are each propagated individually, or an areic sheet source. The specified areic dipole source spectral density level model ($L_{A,dp}$) is provided below [Ainslie (2010), Eq (8.206)],

$$L_{A,dp} = 10\log_{10}\left(\frac{10^{4.12}}{\left(1.5+\left(\frac{\hat{f}}{1000}\right)^{1.59}\right)}\hat{v}^{2.24}\right)\text{dB}. \quad (9)$$

In this equation $\hat{f}$ is the center frequency of the decidecade, $\hat{v}$ is the wind speed at 10 m above the water surface in m/s. Rain noise was outside of scope for this exercise. The conversion of the areic dipole source level to an areic monopole source level is discussed in the papers by Chapman *et al.* (2024) and Siderius *et al.* (2024).

B. Soundscape from wind and vessels scenario

The soundscape scenario defined four steps that were sequentially more complex to generate outputs that could be compared between modeling teams (Table IV). The first step was a comparison of the modeled propagation losses, followed by the SPL versus time at one site, the daily average SPL at each SOW, and the time and spatially averaged SPL as a function of the vessel speed limit. The steps are described in this section. For each step the modelers were requested to provide their results as formatted .csv files to facilitate comparisons between groups. The formats and comparison software are provided in the supplementary material.

1. Comparison of propagation losses

Participants were encouraged to provide propagation loss files for comparison of this fundamental output as part of the comparison between results. The suggested outputs included the coherent losses at the decidecade center frequencies as well as the decidecade averages (Table V).

2. Sound pressure level time-series within a SOW

The first soundscape metric was the time series of the sound pressure levels in the 1 km^2 SOW at site B (Table I) for the week-long vessel tracks at 18 knots and wind speed timeseries. A time series sampling rate of one result every 10 min was suggested. Individual results for the three decidecade bands (nominally centered at 63, 630, and 6300 Hz) were requested for receiver depths of 10 m. Results for just the wind, just the ships, and the combined wind and ships were requested.

TABLE V. Suggested propagation loss calculations for comparisons between teams. Results for water depths of 10, 70, and 319 m were requested.

Case	Notes	Analysis Freq. (Hz)
1: Flat-surface (no surface loss) propagation loss (dB) Ranges 0 to 100 km	Source depth = λ/4 λ=1500/freq.	63.096, 630.96 and 6309.6 (single frequencies)
2: Flat-surface (no surface loss) propagation loss (dB) Ranges 0 to 100 km	Source depth = 6m	63.096, 630.96 and 6309.6 (single frequencies)
3: Flat-surface (no surface loss) propagation loss (dB) Ranges 0 to 100 km	Source depth = λ/4 λ=1500/freq.	Decidecade averages centered at 636 306 300
4: Flat-surface (no surface loss) propagation loss (dB) Ranges 0 to 100 km	Source depth = 6m	Decidecade averages centered at 636 306 300

https://doi.org/10.1121/10.0026597

3. Spatial sound pressure level statistics

The next model results were spatial results, averaged over one day. The TAW of one day each were 8 Sep (high wind day) and 10 Sep (low wind day—see Fig. 3). The SOWs were 1 km × 1 km squares in the 50 km × 50 km SAW (Table I). Modelers computed the SPLs from wind only and from vessels only, then added them to find the total daily SPL, and subtracted the wind from the vessel SPL to find the excess SPL (in dB). These results were requested for all three receiver depths, for the 18-kn speed limit case, for all three decidecade center frequencies. The results were plotted as maps as well as probability distributions.

4. Sound pressure level vs speed limit

The last modeling exercise found the spatial average across the entire analysis area (i.e., using a SAW of 50 km × 50 km) for a TAW of 1-d for each of the speed limits (6, 9, 12, 15, and 18 kn) for the high-wind day of 8 Sep 19 and the low wind day of 10 Sep 2019. Results were requested for each of the decidecade bands and all three water depths. The results were plotted as average SPL vs speed; this was expected to provide a metric for assessing how speed reductions would limit sound exposures.

IV. RESULTS

Eight modeling teams performed elements of the Soundscape Scenario and submitted results for comparison (Table VI). To compare the results of multiple modeling teams, a MATLAB Live-Editor script was written. The script automatically reads the results from each team, assuming it is in a single directory per team, and the results files are in a standardized format. The details of these files formats are included in the supplementary material, along with the script and results that were submitted for this manuscript.

A. Propagation modeling

The full range of propagation modeling methods were employed, with parabolic equation methods being the most common at low frequencies and raytracing, fast field, hybrid energy-flux and normal mode models being added at the mid and high frequencies. A wide variety of approaches to computing the decidecade average propagation losses were also used. The modeled results compared well for ranges of 100–40 km (Fig. 4). Close to the receiver, the normal modes method (group A) predicted higher PL than the PE methods. The PE method used by group E had a very coarse range resolution which affected the quality of their results at longer ranges. The fast-field model SCOOTER employed by group E appears to have overpredicted the PL.

Multiple methods of obtaining the decidecade average propagation losses from single frequency values were employed. groups A and F computed the propagation loss for multiple frequencies in each band and then averaged them, as shown in Eq. (6). Most other groups used some variation on the averaging of the propagation losses in range as a surrogate for averaging in frequency as recommended by Harrison and Harrison (1995). The use of a small (5%, group E) range smoothing window to estimate the decidecade PL resulted in a rapid oscillation in the PL compared to the other estimation methods. Zykov and Martin (2024) explored the decidecade averaging methods further and recommended the use of Gaussian range averaging. Group C provided propagation loss calculations based on the incoherent mode sum at the center frequencies of decade bands without any averaging over range and frequencies. Teams that included seawater absorption in the propagation loss estimation used the Horton-Thorp method as requested. The effects of absorption are not apparent at 63 Hz in Fig. 4, but were present at 6300 Hz with differences of ∼30 dB at 30 km between those that included and omitted absorption.

B. Step 2: Sound pressure level time series within a SOW

Modeling of the time-series of sound from the vessels yield very similar results between research groups at 63 Hz (Fig. 5, left), with larger differences between groups at 6300 Hz, especially when the vessels were further away (Fig. 5, right). Group E levels were consistently lower than the others, as expected from the higher PL values (Fig. 4). There is also more jitter in the Group E results, likely due to the 5% range smoothing to estimate the decidecade average PL instead of the wider windows used by other groups. The increased PL (and hence lower high frequency sound levels) estimated by group B at 6300 Hz was caused by including the time-dependent surface scattering (which was not included in Fig. 4). It is not clear why group D had significantly higher levels when vessels were farther away.

