Marine Pollution Bulletin 184 (2022) 114122

Contents lists available at ScienceDirect

Marine Pollution Bulletin

journal homepage: www.elsevier.com/locate/marpolbul

Acoustic detectability of whales amidst underwater noise off the west coast of South Africa

Fannie W. Shabangu [a,b,*], Dawit Yemane [a,c], George Best [d], Bobbi J. Estabrook [e]

[a] Fisheries Management Branch, Department of Forestry, Fisheries and the Environment, Foreshore, Cape Town, South Africa
[b] Mammal Research Institute Whale Unit, Department of Zoology and Entomology, University of Pretoria, Private Bag X20, Hatfield, Pretoria 0028, South Africa
[c] Department of Biological Sciences and Marine Research Institute, University of Cape Town, Cape Town, South Africa
[d] ORBCOMM, Apartado Correos 128 Port de Pollensa, Balearics 07470, Spain
[e] K. Lisa Yang Center for Conservation Bioacoustics, Cornell Lab of Ornithology, Cornell University, Ithaca, NY 14850, USA

ARTICLE INFO

Keywords:
Underwater noise
Long-term trends
Oceanographic conditions
Vessel traffic
Ocean economy
Noise pollution

ABSTRACT

Anthropogenic underwater noise has been shown to negatively affect marine organisms globally; yet little to no noise research has been conducted in most African waters including South Africa's. This study aimed to quantitatively describe sources of underwater noise and effects of underwater noise on the acoustic detectability of Antarctic blue, fin, minke, humpback, and sperm whales off South Africa's west coast. Noise from vessel traffic (<35 km to the location of recorders) dominated the soundscape below 500 Hz while wind-generated noise increased with wind speed above 5 m s^{-1} and dominated the soundscape above 500 Hz. Acoustic detectability of humpback, minke and sperm whales decreased with increasing ambient noise levels whereas blue and fin whale acoustic detectability increased with the ambient noise levels. We provide baseline information on underwater noise sources and the effects of underwater noise on whale acoustic detectability off the west coast of South Africa.

1. Introduction

Underwater sounds are generated by biological (e.g. crustaceans, marine mammals and fish), anthropogenic (e.g. sonar, vessel traffic, wind energy, and gas and oil exploration operations), and geophonic (e.g. rain, wind and wave actions, sea ice, earthquakes and volcanoes) sources (Wenz, 1962; Urick, 1983; McKenna et al., 2013; Melcón et al., 2012; Duarte et al., 2021). Marine organisms contend with these different sounds and are affected by different anthropogenic underwater noise to varying degrees, including loss of communication space, acoustic masking, and temporary and permanent shifts in their distribution (Southall et al., 2007; Clark et al., 2009; Melcón et al., 2012; Cholewiak et al., 2018). In extreme cases, anthropogenic noise can affect physiological and behavioural responses of marine organisms including hearing loss, disorientation and sometimes death (e.g. Southall et al., 2007; Cerchio et al., 2014; Duarte et al., 2021). Sound is important for marine organisms as they use it to locate prey (i.e. echolocation or stunning), navigation, communication, and predator detection and avoidance (Au and Hastings, 2008); thus, it is essential for these organisms to inhabit areas with suitable soundscapes given the effects of noise on their acoustic ecology and health (Erbe et al., 2016, 2019; Duarte et al., 2021). The west coast of South Africa is rich in marine mammal species biodiversity and acts as an important acoustic habitat (defined by Clark et al. (2011) as an "ecological space acoustically utilised by particular species") for numerous whale species including Antarctic blue whales (*Balaenoptera musculus intermedia*; vocalization frequency: 18–26 Hz), fin whales (*B. physalus*; vocalization frequency: 18–28 Hz), Antarctic minke whales (*B. bonaerensis*; vocalization frequency: 0.05–2 kHz), southern right whales (*Eubalaena australis*; vocalization frequency: 0.03–34 kHz), humpback whales (*Megaptera novaeangliae*; vocalization frequency: 0.02–24 kHz) and sperm whales (*Physeter macrocephalus*; vocalization frequency: 0.01–32 kHz) (Shabangu et al., 2019, 2020, 2021; Shabangu and Andrew, 2020; Letsheleha et al., 2022; Shabangu and Kowarski, 2022). Some of the aforementioned whale species use this region for overwintering, breeding, opportunistic feeding, migration, or year-round habitation.

Compared to other parts of the world, there has been little research evaluating the effects of noise on marine mammals in the southern African region, with only three published studies (Cerchio et al., 2014; Heiler et al., 2016; Koper et al., 2016) to date. A literature review of

* Corresponding author at: Fisheries Management Branch, Department of Forestry, Fisheries and the Environment, Private Bag X2, Vlaeberg 8018, South Africa
E-mail address: fannie.shabangu@yahoo.com (F.W. Shabangu).

https://doi.org/10.1016/j.marpolbul.2022.114122
Received 30 March 2022; Received in revised form 5 September 2022; Accepted 6 September 2022
Available online 22 September 2022

knowledge on the effects of anthropogenic noise on marine mammals showed a lack of research on oceanic noise in South African waters (Koper and Plön, 2012). Furthermore, Purdon (2018) suggests that the South African marine legislation might not be enough to protect marine life from anthropogenic noise associated with seismic surveys. Only one study (Schoeman et al., 2022) investigating underwater soundscape exist from shallow-water environment in South Africa. This disparity in noise research and knowledge is of great concern in the Anthropocene era given there is increasing pressure on the marine ecosystem due to anthropogenic stressors such as shipping and fishing (Purdon et al., 2020). For instance, anthropogenic stressors such as noise pollution associated with Operation Phakisa, a South African government initiative commissioned in 2014 with the aim to inter alia fast track the growth and development of the ocean economy within the South African Exclusive Economic Zone (EEZ) by 2030 (Zuma, 2014), are not quantified yet. This South African government ocean economy growth program works on unlocking the country's ocean economy by developing the following sectors: 1) marine transport and manufacturing; 2) offshore oil and gas exploration; 3) aquaculture; 4) marine protection services and ocean governance; 5) coastal and marine tourism; 6) small

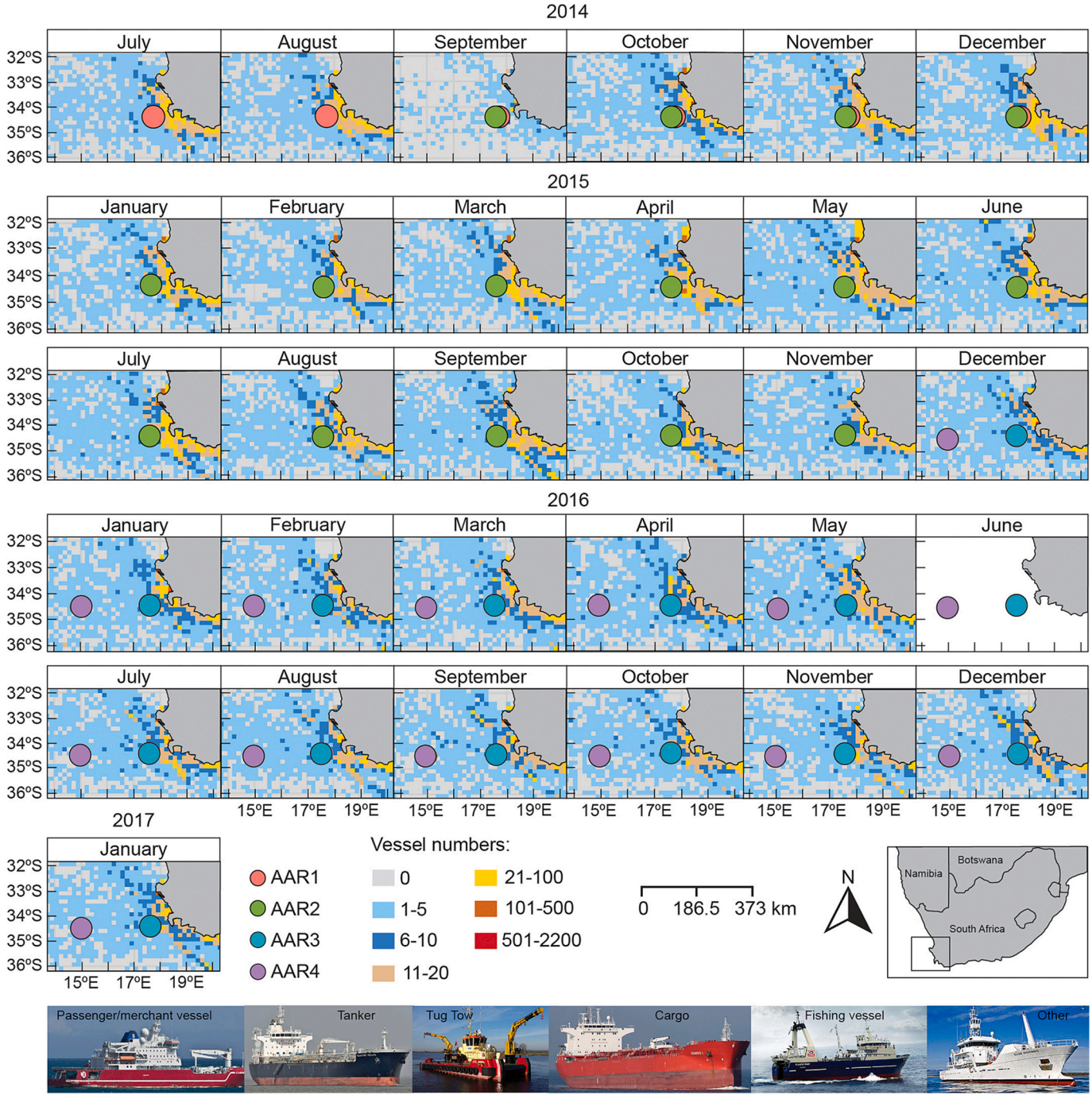

Fig. 1. Maps showing vessel density (number of vessels per 100 nm^2/185 km^2 grid) for each month around each autonomous acoustic recorder (AAR) mooring location, and examples of different vessel types tracked off the west coast of South Africa. AARs 2 and 3 were deployed on the same location, thus AAR2 location is overlaid by AAR3 circle in December 2015 map. Insert map shows location of the study area relative to the whole coast of South Africa. Vessel photo credit: Johannes Köring (passenger/merchant vessel), Bernhard Fuchs (tanker vessel), SSE (tug tow vessel), Hannes van Rijn (cargo vessel), Sea Harvest (fishing vessel), and Rob Tarr (other vessel).

harbours (Zuma, 2014; Operation Phakisa, 2015). Unarguably, most if not all of the above sectors will generate underwater noise directly or as a by-product of their operations, making it critically important now more than ever to identify sources of noise and monitor noise levels within the South African EEZ.

To date, there have been no long-term measurements of offshore noise level trends and effects of noise levels on whale acoustic detection off South Africa's coasts. This study aimed to close this knowledge gap by providing a quantitative overview of the influence of oceanographic conditions and vessel traffic on ambient noise levels, and the effects of noise on the acoustic detection of four baleen whale species (Antarctic blue, fin, Antarctic minke, and humpback whales) and one toothed whale species (sperm whales) off the west coast of South Africa.

