Clicking throughout the year: sperm whale clicks in relation to environmental conditions off the west coast of South Africa

Fannie W. Shabangu[1,2,*], Rex K. Andrew[3]

[1]Fisheries Management Branch, Department of Environment, Forestry and Fisheries, Foreshore, Cape Town 8018, South Africa
[2]Mammal Research Institute Whale Unit, University of Pretoria, Private Bag X20, Hatfield, Pretoria 0028, South Africa
[3]Applied Physics Laboratory, University of Washington, Seattle, WA 98105, USA

ABSTRACT: Knowledge of cetacean occurrence and behaviour in southern African waters is limited, and passive acoustic monitoring has the potential to address this gap efficiently. Seasonal acoustic occurrence and diel-vocalizing patterns of sperm whales in relation to environmental conditions are described here using passive acoustic monitoring data collected off the west coast of South Africa. Four autonomous acoustic recorders (AARs) were deployed on 3 oceanographic moorings from July 2014 to January 2017. Sperm whale clicks were detected year round in most recording sites, with peaks in acoustic occurrence in summer and late winter through spring. Diel-vocalizing patterns were detected in winter, spring and summer. Higher percentages of sperm whale clicks were recorded by AARs deployed at 1100 m water depth compared to those concurrently deployed at 850 and 4500 m, likely inferring that the whales exhibited some preference to water depths around 1100 m. Acoustic propagation modelling suggested a maximum detection range of 83 km in winter for sperm whale clicks produced at 1100 m. Random forest models classified daylight regime, sea surface height anomaly and month of the year as the most important predictors of sperm whale acoustic occurrence. The continuous acoustic occurrence of sperm whales suggests that the study area supports large biomasses of prey to sustain this species' food requirements year round. This is the first study to describe the seasonal acoustic occurrence and diel-vocalizing patterns of sperm whales off the west coast of South Africa, extending knowledge of the species previously available only through whaling records.

KEY WORDS: Sperm whales · Acoustic occurrence · Passive acoustic monitoring · Diel behaviour · Detection ranges · Environmental conditions · Random forest model · South Africa

1. INTRODUCTION

The seasonal acoustic occurrence and diel-vocalizing patterns of sperm whales *Physeter macrocephalus* off the west coast of South Africa have not previously been investigated owing to the lack of acoustic research effort in this region. Current knowledge about seasonal occurrence, distribution and behaviour of sperm whales off the west coast of South Africa derives solely from whale catch statistics (e.g. Best 1969,

1999, 2007, Elwen et al. 2016). Sperm whales were harvested in South African waters on the west coast (Donkergat whaling station) and east coast (Durban whaling grounds), and the total number of whales harvested in the southern African region is unknown (Gambell 1967, Best 1969, 2007, Findlay & Best 2016). Donkergat whaling station (see Fig. 1) was land based in Saldanha Bay (33° 05′ S, 18° 00′ E), and operated from 1909 to 1967, during which time about 10 700 sperm whales were processed (Best 1974). Best (1969)

reported that females and small males were caught mostly in autumn (March/April), whilst medium-sized and large males were caught mostly in autumn (April/May) and spring (August–October).

Open-boat whaling (1712–1920s) had reduced the pre-whaling worldwide population (approximately 1100000) by 29% in 1880, and modern whaling (1904–1988) further reduced the population to 32% of the pre-whaling population by the 1990s (Whitehead 2002). In total, over 400000 sperm whales were killed in the Southern Hemisphere between 1900 and 2005 (Clapham & Baker 2009). Recent quantitative analysis of sperm whale population trends indicate that there are limited signs of recovery (Branch & Butterworth 2001, Whitehead 2002) whereby population recovery might be heavily dependent on female survivorship rates (Chiquet et al. 2013). Thus, both the International Union for the Conservation of Nature Red List of Threatened Species (Taylor et al. 2008) and the Regional Red List (Elwen et al. 2016) currently classifies them as Vulnerable. Presently, there are no reliable estimates of the sperm whale population in the southern African region (Best 2007).

Sperm whales are the largest species of odontocetes (toothed whales), and are cosmopolitan in their distribution range. They tend to inhabit deep oceanic, ice-free waters in both hemispheres and can dive to considerable depths, sometimes >3000 m (Jefferson et al. 1993, 2015), feeding on a variety of prey species, but mainly squids (Jefferson et al. 1993, 2015, Best 1999, 2007). Tagging data in the Atlantic Ocean have revealed that these whales can dive to a maximum foraging depth of 1200 m in an average water depth of 2015 m (range 1287–2786 m), with indication of occasional foraging on the seafloor based on reflections of seafloor echoes (Watwood et al. 2006). Sperm whales are likely to be important in the ecological functioning of the Benguela ecosystem off the west of South Africa because they may play a critical role in the transfer of nutrients from the abyss to the euphotic zone, as observed in the Southern Ocean (Lavery et al. 2010). They are also apex predators, increasing the value of enriching our knowledge about their spatio-temporal patterns and about how and when changes happen (Fais et al. 2016). This species may also be important for the growing ecotourism industry in South Africa, especially in light of the South African government's initiative, Operation Phakisa (which means 'hurry up' in seSotho), which facilitates the rapid growth of the country's blue or oceans economy (van Wyk 2015). Such growth in the ocean economy through Operation Phakisa might increase anthropogenic effects on this species and many other vulnerable, yet recovering species. For example, Purdon et al. (2020a) showed that anthropogenic stressors such as climate-related stressors and shipping stressors associated with blue economy growth are increasing more rapidly across the South African exclusive economic zone, affecting species richness and the protection of multiple cetaceans including sperm whales.

Ecotourism associated with blue economy growth may also affect whales (Purdon et al. 2020a), even when operated under the criteria and regulations set by the Marine Living Resources Act 18 of 1998 that respect protected species in South African waters (Department of Environmental Affairs 2017). The west coast of South Africa is located in the southern Benguela ecosystem in the southeastern Atlantic Ocean (Fig. 1). The Benguela ecosystem extends from the south coast of South Africa equatorward to southern Angola (Shannon 2006). It is a wind-driven upwelling system characterized by the cold, nutrient-rich water of the Benguela Current (Fig. 1) that promotes biological productivity (Andrews & Hutchings 1980, Brown 1992, Shannon 2009). Purdon et al. (2020b) is the only study to have investigated the relationship between sperm whale distribution and environmental conditions in the southern African region. It is important that relationships between sperm whale occurrence and environmental conditions be established specifically in South African waters in light of the recently estimated global increase in climate change-related effects on environmental variables such as sea surface temperature (Halpern et al. 2019).