The vessel sound pressure levels at 63 Hz show slight differences in levels at the closest points of approach and between the group A and group C KRAKEN results when vessels were further away. These are likely due to how the groups approached the distillation of the modeling results from the modeling resolution to the SOW, SAW, TOW, and TAW resolutions (Table VII). Although group C-SOPRANO results used an isovelocity assumption (c = 1500 m/s), SPL results are similar to other models, which was explored previously for shallower waters in Sertlek *et al.* (2016). Group A used a 1 s time resolution and 25 m range resolution for their modeling, then averaged down to the 1-min and 1 km^2 TOW and SOW; the effect of their higher resolutions can be seen by the much higher average SPL obtained at the closest point of approach for vessels, especially for the very close CPA at ∼16:00 on 5 Sep 2019 in Fig. 5.

Systematic convergence tests were performed at site B, using the group D model, to assess the effect of spatial and temporal sampling resolution on the sound level predictions (Fig. 6). Temporal resolution was evaluated for five different sampling intervals (60, 120, 240, 480, and 600 s) over a 1-week TAW, using a spatial resolution of 250 m. Spatial resolution was evaluated for 4 different grid resolutions (100, 250, 500, and 1000 m) over a 1-day TAW, using a temporal resolution of 60 s. The convergence tests showed

TABLE VI. Summary of the methods employed by the modeling teams. PE is Parabolic equation propagation model. KRAKEN\KRAKENC (normal modes), BELLHOP (ray tracing), RAMGEO (a version of PE), SCOOTER (fast-field solution), and SOPRANO (hybrid mode energy flux) are all open-source propagation models that were used by the different teams. SSP is sound speed profile. BW is bandwidth. JOMOPANS is the Joint Monitoring Programme for Ambient Noise in the North Sea. The Horton-Thorp absorption method was used by all modeling teams. Group H performed this analysis as input to the Detection Scenario (Oppeneer *et al.*, 2024) but did not provide results for plotting for this Soundscape Scenario. Note 1: These teams averaged frequencies within the decidecade bands at 0.01 Hz resolution @63 Hz, 0.1 Hz resolution @630 Hz and 1.42 Hz resolution at 6300 Hz.

Group	Wind modeling method	Decidecade averaging method	SSP in wind model?	SSP in vessel model?	Modeling method as a function of frequency band			Included surface roughness in modeling?		
					63 Hz	630 Hz	6300 Hz	63 Hz	630 Hz	6300 Hz
A	OASES + Corr Factor (from ISO velocity test)	Note 1	Yes	Yes	KRAKEN	KRAKEN	KRAKEN	No	No	No
B	AREIC-PSD Ray trace Surface Scatter	Range averaging 25% boxcar	Yes	Yes	PE	RAY	RAY	No	Yes	Yes
C	WRASP (AERIC PSD)	No averaging (Incoherent mode sum)	No	SOPRANO: N KRAKENC: Y	SOPRANO KRAKENC	SOPRANO + KRAKENC	SOPRANO + KRAKENC	No	No	No
D	WRASP (AERIC PSD)	Range averaging 13% Gaussian	No	Yes	PE	BELLHOP	BELLHOP	No	No	No
E	JOMOPANS equation including surface scatter	Avg center + edge; 5% log space	No	Yes	RAMGEO (for grazing angles $\leq 80°$) SCOOTER/FIELDS (for grazing angles $>80°$)	RAMGEO (for grazing angles $\leq 80°$) SCOOTER/FIELDS (for grazing angles $>80°$)	BELLHOP (for grazing angles $\leq 85°$) SCOOTER/FIELDS (for grazing angles $>85°$)	No	No	No
F	AREIC-PSD PE + 6 dB corr	Note 1	Yes	Yes	RAMGEO	RAMGEO	RAMGEO	No	No	No
G	AREIC-PSD + 6 dB corr.	Mean square across band (21 freqs)	Yes	Yes	PE	PE	PE	No	No	No
H	AREIC-PSD Ray trace	Range averaging 25% boxcar	Yes	Yes	RAM	RAM RAY tracer energy flux model mode-stripping	RAM RAY tracer energy flux model mode-stripping	No	No	No

https://doi.org/10.1121/10.0026597

JASA

https://doi.org/10.1121/10.0026597

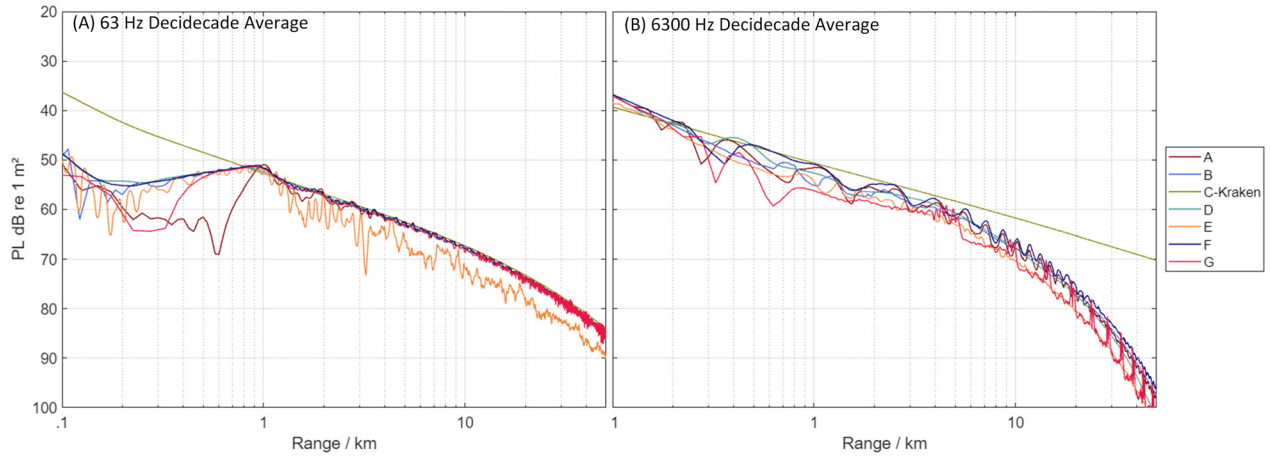

FIG. 4. (Color online) Decidecade average propagation losses for (A) the 63 Hz and (B) the 6300 Hz bands, for a source depth of 6 m and a receiver depth of 319 m. Group C results show incoherent mode sums without averaging over range or frequency; no absorption was employed in this version of the Group C results.

that the SEL at site B was quite sensitive to the choice of time sampling interval, with the result only converging for sampling intervals of 240 s or less. On the other hand, the SEL at site B was relatively insensitive to the choice of spatial grid resolution. Likewise, distributions of the modeled SPL were relatively insensitive to the choice of spatial and temporal sampling resolution.

The modeling of the sound levels from wind yielded the greatest differences between groups and led to additional investigations (Fig. 7). Chapman *et al.* (2024) explore the impact of assumptions made about the wind sound source in papers by Burgess and Kewley (1983) and Chapman and Cornish (1993) that reported estimates of source levels for low frequencies (<400 Hz) derived from experimental data obtained using vertical line arrays. The analysis provides a resolution for the roughly 6 dB difference between the two sets of reported values in terms of the assumptions about the source level—whether distributions of monopoles near the sea surface [as in Burgess and Kewley (1983)], or sheets of dipoles at the surface [as in Chapman and Cornish (1993)]. The reported values in these papers provide experimental benchmarks for the source levels.