2. Materials and methods

2.1. Acoustic data collection

Acoustic data were collected off the west coast of South Africa (Fig. 1) between July 2014 and January 2017 (Table 1) using autonomous acoustic recorders (AARs) deployed on oceanographic moorings as part of the South African Blue Whale Project (Shabangu et al., 2019). We used AARs of Autonomous Underwater Recorder for Acoustic Listening Model 2 version 04.1.3 manufactured by Multi-Électronique, Canada, with a usable frequency range of 10 Hz to 16 kHz. Sampling settings, hydrophone sensitivities, and details about the sampling sites are provided in Table 1. AARs 1–3 were located ~70 km from the shore, whereas AAR4 was ~310 km from the shore (Fig. 1). The water depth of 855 m at AAR1 mooring location was close to the relatively fast (20–70 cm s^{-1}) equatorward Benguela Current whereas the water depth (1118 m) at AARs 2 and 3 mooring location was least affected by that current (Shannon, 2009). On the other hand, the water depth of 4400 m at AAR4 mooring location was positioned on the path of the fast (2–10 Sv) Agulhas leakage rings associated with strong anticyclonic eddies (Shannon, 2009).

2.2. Acoustic detection of whale calls

All acoustic files were visually reviewed and inspected through spectrogram analysis by Shabangu et al. (2019, 2020), Shabangu and Andrew (2020), Letsheleha et al. (2022) and Shabangu and Kowarski (2022) to detect the presence and absence of whale calls (Fig. 2). For blue and fin whales, automated call detection analyses were implemented based on the template method using spectrogram correlation by Shabangu et al. (2019) and Letsheleha et al. (2022) and were manually validated by the lead authors of those studies. False negative rates of those automated detections were comparable between the two studies, making acoustic occurrence results comparable as well. All other whale detections were conducted by FWS. Based on the results of these studies, monthly percentages of acoustic detection of whale sounds were calculated as the number of files with whale acoustic presence divided by the total number of files recorded per month.

2.3. Oceanographic conditions

To estimate the contribution and influence of oceanographic conditions on the ambient noise around AAR locations, we considered the below environmental variables. A maximum radius of 80 km was applied across all AAR mooring locations to extract and average most oceanographic conditions except for ocean current speed measured locally via acoustic Doppler current profilers (ADCPs) deployed on these moorings. The 80 km buffer distance was estimated using an acoustic propagation modelling (BELLHOP model) as the maximum detection range for sperm whale clicks (Shabangu and Andrew, 2020) and likely represented the possible maximum range for most AARs. All oceanographic data were processed using 'raster' package (Hijmans, 2020) in R (version 4.1.1; R Core Team, 2021).

2.3.1. Ocean current speed

Hourly ocean current speeds (cm s^{-1}) used as a proxy of frontal activities were collected by upward looking 75 kHz ADCPs deployed at various depths on oceanographic moorings of the South Atlantic Meridional Overturning Circulation Basin-scale Array (SAMBA) transect. No ADCP was deployed on AAR1 mooring since this was an experimental SAMBA mooring deployment. Ocean current speed data were collected from September 2014 to November 2015 for AAR2 (Lamont and van den Berg, 2021a); December 2015 to April 2017 for AARs 3 and 4 (Lamont and van den Berg, 2021b, 2021c). ADCP current speeds were measured in the upper water column from 69.88 to 581.88 m, 64.56 to 576.57 m, and 60.84 to 428.84 m for AARs 2, 3 and 4 respectively. We only considered hourly values with ≥75% good beam data and averaged current speed data around AARs' depth in the water column using measurements from the two nearby ADCP beams.

2.3.2. Wind speed and total precipitation

Hourly meridional (v10) and zonal (u10) wind speeds (m s^{-1}) and total precipitation (m) measured through the fifth generation of European Centre for Medium-Range Weather Forecasts (ECMWF) reanalyses (ERA5; Copernicus Climate Change Service, 2017) were downloaded from https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=form. ERA5 wind speed data were collected using a spatial coverage of 0.25° × 0.25° global grids at a height of 10 m above the surface of the Earth (Copernicus Climate Change Service, 2017). The absolute wind speed and wind direction were calculated from u10 and v10 vectors using equations in the Supplementary Material. Total precipitation represents the accumulated liquid and frozen water from rain and snow that falls to the ocean's surface and is the sum of large-scale precipitation and convective precipitation sampled at a spatial coverage of 0.25° × 0.25° (Copernicus Climate Change Service, 2017).

2.3.3. Ocean waves

We obtained ERA5 hourly data on the significant height of combined wind waves and swell (m; hereafter wave height) and peak wave period (s; hereafter wave period) with a spatial resolution of 0.5° × 0.5° from

Table 1
Settings and study site details of the four autonomous acoustic recorders (AARs) used in this study. AARs are numbered according to order of their chronological deployment. Hydrophone sensitivities are from factory calibrations of the HTI-96-MIN (High Tech Inc.) hydrophones.

Hydrophone	Latitude (S)	Longitude (E)	Water depth (m)	AAR depth (m)	Sampling rate (Hz)	Sampling schedule (min h^{-1})	Hydrophone sensitivity (dB re 1 V/μPa)	Start recording date	Stop recording date
AAR1	34° 22.21′	17° 37.69′	855	200	4096	30	−164.20	24/07/2014	01/12/2014
AAR2	34° 23.64′	17° 35.66′	1118	300	4096	20	−163.90	16/09/2014	01/12/2015
AAR3	34° 23.64′	17° 35.66′	1118	300	8192	25	−164.10	04/12/2015	01/01/2017
AAR4	34° 30.36′	14° 58.81′	4481	200	8192	25	−164.20	04/12/2015	13/01/2017

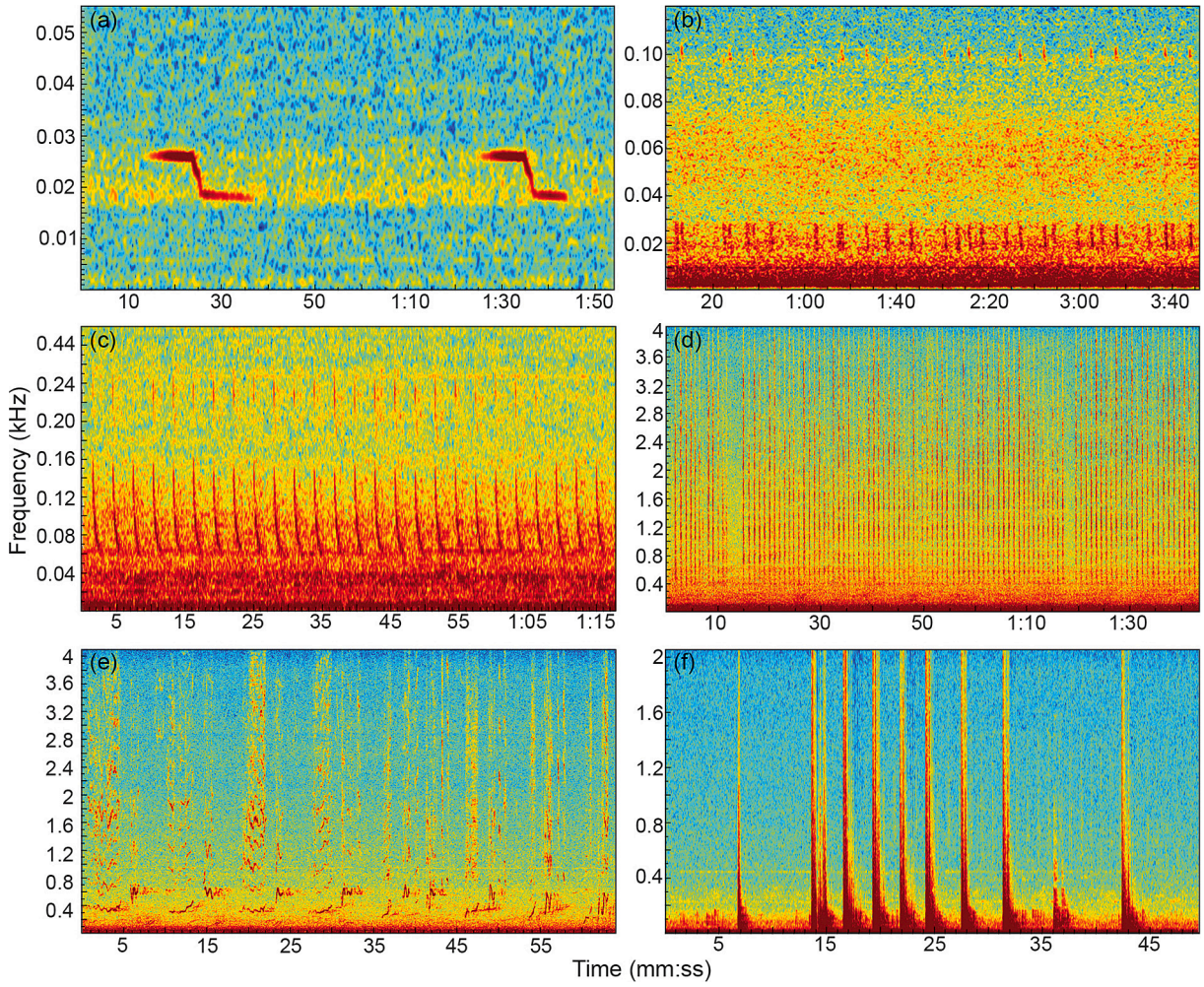

Fig. 2. Examples of whale sounds detected off the west coast of South Africa and used here to define acoustic detection: (a) Antarctic blue whale *Z*-calls; (b) fin whale 20 Hz pulses showing the higher frequency pulse with a peak at 99 Hz that is associated with the eastern Antarctic fin whale acoustic population; (c) Antarctic minke bioduck C train that can consist of up to 98 pulses; (d) sperm whale click train; (e) humpback whale song units; (f) southern right whale gunshot series. X- and y-axes scales are different between plots. Spectrogram parameters (all figures were produced using Hann window and 90% overlap): (a) frame size = 1.51 s and discrete Fourier transform (DFT) size = 8192 samples; (b) frame size = 2.03 s and DFT size = 16,384 samples; (c) frame size = 0.399 s and DFT size = 4096 samples; (d) frame size = 0.170 s and DFT size = 2048 samples; (e) frame size = 0.210 s, DFT size = 2048 samples; (f) frame size = 0.125 s, DFT size = 512 samples.

https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=form. Wave height corresponds to the total wave field that can be observed at sea, which combines the shorter wind waves that are being driven by the wind and the longer swell waves that have been generated by distant winds and are moving across ocean basins. Wave period is the time when most of the wave energy is concentrated and is used here to determine the length of waves which influence the breaking characteristics and interaction with the seabed and potential noise caused by this process.