Sperm whales produce powerful clicks (with a maximum-recorded source level [SL] of 236 dB re 1 μPa @ 1 m; Møhl et al. 2003) that are used for communication and echolocation (Weilgart & Whitehead 1993, Møhl et al. 2000, Jaquet et al. 2001, Zimmer et al. 2005). Sperm whale clicks have been detected in different oceans with maximum detection ranges of between 15 and 35 km, mainly depending on the sea-state conditions (e.g. Mathias et al. 2013, André et al. 2017). These clicks contain energy in the frequency band from 10 Hz to about 32 kHz (Backus & Schevill 1966, Mellinger et al. 2004, André et al. 2017). Sperm whales produce clicks as: 'usual' or 'regular', 'slow', 'codas' and 'creaks'. 'Usual' clicks are the most commonly heard sperm whale clicks, with a 0.5 to 2 s interclick interval (ICI), and are emitted by diving whales (Whitehead & Weilgart 1991). 'Slow' clicks are linked to large, mature males and have an ICI of 5 to 7 s (Weilgart & Whitehead 1988, Whitehead 1993). 'Codas' are short, patterned sequences of 2 to 40 clicks with irregular repetition rates, have an ICI of 0.5 to 2 s, are used mainly for communica-

tion (Watkins & Schevill 1977, Pavan et al. 2000) and may have different functions depending on coda type (Antunes et al. 2011). 'Codas' are largely produced within the social groups (females with young individuals), and can be specific to regions or to individuals (Watkins & Schevill 1977, Rendell & Whitehead 2004, Oliveira et al. 2016). 'Creaks' are sequences of very rapid clicks, with up to 200 clicks per second, used for short-range echolocation (Gordon 1987, Whitehead & Weilgart 1991). The duration of sperm whale clicks depends on the click type and whale body size (Backus & Schevill 1966, Gordon 1991, Møhl et al. 2003, Growcott et al. 2011). These clicks are highly directional at peak frequencies between 10 and 15 kHz (Thode et al. 2002, Møhl et al. 2003), and they have an almost omnidirectional low frequency component between 1 and 4 kHz (Thode et al. 2002, Zimmer et al. 2005).

This study emphasizes the importance of passive acoustic monitoring data for investigating and defining the seasonal acoustic occurrence and diel-vocalizing patterns of sperm whales in relation to environmental conditions off a previously unstudied region of the west coast of South Africa. We identify important habitats and environmental drivers that best predict the seasonal acoustic occurrence of sperm whales off the west coast of South Africa. Additionally, we provide the first description of the seasonal acoustic occurrence and diel-vocalizing patterns of sperm whales in South African waters, which will assist with the protection and conservation of important habitats of this species from activities associated with ocean economy growth.

2. MATERIALS AND METHODS

2.1. Acoustic data collection

Acoustic data were collected over a period of 3 yr in 3 different sites off the west coast of South Africa, southeastern Atlantic Ocean (Fig. 1, Table 1) as part of the South African Blue Whale Project (SABWP) to study the acoustic occurrence and behaviour of Antarctic blue whales (Shabangu et al. 2019). Different sampling rates and sampling protocols (number of minutes recorded per hour) were applied to different autonomous underwater recorders (AARs; Autonomous Underwater Recorder for Acoustic Listening Model 2 version 04.1.3, Multi-Électronique) (Table 1) to record low frequency sounds of baleen whales and to preserve the AAR battery life. AARs were used to record the acoustic data in 3 passive

acoustic monitoring stations (Fig. 1). A gain of 22 dB was applied to all AARs, and factory-provided calibration settings were used (Table 1). AARs were deployed at different depths in the water column (Table 1) on oceanographic moorings. AAR1 was approximately 70 km from the coast whereas AARs 2 and 3 were 75 km from the coast, and the distance between the location of AAR1 and that of AARs 2 and 3 was 4.80 km (Fig. 1). AAR4 was approximately 240 km farther offshore than AARs 1, 2 and 3 (Fig. 1). AARs 2 and 3 were deployed in succession on the same oceanographic mooring at 1118 m water depth (Table 1). Acoustic data collected off the Maud Rise (65° S, 2.5° E), eastern Weddell Sea, Antarctica (Shabangu & Charif 2020, Shabangu et al. 2020a,b) were also analysed for sperm whale clicks. No sperm whale clicks were detected in that dataset, thus it was not further analysed.

2.2. Detection of clicks

Sperm whale clicks were visually detected using spectrograms and verified aurally when clicks were visually identified in Raven Pro (Bioacoustics Research Program 2017). Broadband pulses of 'usual' clicks, 'slow' clicks (Fig. 2), 'codas' and 'creaks' were used to determine acoustic occurrence of sperm whales. Acoustic presence of sperm whales was defined as the detection of any type of click within a sampling interval. A sampling interval is the time in which acoustic data were recorded depending on the set sampling protocol (Table 1); for example, acoustic data were recorded for 30 min for AAR1. Acoustic presence of sperm whales was used to define the acoustic occurrence of whales. Percentage of acoustic occurrence was defined as the ratio of the number of sampling intervals with acoustic presence to the total number of sampling intervals, in a given time period. For example, monthly percentage of acoustic occurrence was defined as the ratio of sampling intervals with acoustic presence per month to the total number of sampling intervals recorded per month, which was later translated to seasonal acoustic occurrence for 3 consecutive months of a season. Likewise, seasonal diel percentage of acoustic occurrence was defined as the number of sampling intervals with sperm whale clicks for each hour of the day for that season divided by the total number of sampling intervals recorded for each hour of the day for that season. The number of days with sperm whale clicks was derived from all days with acoustic occurrence of sperm whales. We did not classify clicks to types, as this