Siderius *et al.* (2024) determined there was about 8 dB difference in results for wind-generated sound pressure levels provided by the participants (as shown in Fig. 7). The differences were caused by the various approaches taken by the participants and the underlying assumptions each made. For example, while all participants used the same definition for source level as a function of wind speed and frequency, subtle differences in the interpretation and implementation of the surface sound can create several decibel discrepancies. Much of the discrepancies were due to participants modeling the wind generated sound using propagation modeling of point sources (monopoles) distributed near the surface and mis-applying the provided source level model that assumes a point-dipole surface distribution. This led to all results being several dB lower than the benchmark solution provided in Siderius *et al.* (2024).

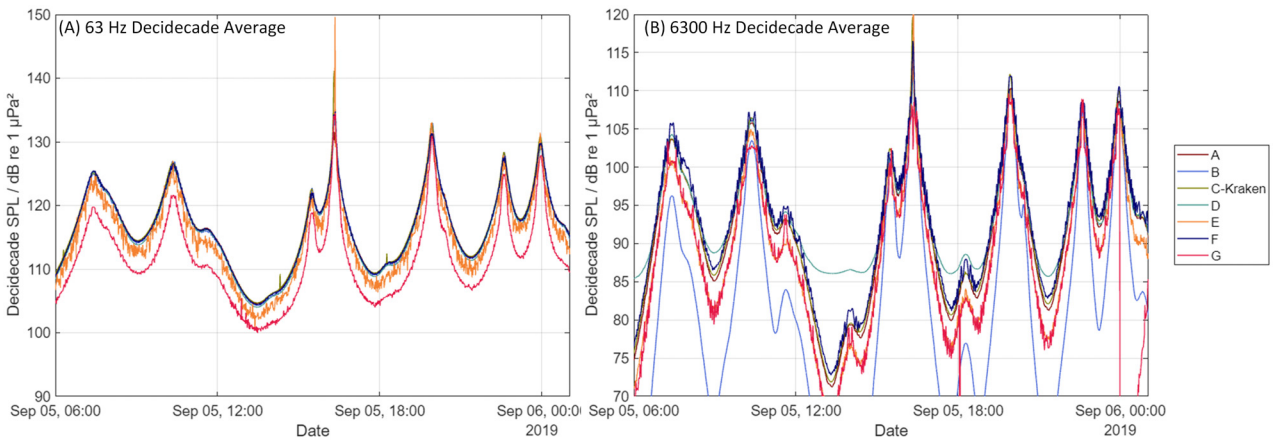

FIG. 5. (Color online) Comparison of the shipping sound levels predicted by the different modeling teams (and models) for (left) 63 Hz and (right) 6300 Hz, for a depth of 10 m and the 18-kn vessel speed limit condition. The increases and decreases are the result of the simulated vessels getting closer to and further from the measurement site. All times/dates are UTC.

TABLE VII. Summary of how the modeling teams performed the soundscape modeling for Step 2. The model Extent is the distance outside of the SAW that was included in the total modeling. Group H performed this analysis as input to the Detection Scenario (Oppeneer *et al.*, 2024) but did not provide results for plotting for this Soundscape Scenario.

Group	Modeling Time resolution (e.g., vessel and wind time step)	Method used to collapse to 1-minute TOW	Modeling range resolution	Model Extent.	Method used to collapse to 1 km SOW
A	1 s	Wind and vessels were first interpolated onto a 1 s time series for computations and then a 1-minute moving average was used	25 m	Modeled 100 km from all vessel positions, including those outside the SAW	Squared-sound pressure values were interpolated from vessel positions to center of 1 km SOW
B	10 min	Wind: interpolation to center of TOW Vessels: SPL at average distance from line segment (obtained by interpolation of vessel track to start and end time of TOW) to SOW square	15.3 m	153 km	SPL at average distance from line segment (obtained by interpolation of vessel track to start and end time of TOW) to SOW square
C	1 min	Wind and vessel data are interpolated into a time series with a 1-minute resolution	1 m		Squared-sound pressure is averaged over the receiver grids
D	1 min	Wind: Linear interpolation of wind speed between two data points (i.e., within TOW). Vessels: interpolation of position based on positions versus time.	Propagation model: 20 m. Sampled at an output grid resolution of 100 m	Propagation model maximum distance: 100 km. Grid extent: 200 km x 200 km	Averaged the square pressure of each grid cell within the SOW; converted back to SPL.
E	1 min	We interpolated wind and vessels to 1 min and modeled that resolution, rather than model finer and average	PL was modeled in 1 m range resolution. To reduce data, we then subsampled logarithmically to ranges: [10:10:1000, 1100:100:10000, 11000:1000:100000] m.	Only modelled within the SAW	We computed SPL at B and at the 4 corners of the 1 km x 1 km SOW surrounding B, and then took the (linear) average of mean square pressure at the 5 points.
F	1 min	Interpolated wind and vessels to 1 min resolution and modeled at that resolution.	Propagation runs were on a [60,6,0.6]m range grid by [1,0.1,0.01]m depth grid (per band); written out at 60 m range increments. Pressures were interpolated to the ship range for the time series	Modeled 100 km from all vessel positions, including those outside the SAW	We computed SPL at B and at the 4 corners of the 1 km x 1 km SOW surrounding B, and then took the (linear) average of mean square pressure at the 5 points.
G	1 min	Computed the wind and vessel SPL at 1 min resolution.	50 m	Modeled 100 km from all vessel positions, including those outside the SAW	Averaged the PL over 1 km.
H	1 min	Interpolated wind and vessels to 1 min resolution and modeled at that resolution.	100 m	Propagation model maximum distance: 110 km	Squared-sound pressure were interpolated from vessel positions to center of 1 km SOW.

C. Step 3: Spatial sound pressure level statistics

The groups computed the daily mean SPL for all SOWs within the SAWs, which could then be represented as surfaces (Fig. 8). There are substantial differences between the results, both close to the traffic lanes and further away. Focusing on the traffic lanes, the SPL predicted by group A has peaks along the lanes, whereas groups C, D, and F are smoother, while B and E are smooth but more diffuse. These differences in the smoothness of the traffic lane lines are likely due to the time or spatial averaging choices (see Tables VII and VIII). The difference in levels further from the traffic lanes for group B is likely due to the increased propagation loss since this group include wind driven scattering, while group E's lower levels are likely due to the higher propagation losses they predicted in general (Fig. 4).