2.4. Noise data analyses

Noise analysis was conducted using a custom MATLAB program (Dugan et al., 2011) to estimate noise levels from each AAR. For AARs 1 and 2, noise was categorized into three frequency bands (10–500, 500–1000, and 1000–2000 Hz) given similar Nyquist frequency of recordings. For AARs 3 and 4, noise was categorized into five frequency bands (10–500, 500–1000, 1000–2000, 2000–3000, and 3000–4000 Hz) given similar Nyquist frequency of recordings. To calculate noise levels, sound data were processed using a Hann window with zero overlap, and a fast Fourier transform with 1 s and 1 Hz resolution. We used equivalent continuous sound pressure level (L_{eq}), in dB re 1 μPa, to represent the average unweighted sound pressure of a continuous time-varying signal, with an averaging time of 10 min. Visual representation of the noise statistics is presented using long-term spectrograms, 1/3-octave spectrograms, and spectral probability density (SPD) plots.

Various sound sources in the marine environment can elevate sound levels within different frequency bands (Urick, 1986; Hildebrand, 2009). For example, most large vessel noise occurs at frequencies below 500 Hz (Wenz, 1962), while frequencies from 500 Hz to 25 kHz can be dominated by sea-surface agitation (Knudsen et al., 1948). To investigate the potential noise contribution of different sound sources in the soundscape of South Africa's west coast, we measured noise levels within the above three or five distinct frequency bands (depending on the Nyquist frequency of recordings) to characterize noise at each recording site. We categorized measured noise into two groups based on ocean current speed (above and below 11 cm s^{-1} where the effect of current speed was flat (or non-existent) as shown in Fig. S1): 1) ambient noise which only includes noise in the soundscape within periods when current speeds was $\leq$11 cm s^{-1}, and 2) recorded noise which includes the pseudo-noise generated on the hydrophone surface by the flow of ocean current (Strasberg, 1979) at >11 cm s^{-1}, which is not part of the soundscape. Ocean current speed was used for filtering data as it was the main predictor of low frequency recorded noise levels (Fig. S1). The

remaining usable data after filtering represented 43% and 37% of the original dataset from AARs 2 and 3 respectively. Measured noise data from AAR1 were excluded from the ambient noise analysis since there were no current speed data to filter out the high pseudo-noise induced by the ocean current, and data from AAR4 were eliminated from ambient noise analysis as there was too much contamination by pseudo-noise (Figs. 3, S2 and S3).

2.5. Vessel traffic

Vessel traffic data acquired using the critical near real-time vessel monitoring, ocean buoy tracking and automatic identification system (AIS) ship tracking system were obtained from ORBCOMM (https://www.orbcomm.com/en/industries/maritime). We determined the number and types of vessels around all AARs within the maximum detection radius of 80 km to correspond to the buffer distance of environmental conditions. Vessel traffic data from the first half of every hour were processed to match the sampling protocol used for each AAR; for example, vessel traffic data from the first 30 min of every hour were extracted for AAR1. Individual vessels were identified using their Maritime Mobile Service Identity number. Distance of the closest vessel to the AAR location within each recording session was used to define the closest point of approach of a vessel(s). Speed over ground of the closest vessel was utilized as an indication of vessel speed near the AAR location, and speeds above 20 knots were eliminated as erroneous values. The location of AARs 2 and 3 was used for AAR1 since these were closely spaced (~5 km apart), and AAR4 was treated as independent sampling location given its isolated location (Fig. 1). To create the vessel density map, grids with resolutions of 10 × 10 nautical miles (nm) were created. Vessel track data were joined to these grids and then the number of vessels within each grid cell was calculated for each month and year and shown with density binned into seven classes (Fig. 1). The monthly number of vessels was calculated at different distance intervals to AAR locations: 0–1, 2–5, 6–20, 21–50, 51–100, and 101–910 km.

2.6. Application of statistical and machine learning models

We modelled the influence of oceanographic conditions and vessel traffic on the ambient L_{eq}, and the influence of ambient L_{eq} on the acoustic detectability of four baleen whale species (Antarctic blue, fin, Antarctic minke, and humpback whales) and one toothed whale (sperm whales) off the west coast of South Africa. The influence of oceanographic conditions on detected seasonal acoustic occurrence of various whale species has already been widely studied (Shabangu et al., 2019; Shabangu and Andrew, 2020; Letsheleha et al., 2022), and the potential influence of ambient L_{eq} on the acoustic detectability of whale sounds is assessed in this study.

Random forest (RF) model (Breiman, 2001) was the chosen modelling framework for modelling the influence of the environmental variables and vessel traffic on the ambient L_{eq} at different frequency bands. As a machine learning technique, the RF model was chosen for this study due to its non-parametric inferences; it inherently includes interactions among predictors; has reasonably good predictive performance; returns estimated variable importance as part of the model fit and handles large and small sample sizes with high dimensional data (Breiman, 2001; Schonlau and Zou, 2020). Additionally, RF when growing trees, at each split, only evaluates a random subset of predictors to identify the best predictor. Overall, this results in decorrelated trees making overfitting less likely (Breiman, 2001; Strobl et al., 2009).

Generalized linear model (GLM) regression (Dobson, 1990) was used to model the influence of ambient L_{eq} on the probability of detecting whales acoustically. GLM is a statistical model that allows the response variable to accept an error distribution other than a normal distribution and expresses the relationship as a linear function/combination of all

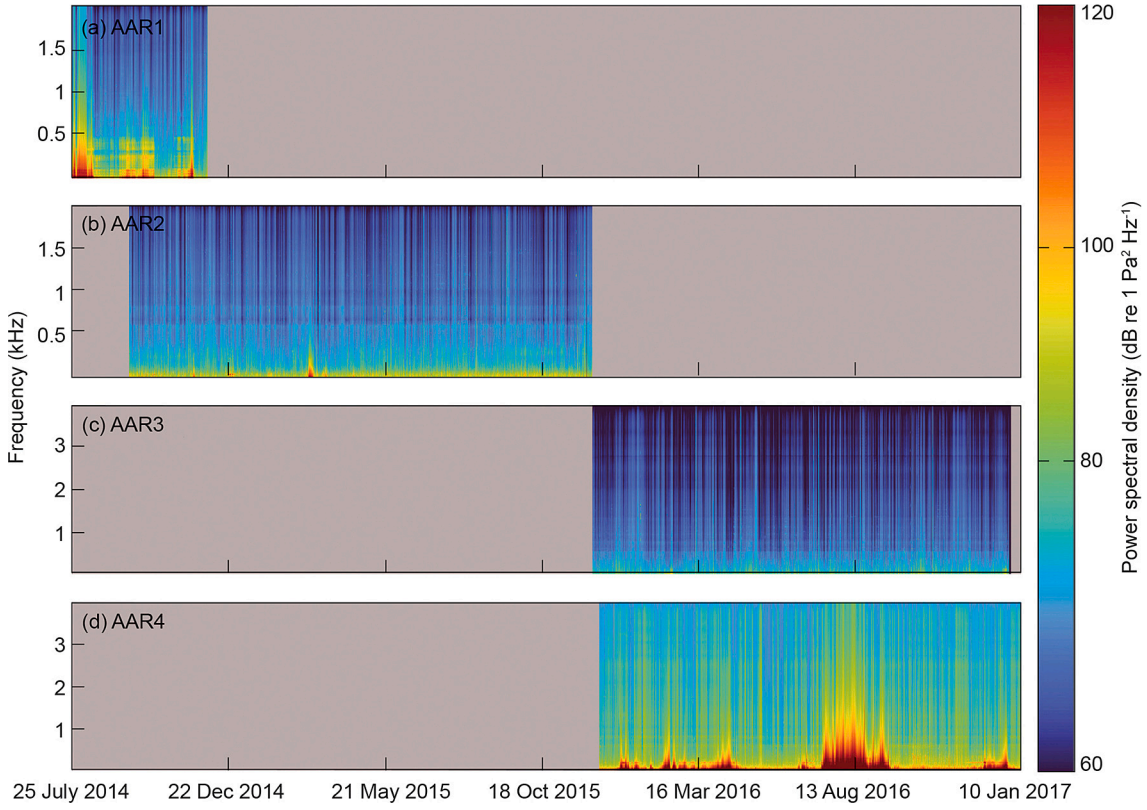

Fig. 3. Long-term spectrograms of recorded noise levels including pseudo-noise expressed as power spectral density (dB re 1 μPa^2 Hz^{-1}), averaged over 1 h time intervals for each recording site: (a) AAR1, (b) AAR2, (c) AAR3 and (d) AAR4 across all sampled frequencies. Grey shaded areas indicate time periods without acoustic recordings. Frequency scales (y-axes) for panels (a) and (b) differ from panels (c) and (d).

the predictors. This model assumes the underlying relationship between the response and the predictors to be linear although one can explicitly incorporate non-linearity in multiple ways (e.g., as polynomial function, B-spline, etc.). GLM allowed us to estimate the relationship between the probability of whale call detection and ambient noise level at the lowest and highest frequency band for each AAR. Percentages of whale acoustic detection that would be affected by masking was estimated based on the relationship between detection probability and noise level. We quantified the probability of whale detections at different noise levels to a point of 0.5 at which a species probability of detection was above or below average (i.e. 50%).

2.6.1. Tuning and testing models

Prior to modelling the influence of environmental variables on ambient L_{eq} at different frequencies, pair-wise correlations among environmental variables were explored for AARs 2 and 3 but no strong correlations were found (Figs. S4 and S5). Principal component analysis (Jolliffe and Cadima, 2016) was applied to L_{eq} at different frequencies since there were strong correlations found among L_{eq} at different frequency bands, especially above the 500–1000 Hz frequency band (Figs. S6 and S7). Principal components (PCs) were plotted as vector plots to summarise the association of principal component and loading of variables (the different frequency bands). To overcome the observed correlation issue and for ease of interpretation, we used the lowest and highest frequency bands for each AAR when modelling the acoustic detection of whales (see sub-Section 3.2). Four different methods of addressing noticeable differences in class imbalance of whale acoustic detection were used: Synthetic Minority Over-sampling TEchnique (SMOTE; Chawla et al., 2002), ADAptive SYNthetic (ADASYN; He et al., 2008), downsampling and upsampling (Nallamuthu, 2020). All models were tuned using 70% of the balanced data for training and the remaining 30% was used for testing. Training data set was further setup for 5-fold cross-validation. In the context of the RF used to model ambient L_{eq} at different frequencies, this was used to tune the model and select the optimal tuning parameter. For the GLM, this helped to compare the performance of the model under each of the class balancing algorithm. Predictive performances of all models were then assessed on the test data set which was not used when tuning models. For the regression RF model that assessed the influence of environmental conditions and vessel traffic on L_{eq} at different frequency bands, root mean squared error, mean absolute error, and coefficient of determination (r-squared; which is the proportion of the variation in the dependent variable that is explained by the independent variable(s) in a regression model) were used as measures of model performance. The best value of tuning parameter was selected as sets that maximize r-squared. For the GLM, influence of ambient L_{eq} on the acoustic detection of whales, there were no tuning parameters, and the cross validation was used to compute model performance when using the different class balancing algorithms. In this context, the area under the curve (AUC) receiver operating characteristic was used as a measure of model performance.