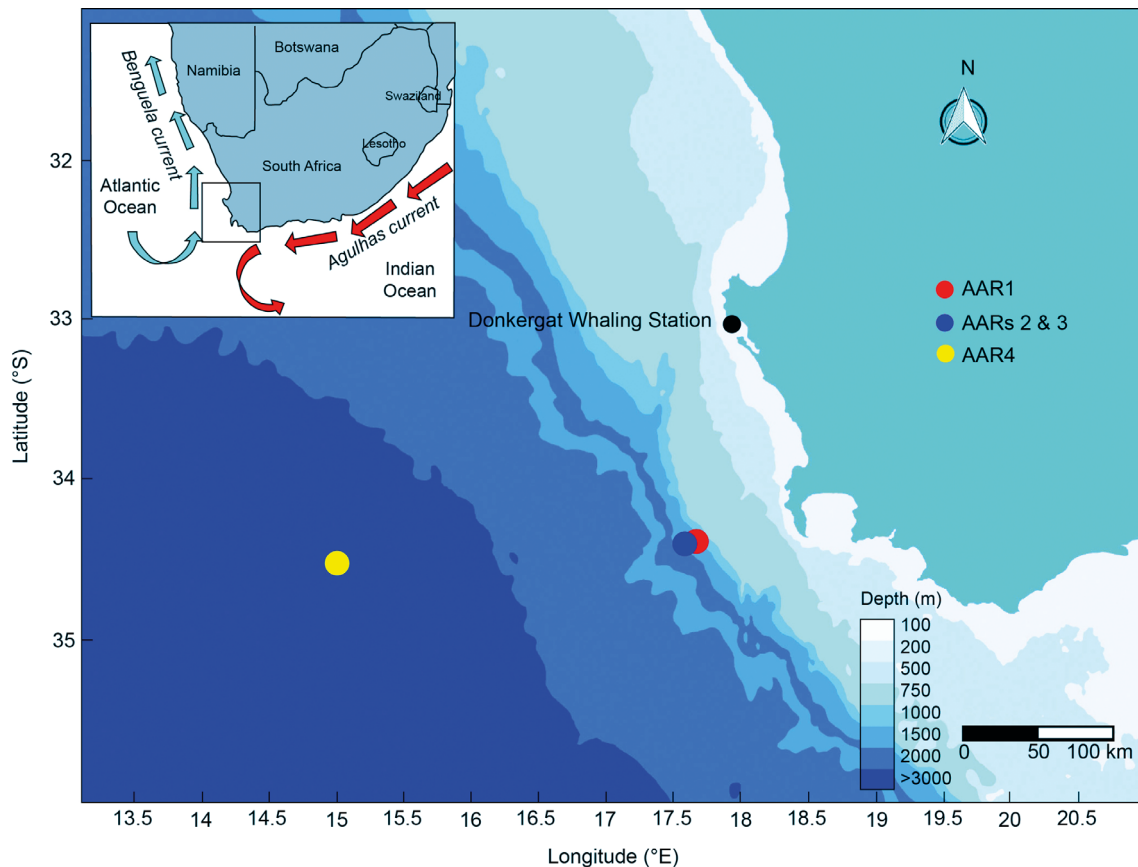Endang Species Res 43: 475–494, 2020

Fig. 1. Deployment positions of autonomous acoustic recorders (AARs) off the west coast of South Africa, Atlantic Ocean. The flow direction of the cold Benguela Current (blue arrows) on the west coast, Atlantic Ocean, and the warm Agulhas Current (red arrows) on the east coast, Indian Ocean, are shown on the insert map. The study area is highlighted with a box in the insert map

Table 1. Deployment details and recording settings of the 4 autonomous acoustic recorders (AARs) used in this study. AARs are numbered (ID) according to the order of their chronological deployment. Hydrophone sensitivities were obtained from the factory calibrations of the HTI-96-MIN hydrophones (High Tech). Dates are given as d/mo/yr

AAR ID	Latitude (S)	Longitude (E)	Water depth (m)	AAR depth (m)	Sampling rate (Hz)	Sampling protocol (min h^{-1})	Hydrophone sensitivity (dB re 1 V μPa^{-1})	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

study focused primarily on the acoustic presence of sperm whales.

Different daylight regimes were classified over different seasons in accordance with the altitude of the sun (dawn [nautical twilight], daytime, dusk [nautical twilight] and nighttime) by averaging hourly sun altitudes over austral seasons. Austral seasons of the year were used to parse the data into seasons: summer (December to February), autumn (March to May), win-

ter (June to August) and spring (September to November). Hourly sun altitudes for each day of the year from 34° 22′ S, 17° 37′ E were used for all AAR locations because all AARs were on the same latitudinal position. Data on sun altitudes were obtained from the United States Naval Observatory Astronomical Applications Department (http://aa.usno.navy.mil). Nautical dawn was defined as the period when the centre of the sun was geometrically between 0 and
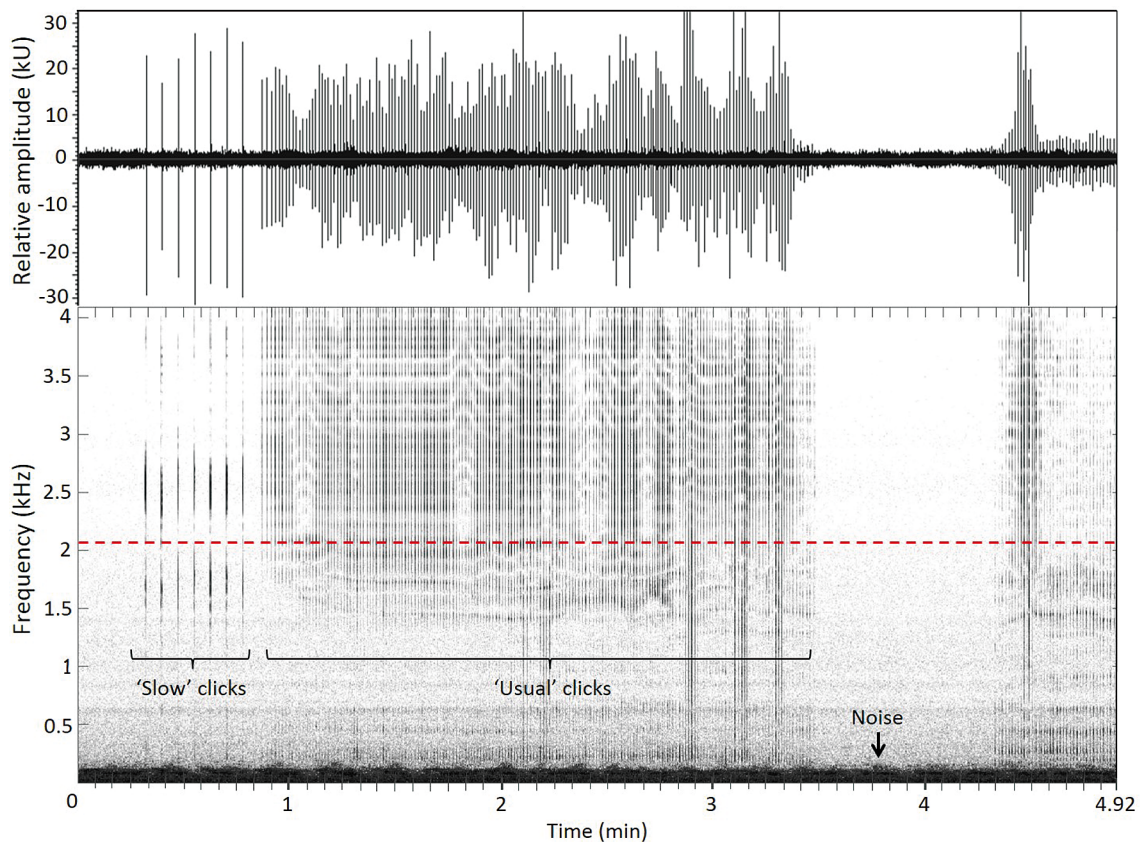
Fig. 2. Examples of sperm whale 'slow' and 'usual' clicks together with low frequency noise recorded over 5 h by autonomous acoustic recorder (AAR) 3 on 8 August 2016. Relative amplitudes of clicks (kU: kilo unit) are shown in the top panel as waveforms. The sampling effective bandwidth (2048 Hz) of AARs 1 and 2 is indicated (Red dashed line). The spectrogram was generated in XBAT (Figueroa 2006), using the parameters: frame size 1.59 s, 25% overlap, fast Fourier transform size 4096 points, Hann window