https://doi.org/10.1121/10.0026597

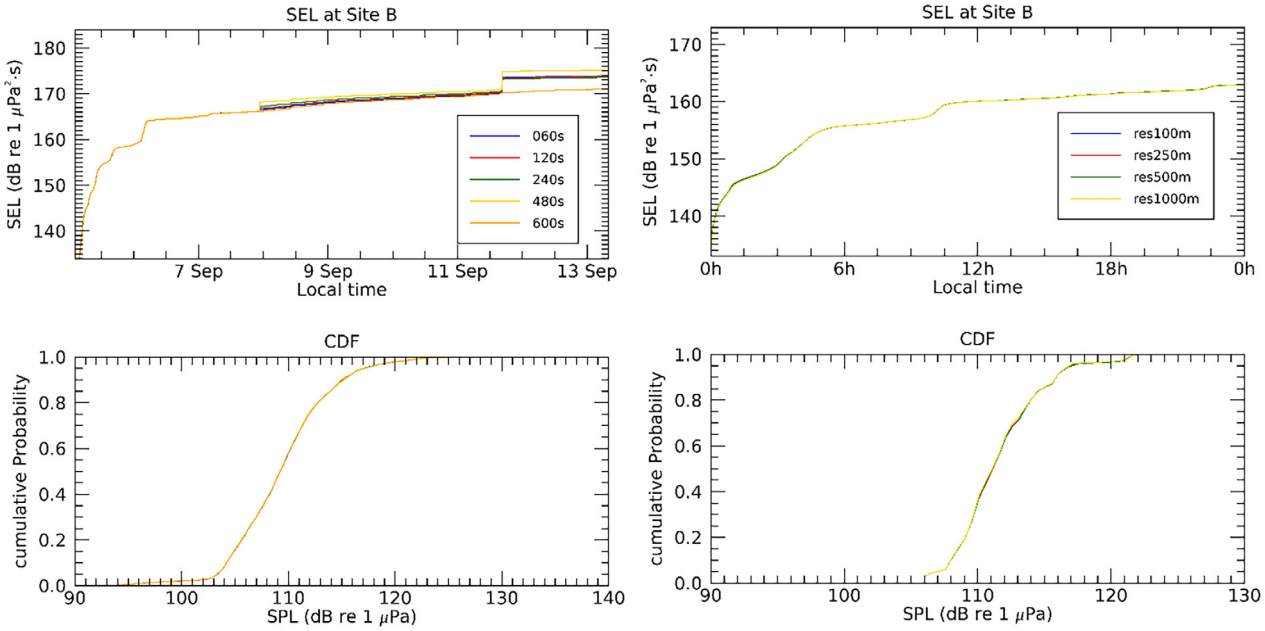

FIG. 6. (Color online) Statistics of modeled sound levels at site B (630 Hz, 10 m receiver depth, 1 km^2 SOW) for different temporal (left) and spatial (right) sampling resolutions using the group D model. The difference in temporal resolution was modeled using a spatial resolution of 250 m; the difference in spatial resolution was model using a temporal resolution of 60 s. The top plots show cumulative SEL versus time. Jumps in SEL versus time correspond to the closest approach of two vessels to site B. Bottom plots show cumulative distributions functions (CDF) of the modeled SPL.

Figure 9 compares distributions of modeled sound pressure levels for the low-wind day of 10 Sep 2019. The distributions are bimodal, with a peak near 100 dB attributed to SOWs distant from the shipping lanes, and a second peak in the 102–104 dB range attributed to the closer shipping traffic. There are clear differences in the high amplitude tails of the distributions. Group A has the longest tails, which are attributed to that group employing a 1-s temporal modeling resolution and 25 m spatial grid, which resulted in more data samples with the vessels at short ranges, increasing the average sound pressure levels in the SOW for each TOW. The group B results are lowest due to the inclusion of surface scattering in the propagation loss calculations. The additional low-amplitude peak in group E's distribution is attributed to their increased propagation losses. The differences between group G and the others are likely related to their higher wind driven sound levels (Fig. 7).

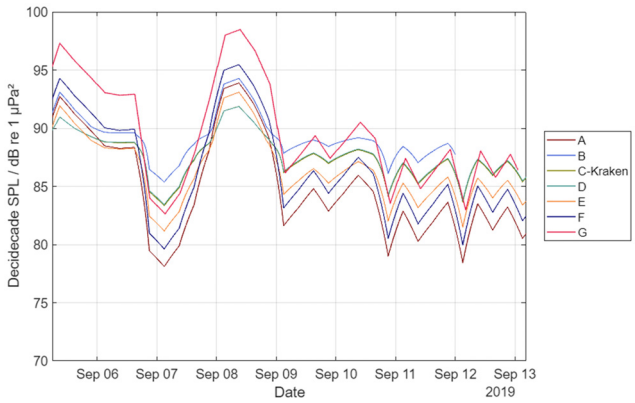

FIG. 7. (Color online) 630 Hz decidecade sound pressure level from wind at 10 m water depth as predicted by the groups.

Large-scale convergence tests were performed, using the group D model, to assess the effect of spatial and temporal sampling resolution on the spatially averaged SPL over the 50 km × 50 km SAW. The spatially averaged SPL was equal to the decibel level of the mean squared sound pressure averaged over all sampling points within the SAW and TAW. The convergence tests showed that spatially averaged SPL was relatively insensitive to the choice of temporal sampling interval, with a difference of less than 0.2 dB between 1-min and 10-min sampling intervals over a 1-week TAW (Table IX). On the other hand, the convergence tests showed that spatially averaged SPL was very sensitive to the choice of spatial grid resolution, with differences of 14–16 dB between the finest and coarsest grid resolutions for 10 m receiver depth. However, the difference in the spatially averaged SPL was not observed at deeper receiver depths (70 and 319 m, not shown), suggesting that the rate of convergence was related to the distance of the receivers from the sources of vessel and ambient noise near the sea-surface (Table X).

D. Step 4: Sound pressure level vs vessel speed limit

The final product of the workshop modeling were the plots of the dependence of sound pressure levels on speed of the vessels (Fig. 10). The results highlight there are increased sound levels from vessels in the shallow sound speed duct near the surface (Fig. 3), as well as at the seabed, compared to the sounds speed minimum at 70 m. There is a clear dependence of vessel sound levels on speed, which