2.6.2. Feature importance

For the RF model used to model the influence of environmental variables and marine traffic on ambient noise, relative importance was assessed via improvement response variance. Vessel type was removed from the final RF models for L_{eq} as it had an insignificant relative importance. Whereas for the GLM model used to model the influence (i.e. dose-response curves) of the low and high frequency band underwater noise on the probability of detecting whales acoustically, the relative importance of predictors was assessed via the firm approach (Greenwell et al., 2018; Scholbeck et al., 2020), where relative flatness of the partial effect of a variable is used to as a measure of its importance. Importance of all features were scaled to the maximum for each model and presented as percent. The GLM and RF model formulas are presented in Table S1. All the data processing, visualization, analysis, and summary of results were performed in R using multiple packages (Table S2).

3. Results

3.1. Noise overview

The average L_{eq} at 10–500 Hz frequency band for AARs 2 and 3 dropped from 110 dB re 1 μPa (range: 97.0–138.9 dB re 1 μPa) to 107 dB re 1 μPa (range: 77.1–109.9 dB re 1 μPa) after filtering out pseudo-noise induced by the ocean current speed >11 cm s^{-1} (Figs. 4, S1, S3, S8 and S9). The power spectra decreased gradually above 100 Hz for AARs 2 and 3 (Fig. 4). Distinct power spectra peaks were evident between 18 and 26 Hz at 25^{th} and 50^{th} percentiles and at 65 Hz for the AAR2 unfiltered data representing recorded noise, whereas peaks were present between 18 and 26 Hz at the 25^{th}–75^{th} percentiles and 22 and 32 Hz at the 95^{th} percentile for AA2 filtered data representing ambient noise (Fig. 4). Power spectra peaks were present between 30 and 40 Hz for AAR3 unfiltered data, between 18 and 26 Hz at the 50^{th}–75^{th} percentiles, and a 40 Hz hump at the 95^{th} percentile for the AAR3 filtered data (Fig. 4). Specific spectra peaks between 18 and 26 Hz correspond to blue and fin whale vocalization frequency (Fig. 2). Overall, the filtered data had lower power spectra and ambient noise levels than unfiltered data for AARs 2 and 3 (Figs. 4, S8 and S9). No seismic survey signals were detected from any of the recording sites (Figs. 3 and S2).

3.2. Ambient noise over frequency bands

The most common feature of the vector plots was that PC2 was commonly associated to ambient L_{eq} at the low frequency band (10–500 Hz) indicating that PC2 represents the low frequency noise whereas PC1 represents the joint medium to high frequency band noise (Fig. 5). When looking specifically at AAR2, PC1 was negatively correlated to L_{eq} at 500–1000 and 1000–2000 Hz frequency bands and PC2 was positively correlated to L_{eq} at 10–500 Hz frequency band (Fig. 5a). For AAR3, PC1 was strongly positively correlated to L_{eq} at 1000–2000, 2000–3000 and 3000–4000 Hz frequency bands whereas PC2 was negatively correlated to L_{eq} at the low frequency bands (10–500 and 500–1000 Hz; Fig. 5b).

3.3. Model performance

Predictive performance of the RF models evaluating the influence of environmental variables and vessel traffic on ambient L_{eq} showed comparable predictive performance for AARs 2 and 3 (Fig. S10). The performance of the RF model on the test set shows a better performance for predicted values (Fig. S11). The AUC results indicate that all sample balancing methods can be used for modelling the five whale species acoustic detection from the two stations (Fig. S12). Additionally, calibration plots indicate that all sample balancing methods can be used for acoustic detectability modelling as they have reasonable fits of predicted data to observed data (Figs. S13 and S14).

3.4. Predictors of ambient noise

The RF models indicated that the 0–35 km distance of the closest vessel to the location of AARs had the highest effect on the ambient noise for all frequency bands, and the partial effect decreased as the distance to the AAR location increased. For example, the AAR3 L_{eq} decreased from 113 dB re 1 μPa at 0 km to 105 dB re 1 μPa at 35 km for the 10–500 Hz frequency band (Fig. 6). Thus, vessel noise contributed 8 dB to the AAR3 ambient noise level for the 10–500 Hz frequency band; similar interpretation is applicable to other frequency bands and other variables. The ocean current speed influence on ambient L_{eq} was constant for all frequency bands although there was a slight increase in influence from 8 to 11 cm s^{-1} for the 10–500 Hz frequency band (Fig. 6). The partial effect of hour of day on L_{eq} was constant throughout different hours for different frequency bands indicating no diel pattern in noise variation. Autumn through mid-spring (April–October) months had the highest partial effect on L_{eq} of 10–500 Hz frequency band from both

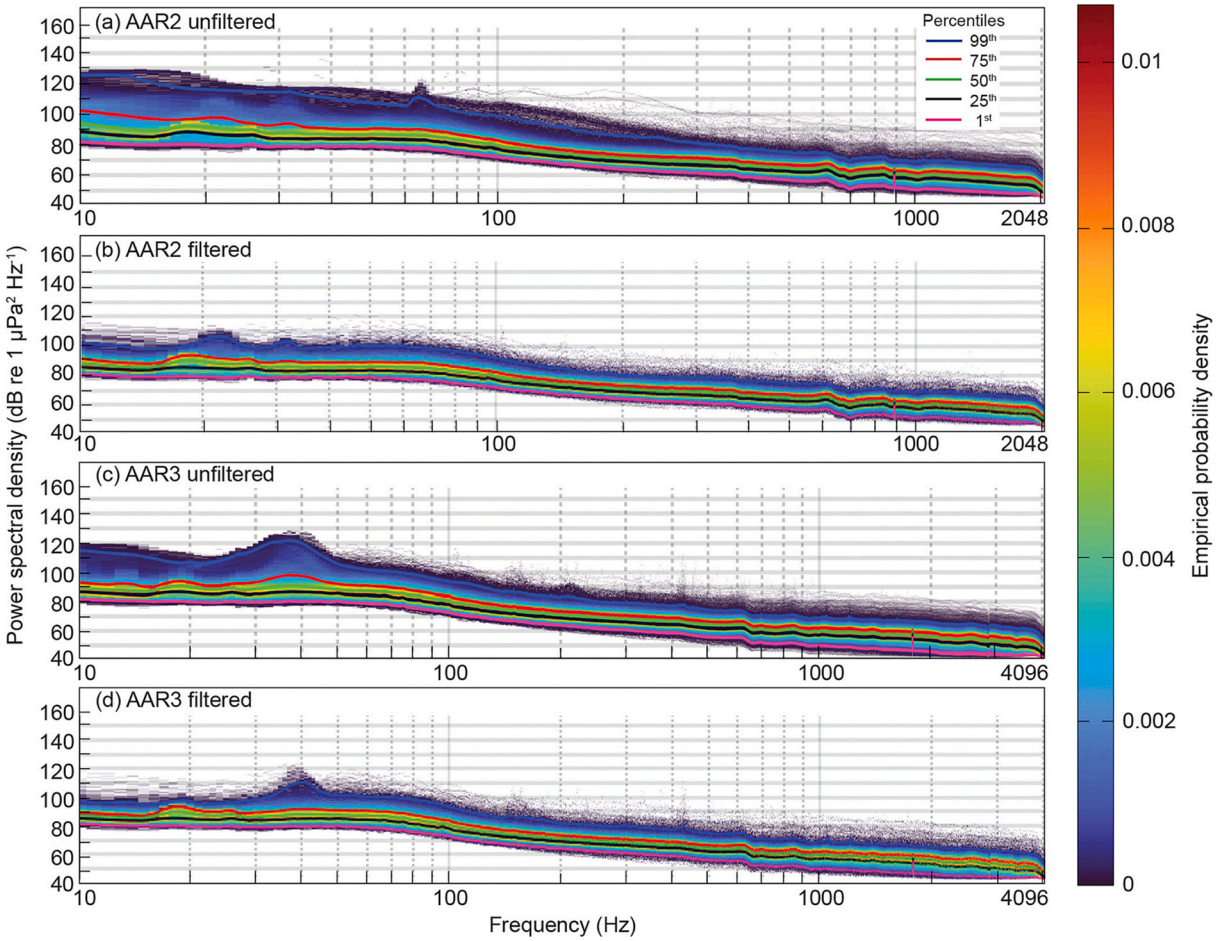

Fig. 4. Spectral probability density plots illustrating the statistical distribution of power spectral density percentiles at a frequency resolution of 1 Hz for (a) AAR2 unfiltered data including pseudo-noise, (b) AAR2 data filtered at $\leq$11 cm s^{-1} current speed to represent the ambient noise, (c) AAR3 unfiltered data including pseudo-noise, and (d) AAR3 data filtered at $\leq$11 cm s^{-1} current speed to represent the ambient noise. Frequency scales (x-axes) for panels (a) and (b) differ from panels (c) and (d).

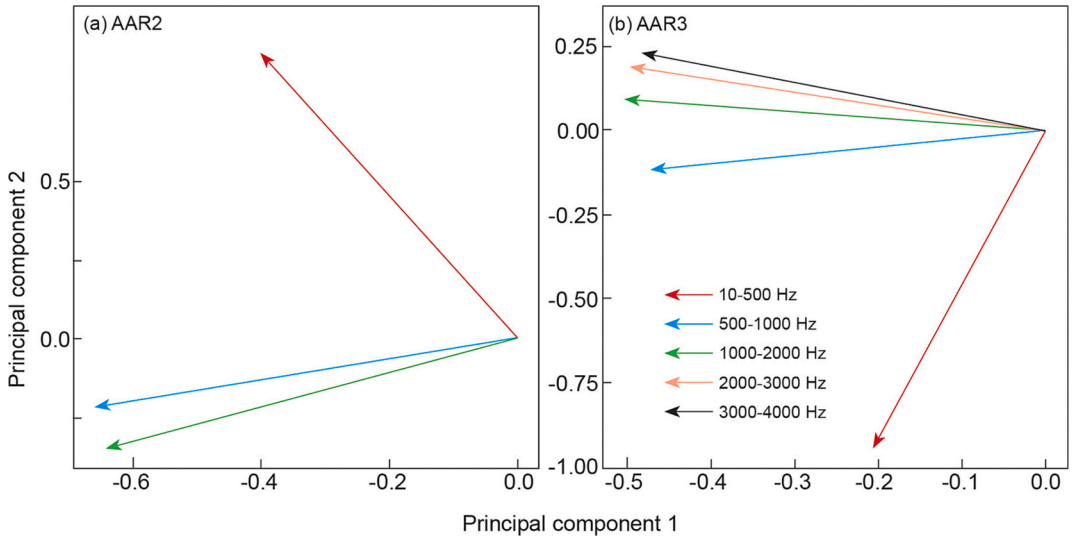

Fig. 5. Principal component analyses vector plots of ambient equivalent continuous sound pressure levels at different frequency bands for autonomous acoustic recorders (AARs) 2 and 3 data filtered at $\leq$11 cm s^{-1} current speed. Each vector points in the direction of the steepest increase of values.