12° below the horizon before sunrise. Daytime was between sunrise and sunset, and nautical dusk was between sunset and the evening (defined as when the sun was less than 12° below the horizon). Nighttime was when the geometric centre of the sun was over 12° below the horizon between dusk and dawn. Since time of day is a circular variable, we smoothed the mean diel sperm whale acoustic occurrence per season through penalized cyclic cubic regression splines (Wood 2017) in a generalized additive model (GAM; Guisan et al. 2002). Diel sperm whale acoustic occurrence per season was calculated as the acoustic presence per hour in a season divided by the total number of sampling intervals for each hour in that season. Welch's 2 sample t-test was conducted to evaluate the inter-annual variability in percentage of acoustic occurrence between AARs 2 and 3 deployed on the same location over 2 consecutive years. GAMs were fitted using the 'mgcv' package (Wood 2001) in R (version 3.6.0; R Core Team 2019). Welch's t-tests were implemented using the 'DescTools' package (Signorell et al. 2020) in R.

2.3. Click detection range modelling

This analysis sought to calculate the seasonal detection range of broadband mid-frequency signals in the Benguela ecosystem, which were used to determine the spatio-temporal scales to extract environmental conditions around each AAR. Detection here is defined as signals with signal-to-noise ratios (SNRs) of 0 dB or greater to be intelligible by conspecifics as found by Miller et al. (1951), since there are no established SNR detection thresholds for sperm whales. Seasonal calculations were made for 4 scenarios to represent the year-round conditions at each deployment site. Signals were defined so as to represent the characteristics of sperm whale clicks as observed in the recordings. SNRs were modelled at the Nyquist frequency of each AAR recording, which was either 2048 or 4096 Hz depending on the sampling rate (Table 1). We did not model detection ranges at the low end of the recording frequency band (300 Hz) as there were high noise levels (NLs) at those frequencies, and the majority of energy of

sperm whale clicks is located at higher frequencies. The signal SL was defined as 190 dB re 1 µPa (equivalent to an SL density of 157.7 dB re 1 µPa^2 Hz^{-1} @ 1 m) at 2000 Hz, and 200 dB re 1 µPa (equivalent to an SL density of 164.3 dB re 1 µPa^2 Hz^{-1} @ 1 m) at 4000 Hz frequency bands as determined by Zimmer et al. (2005). SL density (ρ) was defined as SL normalized by the frequency bandwidth (BW):

$$SL\rho = SL - 10\log_{10}(BW) \qquad (1)$$

where BW was either 2000–300 Hz (for AARs 1 and 2) or 4000–300 Hz (for AARs 3 and 4). The calculations were at 2000 and 4000 Hz instead of 2048 and 4096 Hz for simplicity, which should not make more than 0.1 dB difference and is also less than the error in NL estimate. The lowest frequency at which the auto-spectra from all records seemed free of strum noise was roughly 300 Hz. Therefore, the frequency band of interest was defined to be 300 Hz to the recording's Nyquist frequency.

The transmission loss (TL) was computed using the BELLHOP beam tracing model (Porter 2011). The received level (RL) was calculated using the SL and TL:

$$RL = SL - TL \qquad (2)$$

To map the TL in Eq. (2) to a smooth model, the TL of the BELLHOP model (in dB) was modelled as:

$$TL_R = a[10\log_{10}(R)]^b + cR^2/10^4 \qquad (3)$$

where R is the detection range (m) with the parameters a, b and c estimated via least-squares and a Nelder-Mead optimizer over the range of 1 to 90 km (wherever there are BELLHOP solutions). The standard errors for a, b and c were computed from the observed information matrix, and the fit was generally observed to be quite good.

Environmental parameters used in the BELLHOP model were an annual average sound speed derived from the World Ocean Atlas 2013 (Boyer et al. 2013) and the Thermodynamic Equation of Seawater - 2010 (www.teos-10.org), bathymetry from the Smith-Sandwell database (Smith & Sandwell 1997) and Thorp attenuation (Thorp 1967, Focke et al. 1982). Some environmental variables were either very smooth, or the quantity tabulated was very smooth. For example, the World Ocean Atlas is a long-term average which will smooth out daily, episodic and/or recent features. Furthermore, the atlas values are tabulated on a 2-dimensional (2-D) grid with a resolution of approximately 100 km. Modelling typically uses 2-D interpolation from the 4 nearest grid points to generate profiles at specific places. Unfortunately, for AARs 1 to 3, which were all near the coast, several of these 'nearest' grid points were on land, and did not provide legitimate sound speed values, invalidating the 2-D interpolator. This problem affected all modelling locations near the moorings as well. As a proxy for the sound speed at the mooring, the climatology at −34.37° S, 16.0° E was chosen. As the modelling region was less than 80 km, which is under the resolution of the climatology, this sound speed profile was assumed to be valid throughout the modelling region. This location was chosen to not be a grid-registered location in order to include contributions from surrounding grid-registered values. Additionally, the extracted sound speed profiles did not go deep enough for modelling into the deeper open ocean. In all such cases, the sound speed profile was linearly extrapolated from the lowest 2 valid values down to 6000 m. Deeper levels do not change much from a depth-linear model, and the BELLHOP model also ignores values deeper than the local seafloor.

Given the tagging results of Watwood et al. (2006), vocalizing whales were assumed to be at depths of 200 and 800 m for AAR1; 200 and 1100 m for AARs 2 and 3; and 200, 1100 and 2700 m for AAR4. We did not assume whales to be at depths around 4000 m for AAR4, given the lack of evidence of whales vocalizing beyond 2700 m. A backstep of a few meters was applied to each AAR, so that we did not start simulating from the seafloor. Depths of potential callers along bearings towards the coastline were typically not as deep as the moorings themselves, due to shoaling bathymetry in those directions. For example, within a few tens of km of the moorings, the water depth in the direction of the coast was only about 300 m, and decreasing. Therefore, TL (Eq. 3) was only modelled along a bearing of 225° from the mooring. This corresponds to a direction out into the southern Atlantic Ocean. The seafloor depth becomes greater along this bearing away from the mooring.