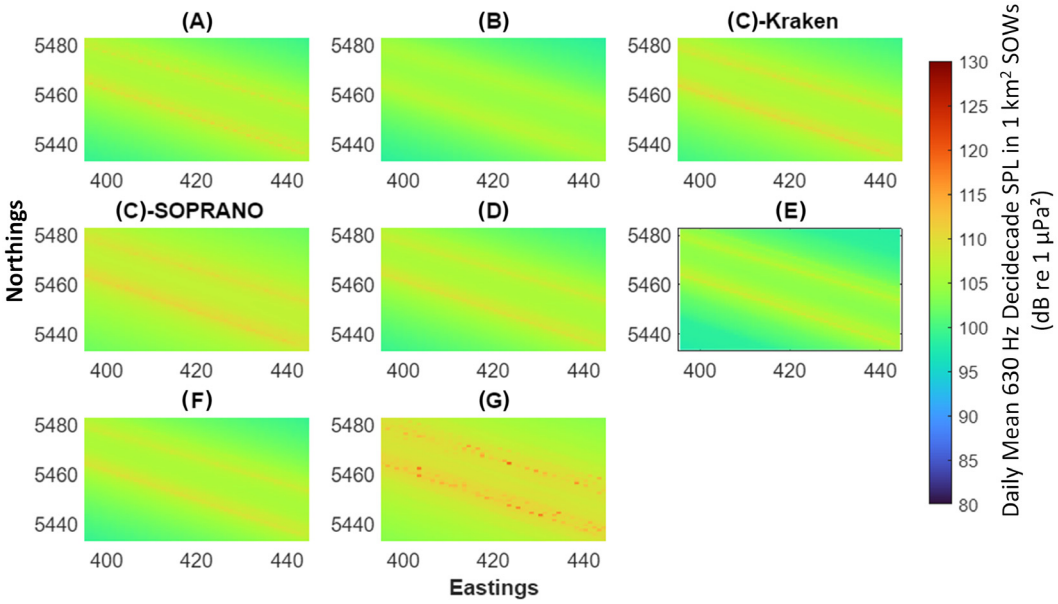

FIG. 8. (Color online) Daily mean SPL in the 630 Hz decidecade band at 10 m depth for 10 Sep 2019.

reaches a constant when the vessels are travelling at their maximum preferred speeds (no further changes in sound levels would occur after 18 knots). Since the vessel source model (MacGillivray and de Jong, 2021) is valid when vessels are cavitating, the lines in Fig. 10 are dashed below 11 kn to indicate where the sound levels are not valid and have been included for illustration of the effects. These results demonstrate how modeling can be employed to understand the impacts of changing vessel operational parameters on the soundscape.

V. DISCUSSION AND CONCLUSION

Soundscape modeling has applications to specific activity-focused management (i.e., assessing the impact of a singular transient human activity), as well as fields such as passive acoustic monitoring. However, ecosystem management frameworks, which consider cumulative and chronic impacts, stand to gain the most. In this context, assessing contributions of human activity over large spatial and temporal scales can enable and support a wide variety of management and decision-making. Accurate soundscape modeling, along with robust long-term field measurements, allow evaluation of noise mitigation strategies, enable the implementation of noise related targets such as the European Union's "Good Environmental Status" (European Commission, 2008), improve understanding of how animals cope with noise, and generally increase the understanding and awareness of human noise sources and their relative contribution to the marine soundscape (Gedamke *et al.*, 2016). Additionally, soundscape models can form the basis of a more complete risk assessment tool that combines sound maps, species densities, and species-specific acoustic sensitivities (Gedamke *et al.*, 2016). Such tools and frameworks have increased utility when accompanied with robust verification, validation and quantification of uncertainty, particularly in the context of risk assessment. Wind and vessels are the two sources with the widest spatial distributions whose sound levels must be estimated to model the baseline ambient soundscape.

This manuscript reports on a model verification workshop held in July 2022 in conjunction with the *Effects of Noise on Aquatic Life* conference. Eight groups of modelers submitted results for a Soundscape Scenario that estimated the sound pressure level from wind and vessel passages at a 1-min resolution in grid-cells 1 km × 1 km. The models were run for a 1-week period with a spatial area of

TABLE VIII. Summary of how the modeling teams summarized the daily mean sound pressure levels for each SOW.

Group	Time resolution of modeling	Method used to collapse to daily TAW
A	1 min	Obtain results for each SOW and TOW as above; compute mean square pressure over TAW
B	1 min	Obtain results for each SOW and TOW as above; compute mean square pressure over TAW
C	1 min	Squared sound pressures are averaged over entire map
D	1 min	Averaged the square pressure of each grid cell over the TAW to get a grid of daily mean squared sound pressure; averaged the square pressure of each grid cell within the SOW; converted back to SPL
E	1 min	Sum mean square pressure, divide by 60 × 24
F	1 min	Sum mean square pressure, divide by 60 × 24
G	1 min	Computed mean square pressure over all SOW and TOW

https://doi.org/10.1121/10.0026597

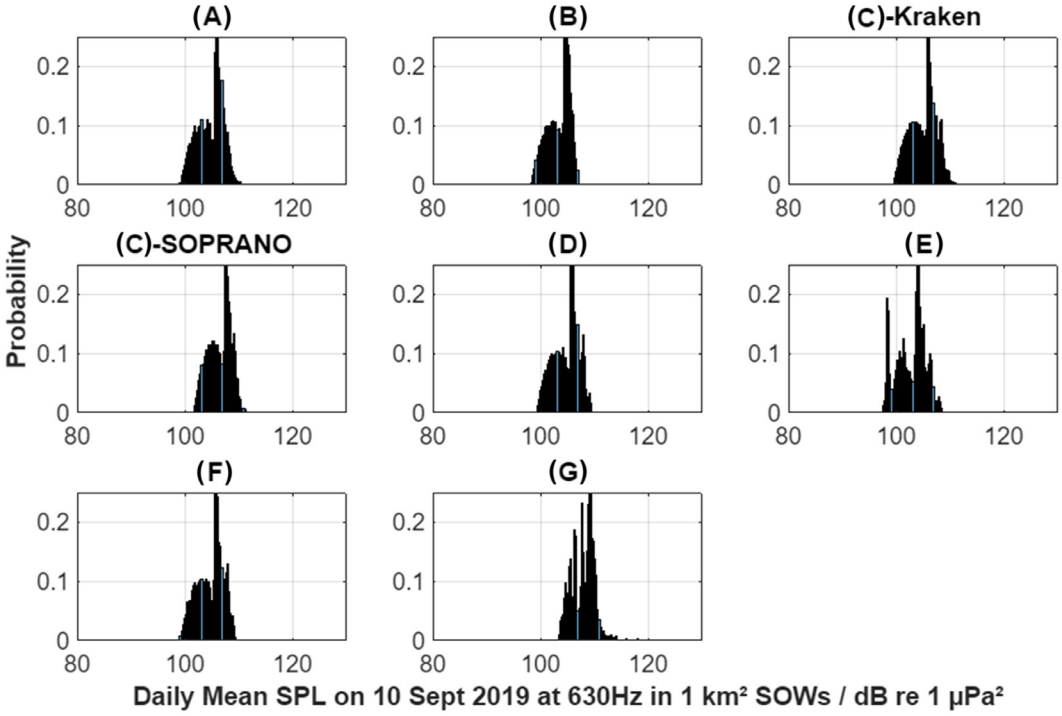

FIG. 9. (Color online) Comparison of the probability distribution functions of the 1 km SOWs, computed with 1-min TOWs, with a daily TAW and 50 km × 50 km SAW. This data is for 10 m depth and 630 Hz, on 10 Sep 2019 (low wind day).

50 km × 50 km. Three groups of modelers also completed a Detection Scenario where the objective was to estimate how far away a submerged target could be detected in noise that varies in time and space. The modelers were provided with wind speeds at a temporal resolution of three hours and vessel tracks at a temporal resolution of 1 min. The remainder of the physical environment was simplified to have a flat sand seabed at 320 m depth, with a sound speed profile that had a shallow duct at ∼15 m and a stronger and deeper duct at 70 m. It was expected that simplifying the propagation environment would make differences in the results depend on the choices made in the soundscape modeling rather than in the propagation modeling, since the propagation losses would be similar for all models.