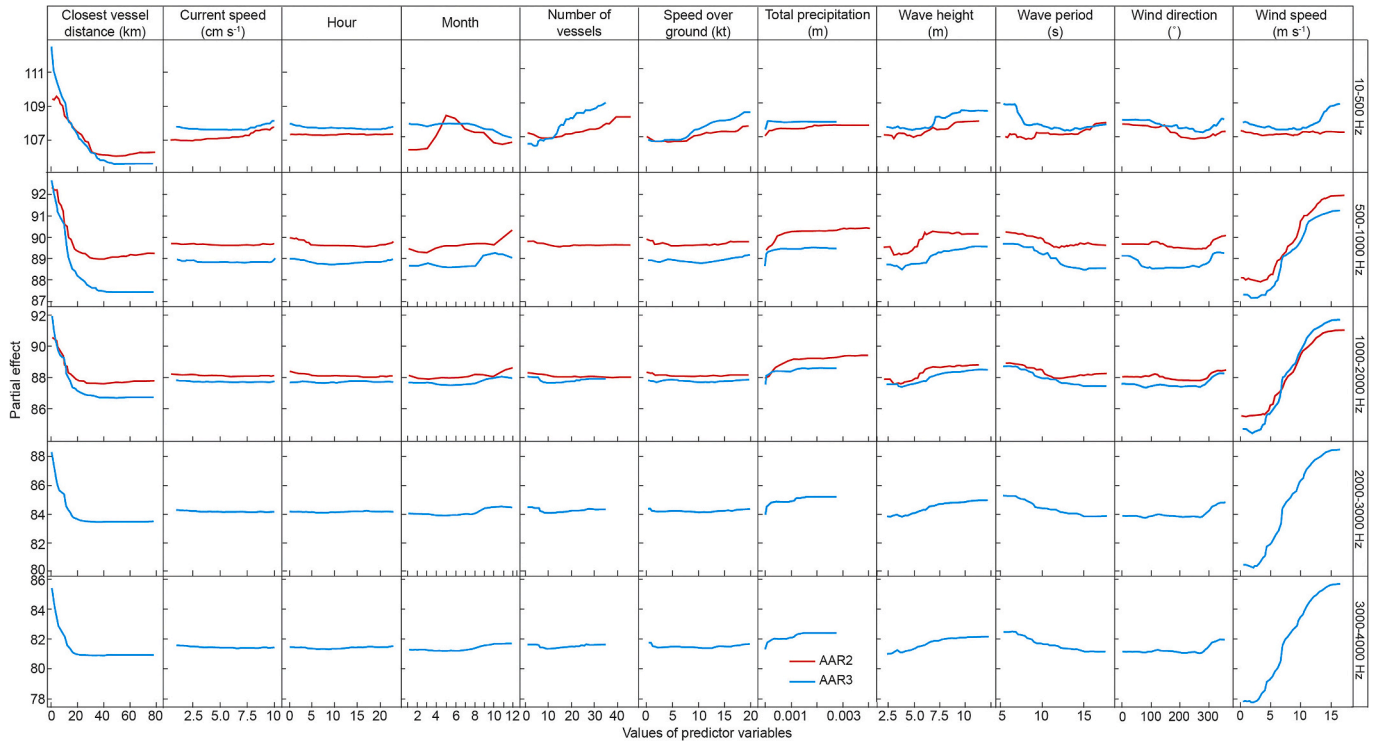

Fig. 6. Random forest (RF) model partial effect of predictor variables (x-axis) on the ambient equivalent continuous sound pressure level (y-axis) at different frequency bands from the three autonomous acoustic recorders (AARs) using data filtered at ≤11 cm s^{-1} current speed. No variance around predicted lines is shown as RF model results have little variation. Kt is for knots.

AARs (Fig. 6). November and December had the highest effect on L_{eq} of other frequency bands from AAR2, and September through December had the highest effect on L_{eq} of other frequency bands from AAR3 (Fig. 6).

Partial effects of vessel traffic on the 10–500 Hz L_{eq} increased with the number of vessels up to 40 vessels within a recording session (20 min per hour) for AAR2 and then plateaued whereas AAR3 partial effects of the vessel traffic on L_{eq} increased with the number of vessels up to 35 vessels within a recording session (25 min per hour). The effect of the number of vessels was flat for the other frequency bands (Fig. 6). The vessel speed over ground influence increased from 8.5 to 20 knots for the 10–500 Hz frequency band, and the influence was flat for the other frequency bands. The total precipitation influence increased sharply from zero to 0.001 m and plateaued throughout different frequency

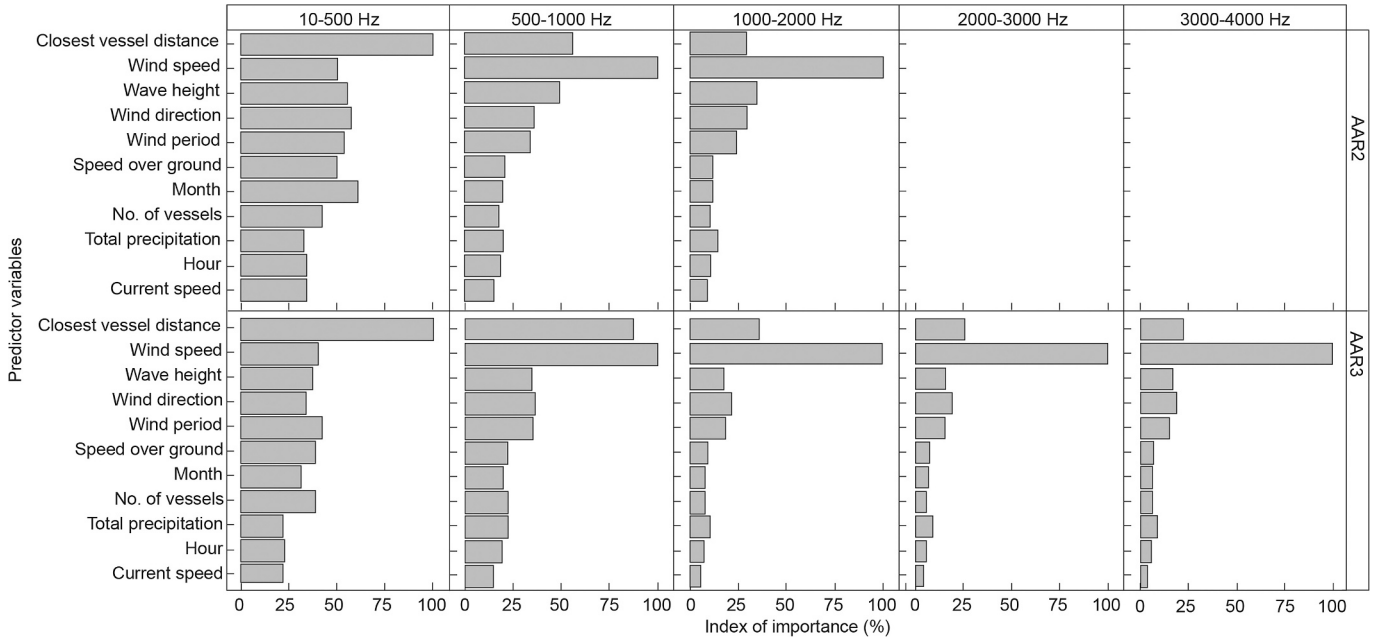

Fig. 7. Random forest (RF) model relative importance of predictor variables on influencing the ambient equivalent continuous sound pressure level at the different frequency bands from the two autonomous acoustic recorders (AARs) using data filtered at ≤11 cm s^{-1} current speed. Blank boxes represent cases when no data were available for AAR2 due to the Nyquist frequency limit of recordings.

bands (Fig. 6). The wave height influence on L_{eq} increased with its increase for all frequency bands (Fig. 6). The partial effects of wave period on L_{eq} decreased with the increase in wave period over different frequency bands (Fig. 6). Wind direction between 0 and 100° and around 360° had the highest influence on L_{eq} for 10–500 Hz frequency band, while direction around 360° had the highest influence for other frequency bands. Wind speed influence on the 10–500 Hz frequency band noise level was flat for both AARs but increased above 12.5 m s^{-1} for 10–500 Hz frequency band from AAR3 (Fig. 6). Partial effects of wind speed on L_{eq} from other frequency bands increased from 5 to 15 m s^{-1} (Fig. 6).

Based on RF models, the distance of the closest vessel to the AAR location was the most important predictor of 10–500 Hz frequency band L_{eq} from both AARs and was the second most important predictor of the 500–1000 Hz frequency band from AAR3 (Fig. 7). Wind speed was the most important predictor of L_{eq} for 500–1000, 1000–2000, 2000–3000, and 3000–4000 Hz frequency bands (Fig. 7). Other remaining variables were the moderate to least important predictors of L_{eq}. Notably, ocean current speed was the least important predictor of ambient noise (Fig. 7). The number of vessels per distance interval indicated that most vessel traffic transited >20 km from AARs 1–3 locations but >100 km to AAR4 location (Fig. 8). Vessel numbers varied from month to month but did not increase over time (2014–2017). Vessel density indicates that more ships were transiting through the area of AARs 1–3 whilst some were entering or exiting the Cape Town Harbour (Fig. 1). Different vessel types were tracked transiting through the west coast of South Africa including tug/tow, tanker, pleasure craft/sailing, passenger, military, fishing, cargo and other (Fig. 1).

3.5. Effects of ambient noise on whale acoustic detectability

The overall monthly percentage of acoustic detection of all whale species relative to recorded L_{eq} was higher in winter and spring than in the other seasons (Fig. S15). The probability of detecting blue and fin whale calls from both AARs 2 and 3 increased with the ambient noise level at the 10–500 Hz frequency band based on different sample balancing methods (Fig. 9). On the other hand, blue and fin whale probability of detection decreased with increasing ambient noise levels at highest frequency bands (1000–2000 and 3000–4000 Hz) from AARs 2 and 3. When evaluating the influence of recorded noise, including pseudo-noise on whale detectability, GLM dose-response curves of all whales including southern right whales showed that whales' detectability decreased with increasing noise levels at the low frequency band (10–500 Hz) except for blue whales from AAR1 that increased with noise level (Figs. S21–S25).

The probability of detecting humpback, minke, and sperm whales decreased with increasing ambient noise levels at both low and high frequency bands apart from humpback whales at the highest frequency band from AAR2 (Fig. 9). The probability of detecting humpback, minke, and sperm whales was low above 110 dB re 1 μPa L_{eq} for the 10–500 Hz frequency band from AARs 2 and 3 based on different sample balancing methods (Fig. 9). Blue and fin whale detectability was high above 110 dB re 1 μPa L_{eq} for the 10–500 Hz frequency band. Sample balancing was critical for improving model performance as shown by better predictive performances of GLMs with balanced sample sizes, whereas unbalanced data showed poor predictability performances compared to sample balanced methods for most whale species (Fig. 9). About 80% of acoustic detectability of humpback, minke, and sperm whales was below the 50% point for AARs 2 and 3 at the low frequency band. The 50% point was passed at low ambient noise levels around 90 dB re 1 μPa at the highest frequency bands for all whales from AARs 2 and 3 based on different sample balancing methods (Fig. 9).