The SNR depends on the signal SL (provided above) and the ambient NLs at the AAR (Fig. 3). The ambient NL was characterized very crudely by using 1 'sample' captured per season for each AAR; data for only 2 seasons were available from AAR1 (Table 1). Samples were 2.5 h long for AAR1, 1.7 h long for AAR2 and 1.3 h long for AARs 3 and 4, and each sample consisted of multiple continuous sub-segments concatenated together by the AAR software during data recording. Sub-segment durations are given in Table 1 as the sampling protocol for each AAR. The raw data were corrected to in-water units using the fac-
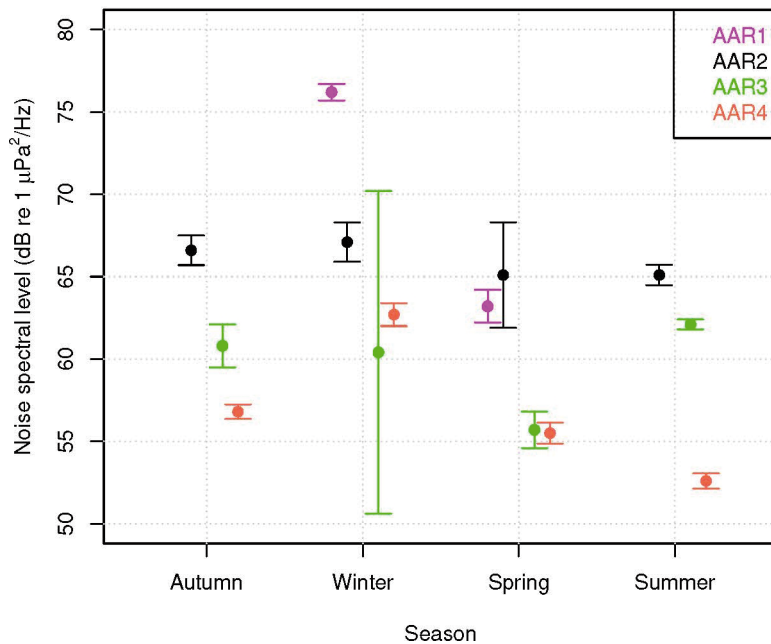
Fig. 3. Ambient noise levels versus austral season at the Nyquist frequency of each autonomous acoustic recorder (AAR). Each circle is an average. The error bars are ±1 standard deviation

tory-supplied hydrophone sensitivities (Table 1) and a gain of 22 dB. The sub-segments were processed independently in order to develop an estimate of the *in situ* ambient NL and a measure of the variability of the estimate. These estimates were then considered appropriate characterizations for the entire season. SLs of sperm whale clicks reported by Zimmer et al. (2005) were assumed to be derived from root mean square (RMS, i.e. broadband) measurements. The RL of RMS measurements can be computed by integrating the received signal spectral energy over the signal frequency band. Therefore, formally, the detection statistic should compare the received RMS energy to the noise spectral energy integrated over the same frequency band, which measures the degree of certainty with which the signal is detected. As a proxy for this calculation, a detection statistic was computed at the Nyquist frequency. Since ambient noise spectra were 'spectral densities' (Fig. 3), the SNR calculation used the levels at exactly these frequencies. The associated signal power was the SL density (Eq. 1). Near-ship contamination significantly changed the apparent NL error, for example in the case of AAR3 in winter.

2.4. Environmental data

We used sea surface temperature anomaly (SSTA), sea surface height anomaly (SSHA), chlorophyll *a*

(chl *a*) and wind speed obtained from satellite data depositories to give an indication of the thermal, circulation and nutrient enrichment conditions around our AAR mooring positions. These variables were used as proxies for oceanographic features that are related to ocean productivity characterizing the Benguela ecosystem (Shannon 2009). Globally gridded products averaged over a daily temporal resolution were selected for all environmental variables; a summary description of each environmental variable used is given in Table 2.

Given the seasonal differences in sperm whale click modelled detection ranges between the AARs (Fig. 4), we treated each AAR position as an independent sampling point. To describe the variation in each environmental variable to which whales were likely exposed, we averaged environmental variables within the average seasonal detection ranges (Fig. 4) for each season from the deepest possible vocalizing depth, as sperm whales are known to produce most of their clicks at greater depths (e.g. Watwood et al. 2006). The seasonal radius of the average detection range around each AAR adjacent to the latitudinal and longitudinal grids of each AAR mooring position gave the spatial scale domain of environmental variable integration per season. For example, we averaged by 15 km grid (detection range at 800 m) for the AAR1 mooring in winter, where the values for four 15 km blocks adjacent to that location were averaged to obtain comparable daily environmental conditions within the AAR average detection range (Shabangu et al. 2019).

Less than 30% of daily chl *a* and wind speed data were missing for some seasons, and these gaps were filled by interpolation of available data from the day before or after the date with the missing environmental data, as there is a strong temporal and spatial autocorrelation in chl *a* (e.g. Kahru et al. 2012) and wind speed (e.g. Monahan 2012) data. Chl *a* was log transformed before any statistical analyses because of data skewness, and will henceforth be abbreviated as lchl *a*. Since environmental data from AAR1 were collected within the second half of 2014, smoothed mean monthly values of environmental variables for AAR1 plots were calculated by the locally weighted polynomial regression (i.e. non-circular smoothing)

Table 2. Summary of daily environmental variables derived from global environmental data repositories. The column 'Usage' indicates the reasons why a particular environmental variable is used in this study. ACRI-ST: Sciences de la Terre (Earth Sciences – Environment); CMEMS: Copernicus Marine Environment Monitoring Service; DUACS: Data Unification and Altimeter Combination System; GHRSST: Group for High Resolution Sea Surface Temperature; OSTIA: Operational Sea Surface Temperature and Sea Ice Analysis; NOAA: National Oceanic and Atmospheric Administration; SSWS: Sea Surface Wind Speed; SSHA: sea surface height anomaly; SSTA: sea surface temperature anomaly

Variable	Abbreviation	Unit	Group: Product	Data repository link	Spatial resolution	Usage	Reference
Chlorophyll a	chl a	mg m^{-3}	ACRI-ST: GlobColour	ftp://ftp.hermes.acri.fr/GLOB/merged/day/	1/24° (~4.63 km)	Phytoplankton pigment concentration	Maritorena et al. (2010)
Sea surface height anomaly	SSHA	m	CMEMS:DUACS	ftp://my.cmems-du.eu/Core/SEALEVEL_GLO_PHY_CLIMATE_L4_REP_OBSERVATIONS_008_057/dataset-duacs-rep-global-merged-twosat-phy-14	0.25° × 0.25° (~28 × 28 km)	Indication of ocean circulation	Taburet et al. (2019)
Sea surface temperature anomaly	SSTA	°C	GHRSST: OSTIA	ftp://nrt.cmems-du.eu/Core/SST_GLO_SST_L4_NRT_OBSERVATIONS_010_001/METOFFICE-GLO-SST-L4-NRT-OBS-ANOM-V2	0.05° (~3.84 km)	Indication of thermal structure of the upper ocean	Donlon et al. (2012)
Wind speed	–	m s^{-1}	NOAA:SSWS	ftp://eclipse.ncdc.noaa.gov/pub/seawinds/SI/uv/daily/netcdf/2000s	0.25° × 0.25° (~28 × 28 km)	Proxy of sea state conditions	Zhang et al. (2006)

using the function 'loess' (Cleveland et al. 1992) in R. Given that AARs 2, 3 and 4 contained year-round data, smoothed mean monthly values of environmental variables for plots were calculated through penalized cyclic cubic regression splines in GAMs.