Soundscape modeling differs from some other types of acoustic modeling in that the sound levels, and hence the propagation losses, must be estimated for broad frequency bands. In the model verification scenarios employed here, decidecade bandwidths were employed. This is typical since the source factor (level) of many human sound sources are reported in decidecade bands, as are the hearing thresholds of marine life. Despite attempts to minimize differences between groups in the propagation modeling, the choices made to obtain decidecade propagation losses (PL) from singe frequency values were the single greatest source of variability between the modeling groups (Table VI). Two basic approaches may be used: averaging PL in frequency within a decidecade band, or modeling at the band center frequency and employing a range average as suggested by Harrison and Harrison (1995) which is computationally more expedient. Zykov and Martin (2024) compared these options. When averaging frequencies, at least 70 frequencies evenly distributed across the decidecade band were found to be required to obtain a stable estimate of PL. When applying the range averaging method, the optimal smoothing window is a Gaussian with a width between 10 to 16% of the range from the source, switching to a constant value beyond 20 km from the source.

Modeling sound from wind yielded substantially different results between groups. As a result, Chapman *et al.* (2024) investigated the source level for wind and resolved the differences between the source level model of Burgess and Kewley (1983) and Chapman and Cornish (1993) which brought the two models into agreement. Siderius *et al.* (2024) provide a derivation of how to apply the areic source level model for wind driven underwater sounds and a reference solution for the soundscape scenario.

TABLE IX. Spatially averaged SPL for a 50 km × 50 km area centered on site B (10 m depth) computed using different temporal sampling intervals with the group D model. The spatial sampling resolution was every 250 m within the 1 km^2 SOW (i.e., 16 points).

	Spatially averaged SPL (1 week, dB re 1 μPa^2)		
Sampling interval (s)	63 Hz	630 Hz	6300 Hz
60	125.73	115.42	104.04
120	125.79	115.46	104.09
240	125.75	115.43	104.06
480	125.74	115.43	104.05
600	125.56	115.29	103.88

TABLE X. Spatially averaged SPL for the 50 km × 50 km SAW (10 m depth) computed using different grid resolutions with the group D model. The temporal sampling interval was 1 min. Note that the extent of the model area was reduced from 200 km × 200 km to 70 km × 70 km for 50 m grid resolution, due to computer memory constraints, which may have reduced the SPL by excluding some distant vessels (> 20 km) from the edge of the 50 km × 50 km evaluated area.

	Spatially averaged SPL (1 day, dB re 1 μPa^2) 10 m water depth			Spatially averaged SPL (1 day, dB re 1 μPa^2) 319 m water depth		
Grid resolution (m)	63 Hz	630 Hz	6300 Hz	63 Hz	630 Hz	6300 Hz
50	120.66	112.54	99.85	120.77	90.77	94.92
100	122.11	113.50	101.00	120.84	111.53	94.92
250	126.61	116.37	105.04	120.84	111.52	94.90
500	132.24	121.09	110.47	120.82	111.51	94.90
1000	138.09	126.66	116.27	120.81	111.61	94.91

In this model verification exercise, the sound level at any point were the sum of the ambient sound level, which depended only on wind speed and depth, and the vessel sound levels which depended on the distance from each vessel in the scenario to the receiver points. Since the vessels are moving, the modeling teams had to make choices on how often to model the sound radiated by the ships, and in the ranges and depths at which the sound levels were stored for later summing to obtain the sound pressure levels at the required temporal (1 min) and spatial (1 km^2) observation windows. The spatial step size of the models in range and depth are dependent on the frequency modeled and the inner working of the model, however, they are generally at a much higher resolution than is needed for the soundscape estimates, and storing the model results at their full resolution, even temporarily, has a substantial impact on computer resources. Thus, the modelers must choose a lower resolution in time and space to store for each source (vessel), which can then be summed across to obtain intermediate results, which are then averaged in time and/or space to obtain the final TOW and SOW outputs. The choice of temporal resolution and spatial resolution for the intermediate outputs has a substantial impact on the variability of the results.

Convergence testing suggested that the soundscape model predictions can be sensitive to the choice of temporal sampling interval and spatial grid resolution. The sensitivity was greatest for those metrics proportional to total sound energy (SEL and spatially averaged SPL) and for shallow receivers close to source of ambient noise near the surface. Statistical distributions of SPL were much less sensitive to the spatial and temporal resolution of the simulation. Likewise, deeper receivers were less sensitive to the spatial and temporal resolution of the simulation. The convergence testing suggested that acceptable accuracy could be achieved for all soundscape modelling scenarios with a maximum sampling interval of 60 s and a maximum grid resolution of 100 m.

The environment in this model verification exercise was radially symmetric and temporally constant. For future model validation exercises that compare against data, or for broader area modeling of real environments, it is expected that range dependent differences in depth, bottom types, and sound speed profiles will add an extra dimension to the

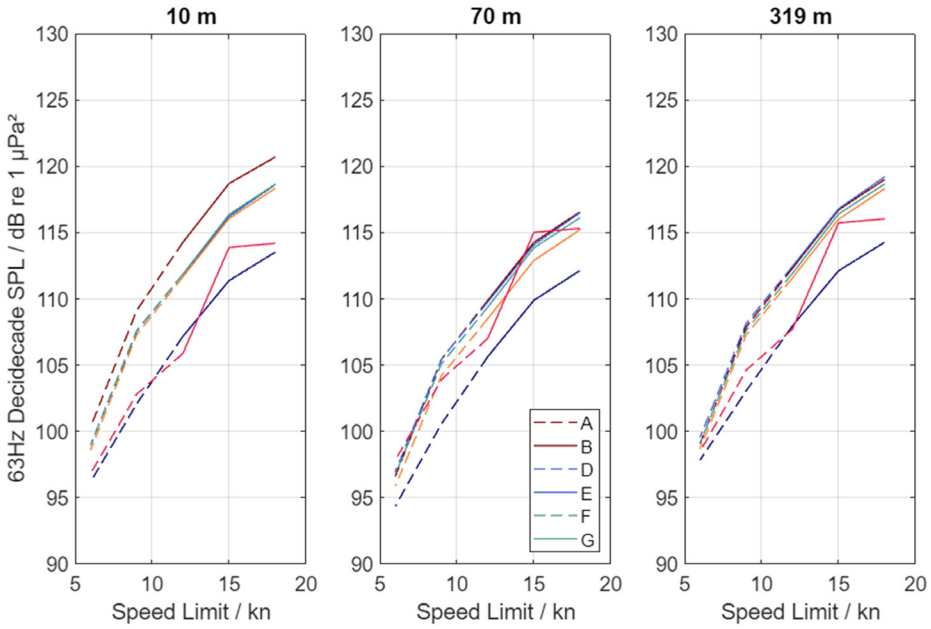

FIG. 10. (Color online) Sound pressure level vs speed limit predicted by the modeling teams for the three measurement depths at 63 Hz. Vessel speeds of 6, 9, 12, 15 and 18 kn are shown. The vessel source level model is valid when the vessels are cavitating, so speeds below 11 kn are shown with dashed lines. Group C did not provide results for this analysis.

models. Convergence testing is recommended to determine the thresholds for how to account for these effects as needed for different environments. For instance, Sertlek *et al.* (2016) found that the sound speed profile did not affect vessel sound maps in shallow environments, but it made a clear difference depending on the measurement depth in the present exercise (Fig. 10). The results of Heaney *et al.* (2024) provide an example of the rapid changes in received levels that occur due to changes in bathymetry and geoacoustic properties.