4. Discussion

In general, our ambient noise level measurements show that temporal trends of the underwater noise levels off the west coast of South Africa were strongly predicted by vessel traffic for the frequency band below 500 Hz and by wind speed for frequency bands above 500 Hz. Results of this study highlight this region as an area with spectrally, temporally, and spatially dynamic ambient noise levels. The probability of acoustic detection of whales suggests species-specific responses to noise levels at different frequency bands. This study provides the baseline information of the underwater noise sources and acoustic detectability of whales for the west coast of South Africa.

4.1. Noise trends and predictors

Density plots and long-term spectrograms showed higher recorded noise levels for AARs 1 and 4 than for AARs 2 and 3, which reflects the

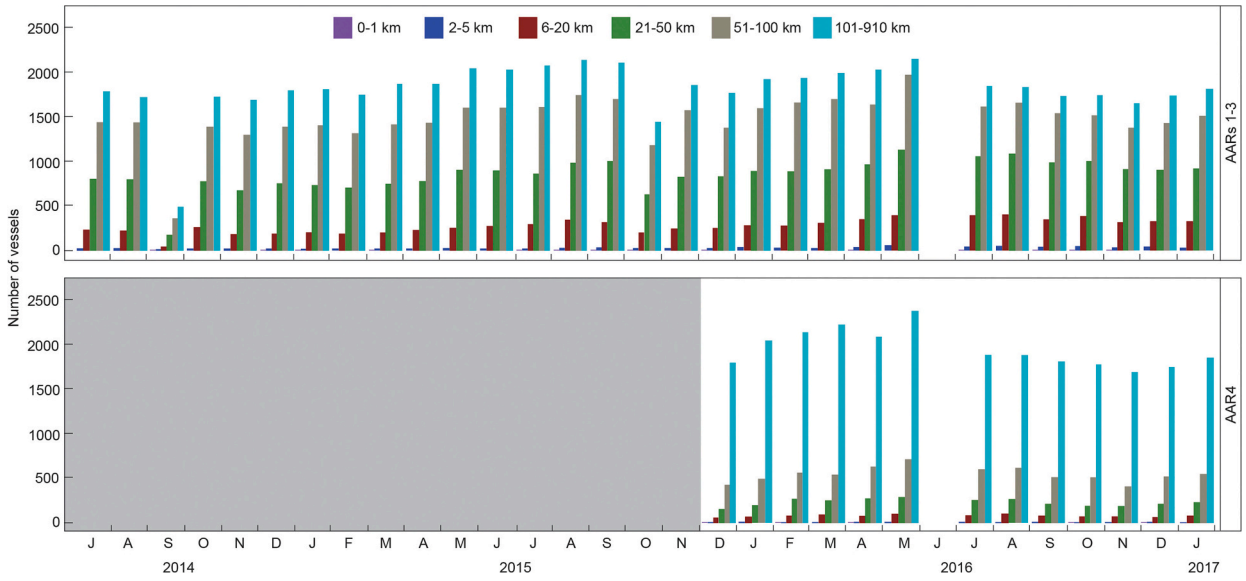

Fig. 8. Number of vessels per month at different distance intervals from the location of autonomous acoustic recorders (AARs) 1–3 (top panel) and AAR4 (bottom panel). No ship tracks were found from June 2016 but does not imply complete absence of ships. Grey shaded area for AAR4 location indicates time without passive acoustic monitoring effort.

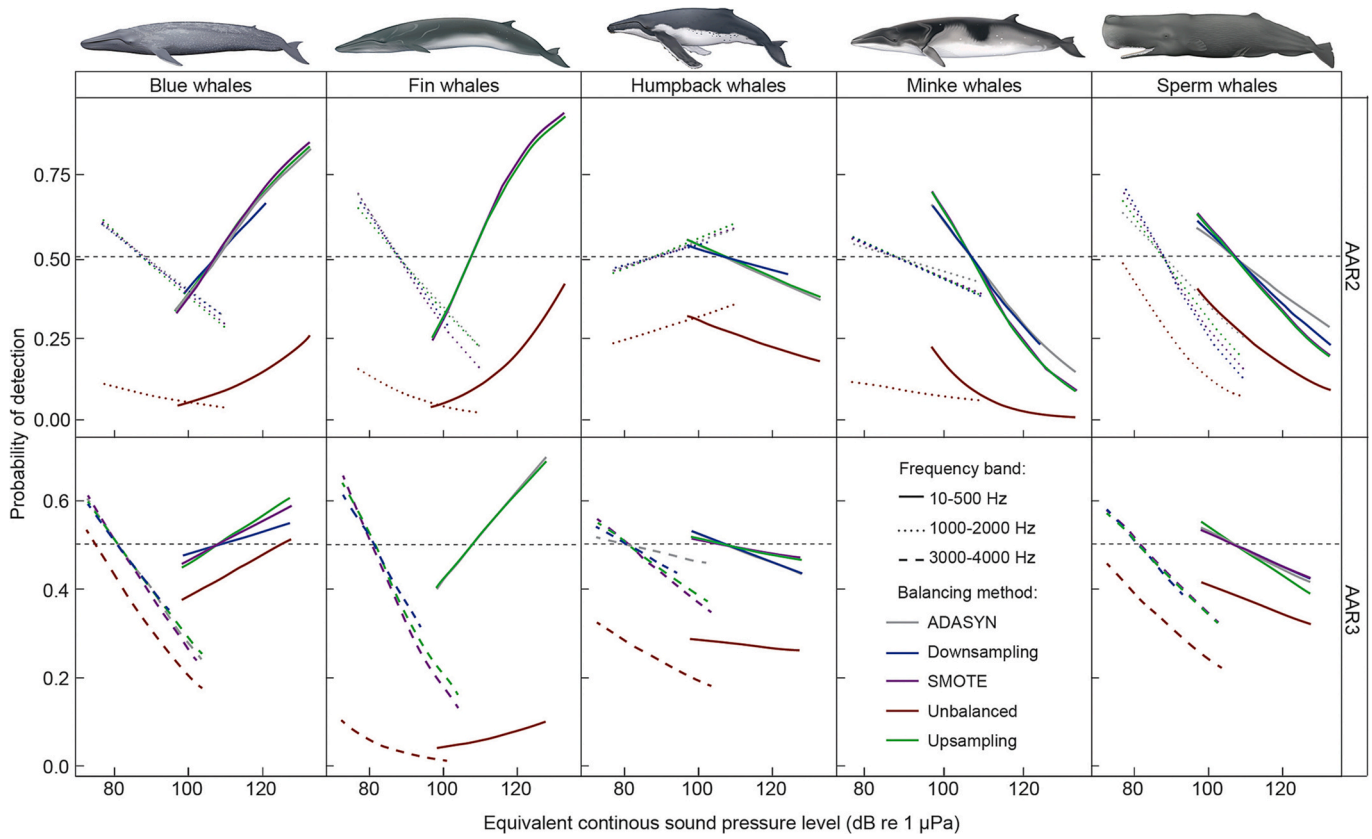

Fig. 9. Generalized linear model (GLM) partial effect of ambient equivalent continuous sound pressure level on probability of detecting whales acoustically at the lowest (continuous lines) and highest (dashed lines) frequency bands for each autonomous acoustic recorder (AAR) stations based on different sample balancing methods. Empty boxes represent cases where whale acoustic detection was <20 occurrences for that species. The dashed horizontal line at 0.5 represents the 50% point for whale acoustic detection at a given noise level. Y- and x-axes scales are different among plots. For legibility purpose, the 95% confidence intervals of these GLM curves are shown in Figs. S16–S20.

influence of ocean current speed associated with water depths where those AAR moorings were deployed in the Benguela ecosystem (Shannon, 2009). Recorded noise levels, including pseudo-noise from AARs 1 and 2, were not comparable during period of recording overlap (Fig. 3), although the two instruments were closely spaced (~5 km apart). This discrepancy could only be explained by the thermocline depth relative to the hydrophone depth in the water column as the thermocline acted as an acoustic barrier to the deeper instrument (Shabangu et al., 2021). On the other hand, sound speed profiles for this region did not change significantly between seasons to have induced seasonal differences in noise detection between AAR locations (Shabangu and Andrew, 2020). Gradual decline in power spectra above 100 Hz for AARs 2 and 3 shown in SPD plots reflects that low frequency sounds below 100 Hz (attributed to shipping) dominated the soundscape of the west coast of South Africa and that acoustic masking would possibly be high for species vocalizing below that frequency level (Clark et al., 2009; Cholewiak et al., 2018). Nonetheless, power spectra peaks in the vocalization frequency range of blue and fin whales before and after filtering out pseudo-noise from AARs 2 and 3 data suggests that whales contribute to the region's soundscape as well.

Recorders deployed approximately at 200 m (AARs 1 and 4) below the sea surface had higher recorded L_{eq} than those deployed at 300 m (AARs 2 and 3), an indication that recorders deployed in relatively shallow waters are more susceptible to contamination by pseudo-noise than those deployed in relatively deep waters where the current speed is slightly weaker as ocean current speeds decrease with depth in this region (Carr, 2017). Ocean current speed induced pseudo-noise is a downside of midwater recorder deployments where noise due to mooring vibration and water flow on the hydrophone can sometimes be very prominent (Strasberg, 1979). It is recommended for future deployments that recorders be deployed as close to the seafloor as possible to avoid such high noise levels induced by ocean current speed in the offshore environment. Implementing flow noise mitigation techniques in the recording equipment like those used in this study is also recommended.

The average L_{eq} at 10–500 Hz frequency band dropped below 110 dB re 1 µPa after filtering out pseudo-noise induced by the ocean current, indicating the effectiveness of the filter at >11 cm s^{-1} current speed at eliminating pseudo-noise. According to results of our RF models, ambient noise levels from the 10–500 Hz frequency band were strongly predicted by the distance of the closest vessel to both AARs 2 and 3 location, an indication that noise pollution by vessel traffic (an anthropogenic activity) is prominent in this region. The influence of distance of the closest vessel to the recorder location on the ambient noise decreased with range due to transmission loss as the sound propagates in the water column (Urick, 1983; McKenna et al., 2012). The 8 dB contribution of vessel traffic to the ambient noise level of AAR3 from vessels within 35 km radius is lower than the 15–20 dB contribution reported by McKenna et al. (2012) from vessel within 4 km distance to the recorder, the difference between the two studies might be due to different types of vessels studied, water depth, environment, and distance of vessels to the recorders. Wind speed was the most important predictor of ambient L_{eq} for the other frequency bands (500–1000, 1000–2000, 2000–3000 and 3000–4000 Hz) from AARs 2 and 3, as wind speed mainly contributes noise above 500 Hz frequency bands (Wenz, 1962; Knudsen et al., 1948; Estabrook et al., 2016; Duarte et al., 2021). Nonetheless, vessel traffic also contributed towards the 500–1000 Hz frequency band noise from AAR3, supporting previous observations that

vessel noise can sometimes contribute to noise above 500 Hz when vessels are at close range (Wenz, 1962; McKenna et al., 2013; Duarte et al., 2021). Ocean wave action contribution to the ambient noise at different frequency bands was moderately important as this source was overshadowed by other more prominent sources. Total precipitation had moderate to low influence on the ambient noise levels as noise from other sources dominated the soundscape. Late autumn through spring (April through October) had the highest influence on the low frequency band (10–500 Hz) ambient noise level from AARs 2 and 3 since those months were characterized by high vessel traffic (Figs. 1 and 8). Spring and summer (September through December) had a high effect on the high frequency bands (1000–2000 and 3000–4000 Hz) as those seasons are associated with high wind speeds (Shannon, 2009).