2.5. Statistical data analyses

The relative effect and importance of 6 predictor variables (daylight regimes, month of the year, lchl a, SSHA, SSTA and wind speed) on the acoustic occurrence of sperm whales from the 4 AARs were investigated using 4 random forest (RF) models (Ho 1995, Breiman 2001). The RF model is an ensemble modelling approach applicable to a wide variety of issues such as classification, regression, time series and survival data with non-parametric inferential properties (Breiman 2001, Hastie et al. 2009, Kane et al. 2014). As a machine learning method, the RF modelling approach provides higher statistical performance (i.e. high prediction accuracy and low prediction error) and has considerable benefits over commonly used regression methods such as the generalized boosted regression trees model (Friedman et al. 2000, Shabangu et al. 2017) and GAMs (Elith et al. 2008, James et al. 2013, Shabangu et al. 2019), owing to its non-parametric inferential properties. The RF modelling approach uses a set of unpruned or unbootstrapped decision trees in the forest that are bootstrapped as they grow with sample training data, and rely on randomly chosen subsets of predictor variables as candidate splitting tree nodes (Breiman 2001, Hastie et al. 2009, James et al. 2013). It is particularly useful and preferred for this kind of data and specific analysis given its observed higher predictive capabilities for modelling acoustic occurrence of other marine mammals (Shabangu et al. 2017, 2019, 2020a,b). The relative importance of each predictor variable in the model was computed by permuting the out-of-the bag (OOB) data, where the prediction error is recorded for each tree as detailed in Shabangu et al. (2017, 2019). The OOB also represents the goodness-of-fit for each RF model.

Prior to RF modelling, the effects of multi-collinearity between predictor variables (daylight regimes, month of the year, lchl a, SSHA, SSTA and wind speed) were determined using generalized variance inflation factors (GVIFs; Fox & Monette 1992) implemented through the 'car' package (Fox & Weisberg 2019). Weak to moderate multi-collinearities were found between predictor variables for AAR1 (highest GVIF was 3.18), AAR2 (highest GVIF was 2.83) and
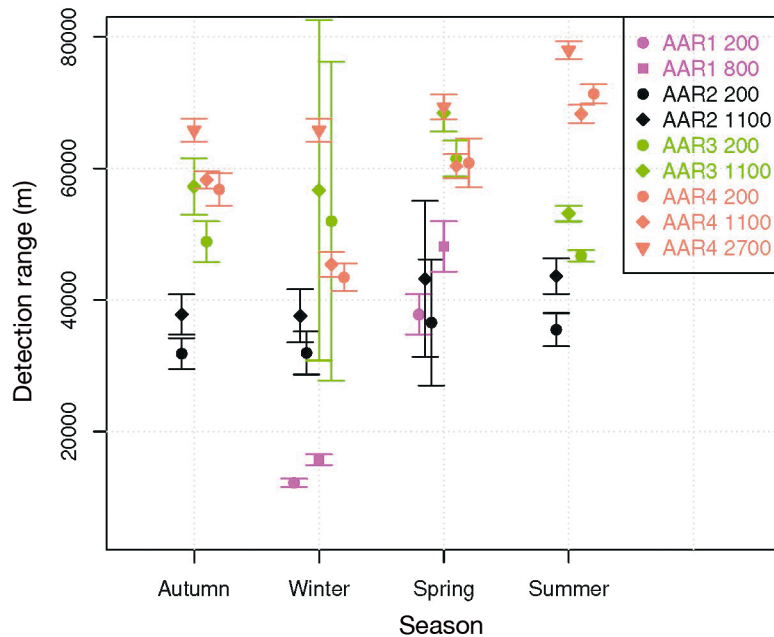
Fig. 4. Average detection ranges of sperm whale clicks at various vocalizing depths: autonomous acoustic recorder (AAR) 1 at 200 and 800 m, AAR2 at 200 and 1100 m, AAR3 at 200 and 1100 m, and AAR4 at 200, 1100 and 2700 m at 2000 (AARs 1 and 2) and 4000 (AARs 3 and 4) Hz. Error bars represent standard deviations

2017). Optimal parameter configurations for each RF model corresponding to each of the 4 AARs are given in Table 3. Using the above derived optimal parameter configurations, RF modelling was performed in R using the 'randomForest' package (Liaw & Wiener 2002). To expand interpretability of RF model outputs, we computed p-values for the feature importance metric through permutation using a technique described by Altmann et al. (2010) to measure the significance of each predictor variable.

3. RESULTS

3.1. Click detection

All 4 types of sperm whale clicks were detected from all 4 AARs over the 31 months of deployment, indicating that all AARs were suitable for recording sperm whale clicks. 'Usual' clicks were the most commonly detected. A total of 13 135 h of acoustic data were recorded from all AARs deployed off the west coast of South Africa, and the total number of hours of data recorded by each AAR per season is given in Table 4. In terms of the total number of hours containing sperm whale clicks, AAR3 produced the most and AAR4 produced the fewest (Table 4). The highest seasonal percentage of hours of data with detected sperm whale clicks was 61 % from AAR3 in spring, followed by 43 % from AAR2 in summer and spring (Table 4). The highest number of days with detected sperm whale clicks was 287 (the majority of those days were in summer and spring) at AAR3, followed by 283 (134 in spring) at AAR2 (Table 4). Spring at AAR3 had the highest seasonal percentage (94 %) of days with clicks, followed by AAR2 with 80 % in spring. Although AAR4 had the highest number of recorded hours of data, it

AAR3 (highest GVIF was 3.04). High multi-collinearity was found between month of the year and other predictor variables for AAR4 as the estimated GVIF value was 12.41 (indicating high monthly changes of variables), but the highest GVIF value dropped to 1.06 after month of the year was excluded as a predictor variable.

The area under the receiver operating characteristic curve (AUC) was used to estimate optimal parameter configurations for each RF model corresponding to each of the 4 AARs used to investigate the effect and importance of predictors on sperm whale acoustic occurrence. The AUC measured the predictive accuracy of a range of RF models with different combinations of the number of growing trees (ntrees; range: 500–3000 by increment of 500), the splitting minimum size of terminal nodes of trees (nodesize; range: 1–5) and the number of acoustic occurrences randomly selected at the tree node (mtry; range: 1–5). The AUC determined how well each model correctly classified the acoustic occurrence, taking values between 0.5 and 1 where values closer to 1 indicate excellent classification ability (DeLong et al. 1988). Values for optimal parameter configuration of RF models were determined using the 'ranger' package as a computational-time-saving method for the implementation of the RF models (Wright & Ziegler

Table 3. Optimal settings used for RF models of each autonomous acoustic recorder (AAR)

AAR ID	mtry	ntree	nodesize
AAR1	5	2000	1
AAR2	1	1000	1
AAR3	5	2500	1
AAR4	5	500	1

Table 4. Seasonal number and percentage of hours of acoustic recordings containing sperm whale clicks for each autonomous acoustic recorder (AAR). Sampling protocols were summed to constitute hours; – represents instances when no acoustic data were collected