When acoustic or soundscape models are employed as inputs for regulatory decision making, an understanding of the uncertainty in the model estimates is essential, but often lacking. This exercise demonstrated that for low frequencies different modeling approaches yielded very similar results when the inputs to the models were all the same. The recommendations made here with respect to estimating decidecade band propagation losses and temporal and spatial sampling scales will help ensure that future modeling is performed to the same minimum standards. The knowledge gained about wind driven source levels will similarly improve the model accuracy and comparability. These results are an essential contribution to reducing acoustic modeling uncertainty.

Taking into account for uncertainty in source levels of vessels and other human sound sources as well as in the actual sound speed profiles and geoacoustic properties remain a significant challenge. Even a qualitative understanding of input sensitivities (i.e., relative importance of geoacoustics, sound speed profiles, etc.) can improve decision making when complete and quantitative assessments of uncertainty are computationally prohibitive. When quantitative analysis is possible, uncertainty can be addressed by comparisons to benchmarks of ground truth data. Alternately, Monte Carlo simulations that draw from distributions of environmental conditions and source levels may be used to characterize the range of possible sound levels that can be anticipated from proposed activities. Investigations that provide an estimate of uncertainty in effects ranges as a function of the variance in source levels are expected to be straight-forward to develop. Characterizing what sound speed profiles yield similar effects ranges, then developing methods of grouping different profiles to estimate the variability as a function of the possible types of profiles will be a greater challenge. Developing lookup tables for propagation loss as a function of depth, geoacoustic properties and sound speed profiles, which can then be compared as function of range may be a method of tackling this uncertainty.

It is recommended that the ambient sound modeling groups, as a next step, compare the estimated sound levels for the real environment in the Honguedo Strait to the measurements made at sites A and B (Fig. 2). For this investigation it is recommended that modelers employ the ERA5 wind speed data (Copernicus Marine Environment Monitoring Service, 2023) as well as vessel tracks from the AIS. Model verification will help understand the difference between models and the real world, as well as illuminating other sources of uncertainty such as variability of vessel source levels and the percentage of vessels that are present in the AIS data.

SUPPLEMENTARY MATERIAL

The supplementary material contains: (1) CSV file containing the sound speed profile used in the modeling, (2) a CSV file containing the wind speed history used in the modeling, (3) a zip-file containing the vessel positions and decidecade source levels for speed limits of 6, 9, 12, 15, and 18 kn, (4) a copy of the original workshop description provided to the participants, (5) a zip-file containing the participant results employed in this manuscript that future modelers may compare against. A copy of the MATLAB LIVEEDITOR script used to generate the figures in the manuscript is included in that zip-file.

ACKNOWLEDGMENTS

The Berlin workshop organizers (M.S., M.A., M.H., L.H., B.M., and M.P.) thank the Workshop Scientific Committee of Ross Chapman, K.H., and Christ de Jong for their insights and support during the development of the workshop. Financial support for organizing and conducting the workshop were provided by the U.S. Office of Naval Research (Ocean Acoustics), Rijkswaterstaat, and U.S. National Oceanographic and Atmospheric Administration. Thanks to Carmen Lawrence for preparing Fig. 1.

AUTHOR DECLARATIONS

Conflict of Interest

The Authors have no conflicts to disclose.

Author Contributions

S. Bruce Martin: Conceptualization (supporting), data curation, methodology, software, supervision, visualization, writing/original draft preparation. **Martin Siderius**: Conceptualization, funding acquisition, formal analysis, methodology, project administration, software, supervision, writing/review & editing. **Michael A. Ainslie**: Conceptualization, funding acquisition, methodology, supervision, writing/review & editing. **Michele B. Halvorsen**: Conceptualization (supporting), writing/review & editing. **Leila Hatch**: Conceptualization, writing/review & editing. **Mark K. Prior**: Conceptualization (supporting), formal analysis, methodology, software, supervision, visualization, writing/review & editing. **Daniel Brooker**: Formal analysis, software, writing/review & editing. **James Caplinger**: Writing/review & editing. **Christine Erbe:** Formal analysis, software, writing/review & editing. **John Gebbie**: Formal analysis, software, writing/review & editing. **Kevin D. Heaney**: Formal analysis, software, writing/review & editing. **Alexander O. MacGillivray**: Formal analysis, software, writing/review & editing. **Marie-Noel Matthews**: Formal analysis, software, writing/review & editing. **Victor**

O. Oppeneer: Formal analysis, software, visualization, writing/ original draft preparation. **Alexandra Schäfke**: Formal analysis, software, writing/review & editing. **Renée P. Schoeman**: Formal analysis, software, writing/review & editing. **H. Özkan Sertlek**: Formal analysis, software, writing/review & editing.

DATA AVAILABILITY

All data required to recreate the modeling in this manuscript are available in the supplementary material.

[1] 1 kn = 1852 m/(3600 s) ≃ 0.5144 m/s.

Ainslie, M. A. (**2010**). *Principles of Sonar Performance Modeling* (Springer, Berlin).

Ainslie, M. A., Halvorsen, M. B., and Dekeling, R. P. A. (**2016**). "International airgun modelling workshop: Validation of source signature and sound propagation models," in *4th International Conference on the Effects of Noise on Aquatic Life*, Dublin, Ireland.

Ainslie, M. A., Laws, R. M., and Sertlek, H. Ö. (**2019**). "International airgun modeling workshop: Validation of source signature and sound propagation models—Dublin (Ireland), July 16, 2016—Problem description," IEEE J. Oceanic Eng. **44**, 565–574.

Ainslie, M. A., Laws, R. M., Smith, M. J., Alexander, O., and MacGillivray, A. O. (**2024**). "Source and propagation modelling scenarios for environmental impact assessment: Model verification," J. Acoust. Soc. Am. **156**, 1489–1508.

Burgess, A. S., and Kewley, D. J. (**1983**). "Wind generated surface noise source levels in deep water east of Australia," J. Acoust. Soc. Am. **73**, 201–210.

Chapman, N. R., and Cornish, J. S. (**1993**). "Wind dependence of deep ocean ambient noise at low frequencies," J. Acoust. Soc. Am. **93**, 782–789.