We found a good relationship between ambient noise levels at 10–500 Hz frequency band and the number of vessels for AARs 2 and 3 as found by previous studies (McKenna et al., 2013; Merchant et al., 2012). Our RF model results for ambient noise indicated that marine vessel traffic contributed significantly to underwater noise below 500 Hz off the west coast of South Africa (an important acoustic habitat) between 2014 and 2017, which is concerning given its effect on marine mammal communication space (e.g. Cholewiak et al., 2018). Reduction of vessel cruising speed might be critical for reducing noise levels in this region as moderate to high speeds (8.5–20 knots) were highly influential at predicting ambient noise levels but that will result in vessels spending more time in the region and challenging operability of ships (McKenna et al., 2013). Although the vessel traffic data used in this study were derived from multiple sources, these data possibly still excluded smaller vessels and those vessels that did not log their AIS positions. Correspondingly, our results provide conservative but realistic effects of vessel traffic on ambient noise levels. High vessel traffic around AARs 2 and 3 (Figs. 1 and 8) of >1100 vessels per month within the 80 km radius increased ambient noise level in this region and in turn will likely result in fatal vessel collisions with whales as recently (April 2021) seen with an Antarctic blue whale off Namibia and previously observed with southern right whales in South African waters (Best et al., 2001).

Miksis-Olds and Nichols (2016) found no common global trend in noise level variation as noise increased while it decreased in certain parts of the world, our results indicated that the ambient noise levels vary seasonally but not inter-annually between AARs 2 and 3 off the west coast of South Africa. Importantly, these results highlight the need for acoustic monitoring over decades to delineate the possible temporal and spatial trends in noise levels and effects of anthropogenic activities. Multidecadal underwater noise research conducted off California coast by Andrew et al. (2002) is a good example of the kind of research needed in this region. No seismic survey signals were detected from any of the AARs as either there were no seismic surveys or surveys were conducted far from AAR locations (330 and 510 km away from AARs 3 and 4 respectively to the nearest track line of the 2D seismic survey conducted in summer 2015/2016) during the deployment period of recorders (Fig. S26). It is encouraging that no noise pollution due to seismic survey signals was detected off the west coast of South Africa, a signal that this ecosystem was then not negatively affected by that anthropogenic activity and that the environment in South African waters does not support long range propagation of seismic signals.

4.2. Whale detection relative to ambient noise levels

Specific spectra peaks were observed within vocalization frequencies of blue and fin whales in the ambient SPD, indicating that they contributed to the soundscape of this region. The lack of specific spectra peaks within vocalization frequencies of other whales in the recorded SPD indicates that biological sounds were often overshadowed by the ambient noise. Nonetheless, whale calls were detected in noisy conditions (Fig. S15) indicating that some calls potentially originating from nearby whales can be detected even in noisy conditions and that visual detection of whale calls was less influenced by noise compared to automated call detectors. Specifically, blue and fin whale acoustic detectability increased with ambient noise level (highly predicted by the closest vessel distance to the recorder location) in the low frequency band (10–500 Hz), an indication that these whales could have increased their vocalization rate in the presence of noise from vessels to overcome the ambient noise levels (Lombard response) to maintain acoustic contact with conspecifics. Similar vocalization rate increase in the presence of vessels has been observed for blue whales in the Southern California Bight (Melcón et al., 2012). Fin whales were previously reported to change their call characteristics in response to shipping and seismic noise (Castellote et al., 2012).

Blue and fin whale probability of detection decreased with the increase in highest frequency band (1000–2000 and 3000–4000 Hz) ambient noise levels from AARs 2 and 3 although those bands are above the vocalization frequency of these whales but corresponds to noise induced by wind speed. This observed decrease in detectability of blue and fin whale calls with increasing ambient noise at the highest frequency bands indicates that whales can be affected by noise above their vocalization frequency (e.g. Melcón et al., 2012). Likewise, Shabangu et al. (2017, 2019) found call rates of blue and fin whales to decrease with increasing wind speed in the Antarctic and South African waters, which indicates that geophonic noise affected whale call detectability. Sub-surface air bubbles that were formed during high wind speeds could also have attenuated whale calls that should have been reflected into the water column by the sea surface in good weather conditions as Shabangu et al. (2014) have shown that echosounder pings transmitted on the sea surface were attenuated by wind-induced air bubbles.

The probability of detecting whale calls at the lowest and highest frequency bands indicated that ambient noise levels reduced the detectability of humpback, minke and sperm whales and would have also masked whale calls that could have been detected by AARs 2 and 3. Humpback whale probability of detection from AAR2 showed a different pattern to the above where detectability increased with the noise associated with wind speed at the highest frequency band, displaying Lombard response to geophonic noise. Sperm whale acoustic detectability was mainly influenced by noise levels from both low and high frequency bands, as clicks of this odontocete sweep through a wide frequency range and their acoustic detection decreased with increasing wind speed in South African waters (Shabangu and Andrew, 2020). Responses of the aforementioned whale species to this vessel noise will likely not only affect their health and welfare, but also increase their risk of colliding with vessels when considering Lloyd's mirror effect (Gerstein et al., 2005).

Species-specific acoustic detectability was observed as acoustic detectability increased with noise level for some species while it decreased for some species, indicating that species-specific noise mitigation measures are needed. Different noise levels at which the average 50% point was passed for each species between sites further suggest that there was some site-specific difference in whale detection. The acoustic communication space (detection range) of these whales was significantly reduced by 80% or more by underwater noise as indicated by this study (Fig. 9) and elsewhere (Cholewiak et al., 2018). Additionally, as the whale acoustic occurrence peaked in winter and spring, whale calls contributed to the soundscape before and after filtering out pseudo-noise. GLM curves on acoustic detectability of whales including pseudo-noise show that whales' detectability decreased further with increasing noise (Figs. S21–S25), an important technical consideration for this study but not for the global whale detectability as that noise is not audible in the soundscape. Deployment of animal-borne passive acoustic recording tags will provide direct information on the behavioural responses of these whales to both anthropogenic and geophonic noise. Different responses of the studied whales to noise indicate that species-specific and flexible mitigation measures are necessary to protect these marine mammals and other marine organisms from the ever-changing anthropogenic impacts on the marine ecosystem, including that of the west coast of South Africa which is currently undergoing

ocean economic development and growth.

5. Conclusion

We provide the first empirical ambient noise level measurements for the west coast of South Africa, and the quantitative description of predictors of ambient noise levels together with the effects of underwater noise on acoustic detection of some of the endangered, threatened, and vulnerable whale species in the Southern Hemisphere. Underwater noise levels varied spectrally and temporally between recording sites. Underwater noise pollution from vessel traffic contributed significantly to the ambient noise level below 500 Hz. Noise induced by wind speed dominated soundscape above 500 Hz, an indication of the contribution of geophonic sources to the soundscape of this region. The four baleen whale species showed species-specific responses to ambient noise levels. On the other hand, sperm whale acoustic detection was influenced by noise levels from both the low and high frequency bands given its broadband vocalization frequency range. The soundscape ecology paradigm of the west coast of South Africa appears not to be affected by seismic survey activities but by vessel traffic. Results of this study advance the scientific understanding of the relationship among anthropogenic activities, natural processes, and marine mammals, which is critically important for the conservation, management, and sustainability of this marine ecosystem.

Funding

This work was supported by the South African National Antarctic Programme [Grant/Award Number: SNA 2011112500003].

CRediT authorship contribution statement

Fannie W. Shabangu: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Dawit Yemane:** Data curation, Formal analysis, Methodology, Software, Resources, Validation, Visualization, Writing – original draft, Writing – review & editing. **George Best:** Data curation, Resources, Writing – original draft. **Bobbi J. Estabrook:** Data curation, Formal analysis, Methodology, Software, Resources, Validation, Visualization, Writing – original draft.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

We are sincerely grateful to Prof Ken Findlay, Meredith Thornton, Marcel van den Berg, Bradley Blows, Chris Wilkinson, Captains, and crew of RV *Algoa* for their invaluable contribution towards the collection of the acoustic data used in this study. South African oceanographers involved in the South Atlantic Meridional Overturning Circulation global project are acknowledged for kindly deploying AARs on their moorings. Special gratitude to Ishmail Letsheleha for providing output on blue and fin whale acoustic detection for AARs 3 and 4. James Maitland of South African Environmental Observation Network is acknowledged for his assistance to acquire and download the ocean current speed data. We thank Catherine Boucher for her professional graphical assistance with figures in this manuscript. Many thanks to Uko Gorter for generously providing whale illustrations. John Kim of ORBCOMM is thanked for responding to our numerous technical questions about the vessel traffic data. We appreciate invaluable comments provided by Dimitri Ponirakis and Tom Purdon on earlier versions of this manuscript. Two anonymous reviewers are acknowledged for providing helpful comments and suggestions on this manuscript.

Appendix A. Supplementary material

Supplementary material to this article can be found online at https://doi.org/10.1016/j.marpolbul.2022.114122.

References

Andrew, R.K., Howe, B.M., Mercer, J.A., Dzieciuch, M.A., 2002. Ocean ambient sound: comparing the 1960s with the 1990s for a receiver off the California coast. Acoust. Res. Lett. Online 3, 65.

Au, W.W.L., Hastings, M.C., 2008. Principles of Marine Bioacoustics. Springer, New York.

Best, P.B., Peddenmors, V.M., Rice, N., 2001. Mortalities of right whales and related anthropogenic factors in south African waters, 1963–1998. J. Cetacean Res. Manage. (Special Issue) 2, 171–176.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.

Carr, M.D., 2017. Investigating the Local Circulation of the Southeast Cape Basin. University of Cape Town. Master Thesis. 81 pp.

Castellote, M., Clark, C., Lammers, M., 2012. Acoustic and behavioural changes by fin whales (Balaenoptera physalus) in response to shipping and airgun noise. Biol. Conserv. 147, 115–122.

Cerchio, S., Strindberg, S., Collins, T., Bennett, C., Rosenbaum, H., 2014. Seismic surveys negatively affect humpback whale singing activity off northern Angola. PLoS ONE 9 (3), e86464.

Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P., 2002. SMOTE: synthetic minority over-sampling technique. J. Artif. Intell. Res. 16, 321–357.

Cholewiak, D., Clark, C.W., Ponirakis, D., Frankel, A., Hatch, L.T., Risch, D., Stanistreet, J.E., Thompson, M., Vu, E., van Parijs, S.M., 2018. Communicating amidst the noise: modeling the aggregate influence of ambient and vessel noise on baleen whale communication space in a national marine sanctuary. Endang. Species Res. 36, 59–75.