AAR ID	Season	No. of hours recorded	No. of hours with clicks	% of hours with clicks	No. of days recorded	No. of days with clicks	% of days with clicks
AAR1	Summer	10	0	0	1	0	0
	Autumn	–	–	–	–	–	–
	Winter	465	91	19.57	39	18	46.15
	Spring	1092	242.50	22.21	91	52	57.14
AAR2	Summer	715	308.88	43.20	91	69	75.82
	Autumn	729	122.76	16.84	92	34	36.96
	Winter	729	180.18	24.72	92	46	50
	Spring	1317	562.32	42.69	167	134	80.24
AAR3	Summer	1211	476.28	39.33	120	85	70.83
	Autumn	927	250.74	27.05	92	62	67.39
	Winter	927	258.30	27.86	92	55	59.78
	Spring	917	562.80	61.37	91	85	94.41
AAR4	Summer	1324	100.80	7.61	132	38	28.79
	Autumn	927	12.60	1.36	92	5	5.44
	Winter	927	11.34	1.22	92	5	5.44
	Spring	917	96.18	10.49	91	34	37.36

produced the lowest percentage and number of hours and days with detected sperm whale clicks (Table 4).

3.2. Click detection ranges

Overall, there was a strong negative correlation (Pearson's r = −0.99) between ambient NL and modelled detection ranges. AAR2 generally had consistent and stronger noise; the modelled detection range was therefore consistently shorter (Figs. 3 & 4). Near-ships created large swings in ambient NL, specifically in winter, which was then reflected in large uncertainties in the modelled detection range for AAR3 in winter (Fig. 4). Modelled detection ranges changed seasonally for each AAR in relation to NLs, where the highest detection ranges for AARs 1 and 3 were in spring and those for AARs 2 and 4 were in summer (Fig. 4). AAR1 had the shortest average detection range of 15 km for whales vocalizing at 800 m depth in winter (Fig. 4), which was characterized by a high NL (Fig. 3). There was a slight, and sometimes significant, increase in detection range for deeper sources, mainly for AARs 3 and 4 in spring (Fig. 4). The average modelled detection range was greatest (78 km) for a source at 2700 m calling from AAR4 in summer (Fig. 4), although it should be noted that ambient NLs on AAR4 were generally the lowest (Fig. 3).

3.3. Environmental conditions around AARs

Environmental conditions varied monthly around AAR positions (Fig. 5). There was a higher overall lchl *a* concentration around AAR1 than around other AARs (Fig. 5a,e,i,m). A higher SSHA was observed in September around AAR1 (Fig. 5b), in February around AARs 2 and 3 (Fig. 5f,j), and in April around AAR4 (Fig. 5n). Negative (cold) SSTA was observed in December around AAR1 (Fig. 5c), and from March through September around AAR2 (Fig. 5g). SSTA was positive (warm) for most months around AAR3 but dropped to 0°C between October and November (Fig. 5k), and a negative SSTA was observed between April and June, and between October and November around AAR4 (Fig. 5o). Monthly wind speeds were comparable for AARs 1 and 2 between July and December (Fig. 5d,h), and changed similarly between corresponding months for AARs 3 and 4 (Fig. 5l,p).

3.4. Observed temporal variability

Sperm whale clicks were detected year round with a considerably higher percentage of acoustic occurrence in summer and late winter through spring (Fig. 6). August through October had the highest percentage of acoustic occurrence for AAR1 (Fig. 6). Clicks were acoustically absent from AAR1 in July and December 2014 as the instrument was deployed
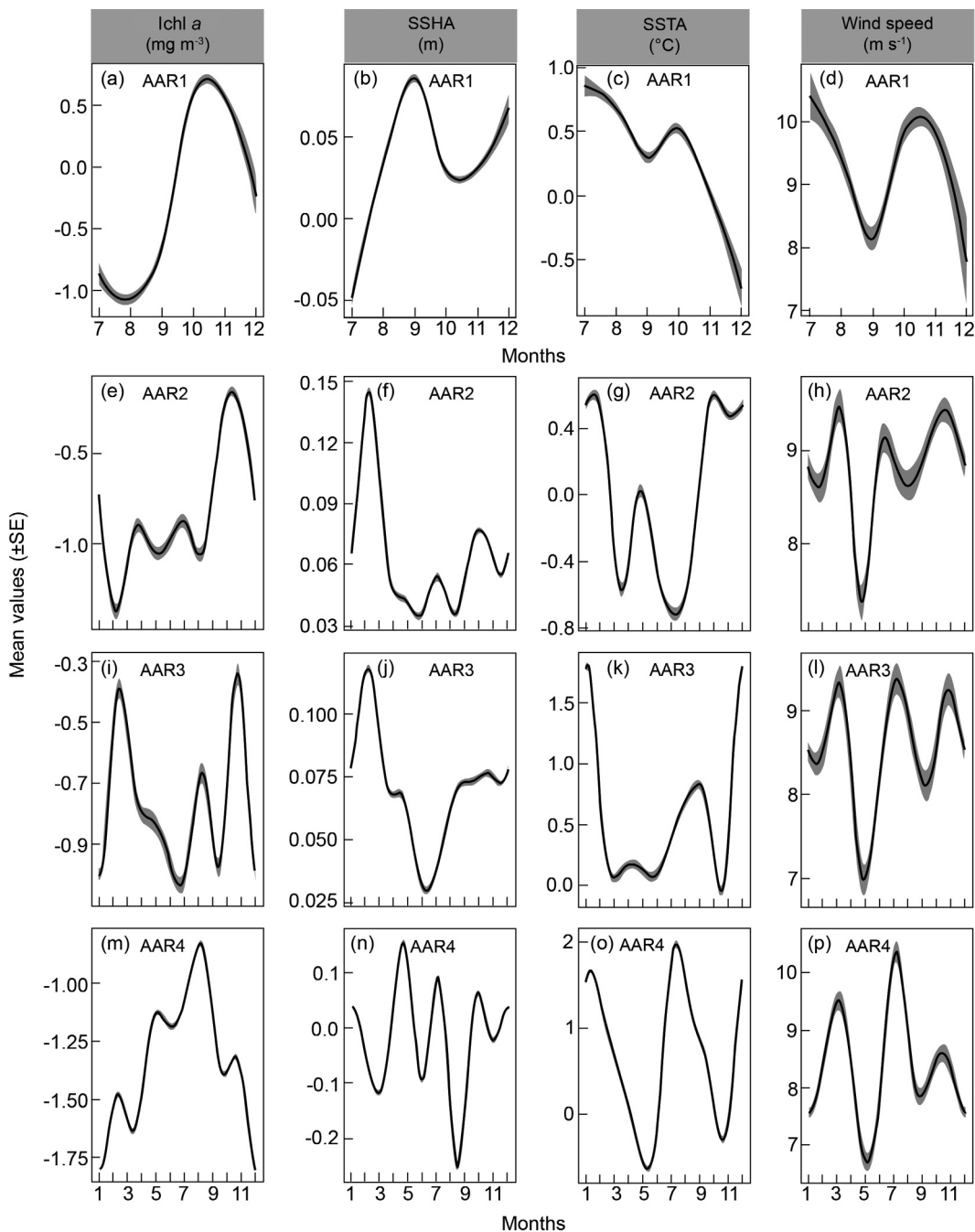
Fig. 5. Locally weighted polynomial regression smoothed monthly mean values of different environmental variables around autonomous acoustic recorder (AAR) 1 (a–d), and circular smoothed monthly mean values of different environmental variables around AAR2 (e–h), AAR3 (i–l) and AAR4 (m–p). See Table 2 for definitions of variable abbreviations. SE: standard error (grey shading); lchl *a*: log-transformed chl *a*. Note the different scales for the *x*- and *y*-axes