Chapman, N. R., Ainslie, M. A., and Siderius, M. (**2024**). "Source level of wind-generated ambient sound in the ocean," JASA Express Lett. **4**, 010001.

Copernicus Marine Environment Monitoring Service (**2023**). "Global ocean physics reanalysis," https://doi.org/10.48670/moi-00021 (Last viewed May 9, 2024).

Dahl, P. H., de Jong, C. A. F., and Popper, A. N. (**2015**). "The underwater sound field from impact pile driving and its potential effects on marine life," Acoust. Today **11**, 18–25.

Duarte, C. M., Chapuis, L., Collin, S. P., Costa, D. P., Devassy, R. P., Eguiluz, V. M., Erbe, C., Gordon, T. A. C., Halpern, B. S., Harding, H. R., Havlik, M. N., Meekan, M., Merchant, N. D., Miksis-Olds, J. L., Parsons, M., Predragovic, M., Radford, A. N., Radford, C. A., Simpson, S. D., Slabbekoorn, H., Staaterman, E., Van Opzeeland, I. C., Winderen, J., Zhang, X., and Juanes, F. (**2021**). "The soundscape of the Anthropocene ocean," Science **371**, eaba4658.

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

Erbe, C., Williams, R., Sandilands, D., and Ashe, E. (**2014**). "Identifying modeled ship noise hotspots for marine mammals of Canada's Pacific Region," PLos One **9**, e89820-10.

European Commission (**2008**). "Directive 2008/56/EC of the European Parliament and of the Council of 17 June 2008 establishing a framework for community action in the field of marine environmental policy (Marine Strategy Framework Directive)," Official Journal of the European Union L 164/19.

Fisher, F. H., and Simmons, V. P. (**1977**). "Sound absorption in sea water," J. Acoust. Soc. Am. **62**, 558–564.

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

Gedamke, J., Harrison, J., Hatch, L. T., Angliss, R. P., Barlow, J. P., Berchok, C. L., Caldow, C., Castellote, M., Cholewiak, D. M., DeAngelis, M. L., Dziak, R. P., Garland, E. C., Guan, S., Hastings, S., Holt, M. M., Laws, B., Mellinger, D. K., Moore, S. E., Moore, T. J., Oleson, E. M., Pearson-Meyer, J., Piniak, W. E. D., Redfern, J. V., Rowles, T. K., Scholik-Schlomer, A. R., Smith, A., Soldevilla, M. S., Stadler, J. H., Van Parijs, S. M., and Wahle, C. (**2016**). *Ocean Noise Strategy Roadmap* (National Oceanic and Atmospheric Administration, Washington, DC).

Gisiner, R. C. (**2016**). "Sound and marine seismic surveys," Acoust. Today **12**, 10–18.

Harrison, C. H., and Harrison, J. A. (**1995**). "A simple relationship between frequency and range averages for broadband sonar," J. Acoust. Soc. Am. **97**, 1314–1317.

Heaney, K. D., Ainslie, M., Murray, J. J., Heaney, A. J., Miksis-Olds, J., and Martin, B. (**2024**). "Regional soundscape modeling of the Atlantic Outer Continental Shelf," J. Acoust. Soc. Am. **156**, 378–390.

Heinis, F., de Jong, C. A. F., and von Benda-Beckmann, A. M. (**2022**). "Framework for Assessing Ecological and Cumulative Effects 2021 (KEC 4.0)—Marine mammals," in *KEC Update Underwater Noise* (TNO), p. 85.

Horton, J. W. (**1959**). *Fundamentals of Sonar* (U.S. Naval Institute, Annapolis, MD).

International Maritime Organization (**2014**). "MEPC 66/17. Noise from commercial shipping and its adverse impacts on marine life. Outcome of DE 57," p. 7.

International Organization for Standardization (**2017**). "ISO 18405:2017. Underwater acoustics—Terminology" (International Organization for Standardization, Geneva), p. 51.

Joy, R., Tollit, D. J., Wood, J. D., MacGillivray, A. O., Li, Z., Trounce, K., and Robinson, O. (**2019**). "Potential benefits of vessel slowdowns on endangered southern resident killer whales," Front. Mar. Sci. **6**, 344.

Kinneging, N. A., and Tougaard, J. (**2021**). "Assessment North Sea," Report of the EU INTERREG Joint Monitoring Programme for Ambient Noise North Sea, Jomopans.

MacGillivray, A. O., and de Jong, C. A. F. (**2021**). "A reference spectrum model for estimating source levels of marine shipping based on automated identification system data," J. Mar. Sci. Eng. **9**, 369.

Mikkelsen, L., Johnson, M. P., Wisniewska, D. M., van Neer, A., Siebert, U., Madsen, P. T., and Teilmann, J. (**2019**). "Long-term sound and movement recording tags to study natural behavior and reaction to ship noise of seals," Ecol. Evol. **9**, 2588–2601.

Oppeneer, V. O., Schäfke, A., Brooker, D., and Prior, M. (**2024**). "Detection range of a passing emitter of underwater sound," J. Acoust. Soc. Am. **156**, 1646–1656.

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

Ross, D. (**1987**). *Mechanics of Underwater Noise* (Peninsula Publishing, Newport Beach, CA).

Sertlek, H. Ö., Binnerts, B., and Ainslie, M. A. (**2016**). "The effect of sound speed profile on shallow water shipping sound maps," J. Acoust. Soc. Am. **140**, EL84–EL88.

Siderius, M., Ainslie, M. A., Gebbie, J., Schäfke, A., Ross Chapman, N., Martin, B., and Gemba, K. L. (**2024**). "Techniques for modeling ocean soundscapes: Detailed description for wind contributions," J. Acoust. Soc. Am. **156**, 3446–3458.

Simpson, S. D., Radford, A. N., Holles, S., Ferarri, M. C. O., Chivers, D. P., McCormick, M. I., and Meekan, M. G. (**2016**). "Small-boat noise impacts natural settlement behavior of coral reef fish larvae," in *The Effects of Noise on Aquatic Life II*, edited by A. N. Popper and A. D. Hawkins (Springer, New York), pp. 1041–1048.

Thorp, W. H. (**1965**). "Deep-ocean sound attenuation in the sub- and low-kilocycle-per-second region," J. Acoust. Soc. Am. **38**, 648–654.

Wenz, G. M. (**1962**). "Acoustic ambient noise in the ocean: Spectra and sources," J. Acoust. Soc. Am. **34**(12), 1936–1956.

Wood, J. D., Southall, B. L., and Tollit, D. J. (**2012**). "PG&E offshore 3-D Seismic Survey Project Environmental Impact Report–Marine Mammal Technical Draft Report," Report by SMRU Ltd., p. 121.

Zykov, M. M., Martin, S. B. (**2024**). "Range versus frequency averaging of underwater propagation loss for soundscape modeling," J. Acoust. Soc. Am. **156**, 3439–3445.