Clark, C.W., Ellison, W.T., Southall, B.L., Hatch, L., van Parijs, S.M., Frankel, A., Ponirakis, D., 2009. Acoustic masking in marine ecosystems: intuitions, analysis, and implication. Mar. Ecol. Prog. Ser. 395, 201–222.

Clark, C.W., Rice, A.N., Ponirakis, D.W., Dugan, P.J., 2011. Marine acoustic ecologies and acoustic habitats: concepts, metrics, and realities. J. Acoust. Soc. Am. 130, 2320.

Copernicus Climate Change Service, 2017. ERA5: fifth generation of ECMWF atmospheric reanalyses of the global climate. copernicus climate change service climate data store (CDS). Available at:. ECMWF https://cds.climate.copernicus.eu/cdsapp#!/home. (Accessed 20 January 2021).

Dobson, A.J., 1990. An Introduction to Generalized Linear Models. Chapman and Hall, London.

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., Juanes, F., 2021. The soundscape of the Anthropocene Ocean. Science 371, eaba4658.

Dugan, P.J., Ponirakis, D.W., Zollweg, J.A., Pitzrick, M.S., Morano, J.L., Warde, A.M., Rice, A.N., Clark, C.W., van Parijs, S.M., 2011. SEDNA—bioacoustic analysis toolbox. IEEE OCEANS 2011, 1–10.

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

Erbe, C., Marley, S.A., Schoeman, R.P., Smith, J.N., Trigg, L.E., Embling, C.B., 2019. The effects of ship noise on marine mammals - a review. Front. Mar. Sci. 6, 606.

Estabrook, B.J., Ponirakis, D.W., Clark, C.W., Rice, A.N., 2016. Widespread spatial and temporal extent of anthropogenic noise across the northeastern Gulf of Mexico shelf ecosystem. Endanger. Species Res. 30, 267–282.

Gerstein, E.R., Blue, J.E., Forysthe, S.E., 2005. The acoustics of vessel collisions with marine mammals. In: Proceedings of the Oceans, 2005. MTS/IEEE, Washington, DC, pp. 1190–1197.

Greenwell, B.M., Boehmke, B.C., McCarthy, A.J., 2018. A Simple and Effective Model-based Variable Importance Measure. arXiv preprint arXiv:1805.04755.

He, H., Bai, Y., Garcia, E.A., Li, S., 2008. ADASYN: Adaptive synthetic sampling approach for imbalanced learning. In: Proc. 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence), Hong Kong, pp. 1322–1328.

Heiler, J., Elwen, S.H., Kriesell, H.J., Gridley, T., 2016. Changes in bottlenose dolphin whistle parameters related to vessel presence, surface behaviour and group composition. Anim. Behav. 117, 167–177.

Hijmans, R.J., 2020. raster: geographic data analysis and modeling. R package version 3.1-5. https://CRAN.R-project.org/package=raster. (Accessed 21 January 2021).

Hildebrand, J.A., 2009. Anthropogenic and natural sources of ambient noise in the ocean. Mar. Ecol. Prog. Ser. 395, 5–20.

Jolliffe, I.T., Cadima, J., 2016. Principal component analysis: a review and recent developments. Phil. Trans. R. Soc. A 374, 20150202.
Knudsen, V.O., Alford, R.S., Emling, J.W., 1948. Underwater ambient noise. J. Mar. Res. 7, 410–429.
Koper, R.P., Plön, S., 2012. The potential impacts of anthropogenic noise on marine animals and recommendations for research in South Africa. In: EWT Research & Technical Paper No. 1. Endangered Wildlife Trust, South Africa, 103 pp.
Koper, R.P., Erbe, C., Preez, D.R.D., Plön, S., 2016. Comparison of soundscape contributors between two neighboring southern right whale nursing areas along the south african coast. Proc. Meet. Acoust. 27, 070014.
Lamont, T., van den Berg, M., 2021a. Long-term Observations of Hourly Currents Along the SAMBA Transect at SAMBA Mooring 7, September 2014 - December 2015 [Data Set]. Department of Environment, Forestry and Fisheries. https://doi.org/10.15493/DEA.MIMS.20210321.
Lamont, T., van den Berg, M., 2021b. Long-term Observations of Hourly Currents Along the SAMBA Transect at SAMBA Mooring 7, December 2015 - April 2017 [Data Set]. Retrieved from. Department of Environment, Forestry and Fisheries. https://doi.org/10.15493/DEA.MIMS.20210322.
Lamont, T., van den Berg, M., 2021c. Long-term Observations of Hourly Currents Along the SAMBA Transect at SAMBA Mooring 10, December 2015 - April 2017 [Data Set]. Department of Environment, Forestry and Fisheries. https://doi.org/10.15493/DEA.MIMS.20210335.
Letsheleha, I.S., Shabangu, F.W., Farrell, D., Andrew, R.K., la Grange, P.L., Findlay, K.P., 2022. Year-round acoustic monitoring of Antarctic blue and in whales in relation to environmental conditions off the west coast of South Africa. Mar. Biol. 169, 41.
McKenna, M.F., Ross, D., Wiggins, S.M., Hildebrand, J.A., 2012. Underwater radiated noise from modern commercial ships. J. Acoust. Soc. Am. 131, 92–103.
McKenna, M., Wiggins, S., Hildebrand, J., 2013. Relationship between container ship underwater noise levels and ship design, operational and oceanographic conditions. Sci. Rep. 3, 1760.
Melcón, M.L., Cummins, A.J., Kerosky, S.M., Roche, L.K., Wiggins, S.M., Hildebrand, J. A., 2012. Blue whales respond to anthropogenic noise. PLoS ONE 7 (2), e32681.
Merchant, N.D., Witt, M.J., Blondel, P., Godley, B.J., Smith, G.H., 2012. Assessing sound exposure from shipping in coastal waters using a single hydrophone and automatic identification system (AIS) data. Mar. Pollut. Bull. 64 (7), 1320–1329.
Miksis-Olds, J.L., Nichols, S.M., 2016. Is low frequency ocean sound increasing globally? J. Acoust. Soc. Am. 139, 501–511.
Nallamuthu, N., 2020. Handling imbalanced data - Machine learning, computer vision and NLP. https://www.analyticsvidhya.com/blog/2020/11/handling-imbalanced-data-machine-learning-computer-vision-and-nlp/. (Accessed 8 September 2021).
Operation Phakisa, 2015. Oceans economy review workshop. Oceans economy review workshop. Operation Phakisa, Pretoria. https://www.operationphakisa.gov.za/operations/oel/Ocean%20Economy%20Lab%20Documents/0.%20Oceans%20Economy%20Review%20Workshop%2015%20October%202015%20Summary%20Report.pdf. (Accessed 22 January 2022).
Purdon, J., 2018. Calming the waves: using legislation to protect marine life from seismic surveys. In: South African Institute of International Affairs: Policy Insight, 58, pp. 1–16.
Purdon, J., Shabangu, F.W., Pienaar, M., Somers, M.J., Findlay, K., 2020. Cetacean species richness in relation to anthropogenic impacts and areas of protection in South Africa's mainland exclusive economic zone. Ocean Coast. Manage. 197, 105292.
R Core Team, 2021. R: a language and environment for statistical computing. URL. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/. (Accessed 8 October 2021).
Schoeman, R.P., Erbe, C., Plön, S., 2022. Underwater chatter for the win: a first assessment of underwater soundscapes in two bays along the eastern Cape Coast of South Africa. J. Mar. Sci. Eng. 10, 746.
Scholbeck, C.A., Molnar, C., Heumann, C., Bischl, B., Casalicchio, G., 2020. Sampling, intervention, prediction, aggregation: a generalized framework for model-agnostic interpretations. In: Cellier, P., Driessens, K. (Eds.), Machine Learning and Knowledge Discovery in Databases. Springer, Cham, pp. 205–216.
Schonlau, M., Zou, R.Y., 2020. The random forest algorithm for statistical learning. Stata J. 20, 3–29.
Shabangu, F.W., Andrew, R.K., 2020. Clicking throughout the year: sperm whale clicks in relation to environmental conditions off the west coast of South Africa. Endanger. Species Res. 43, 475–494.
Shabangu, F.W., Kowarski, K.A., 2022. The beat goes on: humpback whale song seasonality in Antarctic and south african waters. Front. Mar. Sci. 9, 827324.
Shabangu, F.W., Ona, E., Yemane, D., 2014. Measurements of acoustic attenuation at 38 kHz by wind-induced air bubbles with suggested correction factors for hull-mounted transducers. Fish. Res. 151, 47–56.
Shabangu, F.W., Yemane, D., Stafford, K.M., Ensor, P., Findlay, K.P., 2017. Modelling the effects of environmental conditions on the acoustic occurrence and behaviour of Antarctic blue whales. PLoS One 12 (2), e0172705.
Shabangu, F.W., Findlay, K.P., Yemane, D., Stafford, K.M., van den Berg, M., Blows, B., Andrew, R.K., 2019. Seasonal occurrence and diel calling behaviour of Antarctic blue whales and fin whales in relation to environmental conditions off the west coast of South Africa. J. Mar. Syst. 190, 25–39.
Shabangu, F.W., Findlay, K., Stafford, K.M., 2020. Seasonal acoustic occurrence, diel-vocalizing patterns and bioduck call-type composition of Antarctic minke whales off the west coast of South Africa and the Maud Rise, Antarctica. Mar. Mammal Sci. 36, 658–675.
Shabangu, F.W., Andrew, R.K., Findlay, K., 2021. Acoustic occurrence, diel-vocalizing pattern and detection ranges of southern right whale gunshot sounds off South Africa's west coast. Mar. Mammal Sci. 37, 733–750.
Shannon, V., 2009. Benguela current. In: Steele, J.H., Thorpe, S.A., Turekian, K.K. (Eds.), Encyclopedia of Ocean Sciences, 2nd edition. Academic Press, London, pp. 316–327.
Southall, B., Bowles, A., Ellison, W., Finneran, J., Gentry, R., Greene, C., Kastak, D., Ketten, D., Miller, J., Nachtigall, P., 2007. Marine mammal noise exposure criteria: initial scientific recommendations. Aquat. Mammal. 33, 411–521.
Strasberg, M., 1979. Nonacoustic noise interference in measurements of infrasonic ambient noise. J. Acoust. Soc. Am. 66 (5), 1487–1493.
Strobl, C., Malley, J., Tutz, G., 2009. An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. Psychol. Methods 14 (4), 323–348.
Urick, R.J., 1983. Principles of Underwater Sound, 3rd ed. McGraw-Hill, New York.
Urick, R.J., 1986. Ambient Noise in the Sea. Peninsula Publishing, Los Altos Hills, CA.
Wenz, G.M., 1962. Acoustic ambient noise in the ocean: spectra and sources. J. Acoust. Soc. Am. 34, 1936–1956.
Zuma, J., 2014. Address by President Zuma at the operation Phakisa: unlocking the economic potential of the ocean economy open day. Retrieved from. https://www.operationphakisa.gov.za/cc/Documents/Open%20Day%20Operation%20Phakisa%20President%20Speech.pdf. (Accessed 20 January 2022).