towards the end and beginning of those months respectively. During the period of deployment overlap between AARs 1 and 2 in 2014, the average percentage of acoustic occurrence was 17% for AAR1 but 39% for AAR2. Two peaks in sperm whale acoustic occurrence were observed for AARs 2 and 3 in sum-

mer (December through February), and at the end of winter through spring (August through November) (Fig. 6). November and December produced the highest percentage of acoustic occurrence for AAR4 (Fig. 6). AAR2 had the highest percentage of acoustic occurrence in 2015, whilst AAR3 had the highest in 2016
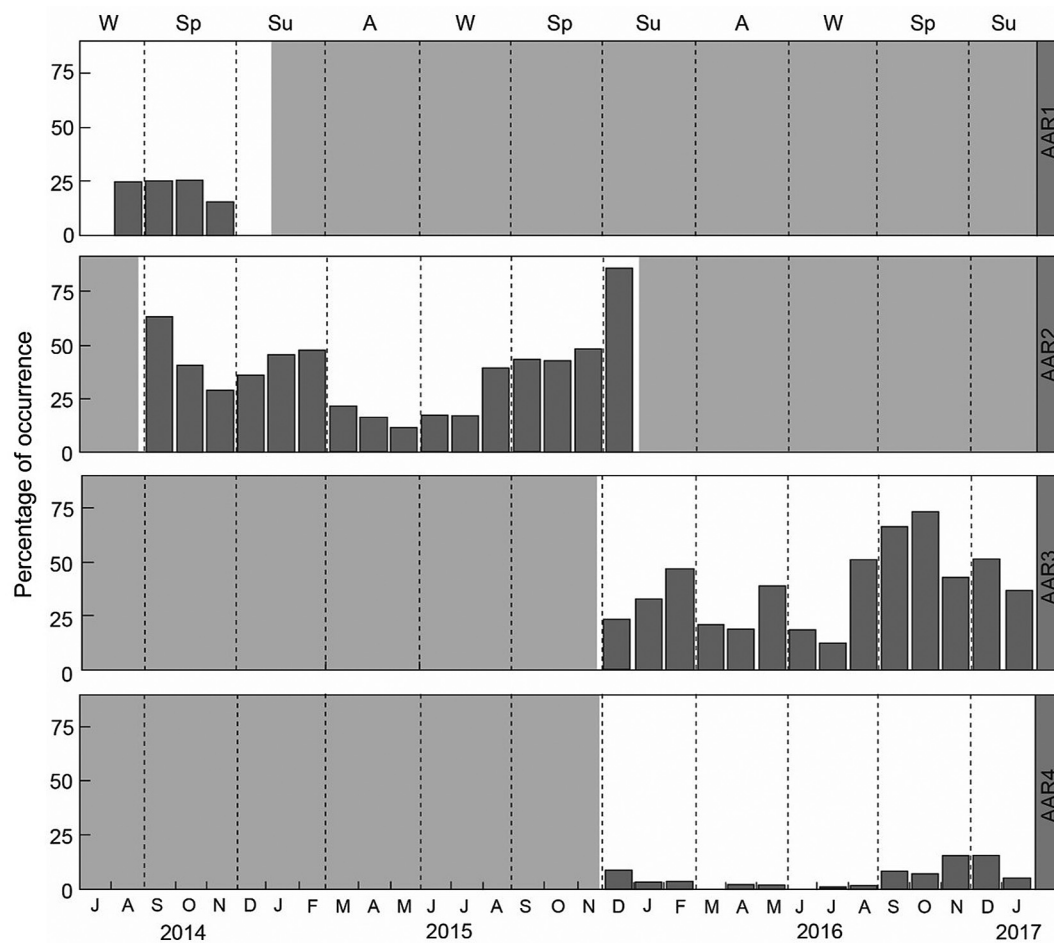
Fig. 6. Monthly percentages of sperm whale acoustic occurrence (dark grey bars) from the 4 autonomous acoustic recorders (AARs) off the west coast of South Africa. Light grey shaded areas indicate periods without passive acoustic monitoring effort. Seasons (Su: summer; A: Autumn; W: winter; Sp: spring) are shown on the top axis and outlined by dashed lines, and months and years are indicated on the bottom axis

(Fig. 6). AAR4 had the lowest percentage of acoustic occurrence of all AARs (Fig. 6). There was no inter-annual variability in the acoustic occurrence of sperm whales between AARs 2 and 3 (p-value > 0.05; Fig. 6), indicating that whale acoustic occurrence did not change over time at this water depth. Nonetheless, AAR3 had a higher average percentage of acoustic occurrence of 40% over all months in 2016, whereas AAR2 had an average of 37% over all months in 2015. For AAR4, the average over all months in 2016 was 5%, while AAR1 had an average of 15% over all months.

Seasonal diel percentages of sperm whale acoustic occurrence were observed to be slightly higher during the daytime for AARs 2 and 3 in summer, and for AARs 1 to 3 in spring, but there was no clear diel pattern in autumn for any AARs (Fig. 7). In winter, sperm whale acoustic occurrence was observed to increase from around 07:00 h to dusk for AAR1, from 07:00 to 20:00 h for AAR 2, and from 07:00 to 17:00 h for AAR3

(Fig. 7). Diel percentage of acoustic occurrence did not change with time of day in any season for AAR4, although there was a slight increase in acoustic oc-currence during the day for most seasons (Fig. 7).

3.5. Predictors of sperm whale acoustic occurrence

SSHA above 0.05 m, dawn, positive (warm) SSTA above 1.2°C, wind speeds below 9 m s^{-1}, lchl a values between 0.3 and 0.7 mg m^{-3}, and month of the year (September and October) had the highest effect on sperm whale acoustic occurrence for AAR1 (Fig. 8a–f). Month of the year (January, February, August and September), low lchl a (around −1.8 mg m^{-3}), dawn and dusk, wind speeds below 7 m s^{-1}, SSTA above 1°C and SSHA above 0.14 m had the highest effect for AAR2 (Fig. 8g–l). September and October, dawn and dusk, lchl a around 0 and 0.5 mg m^{-3}, SSHA
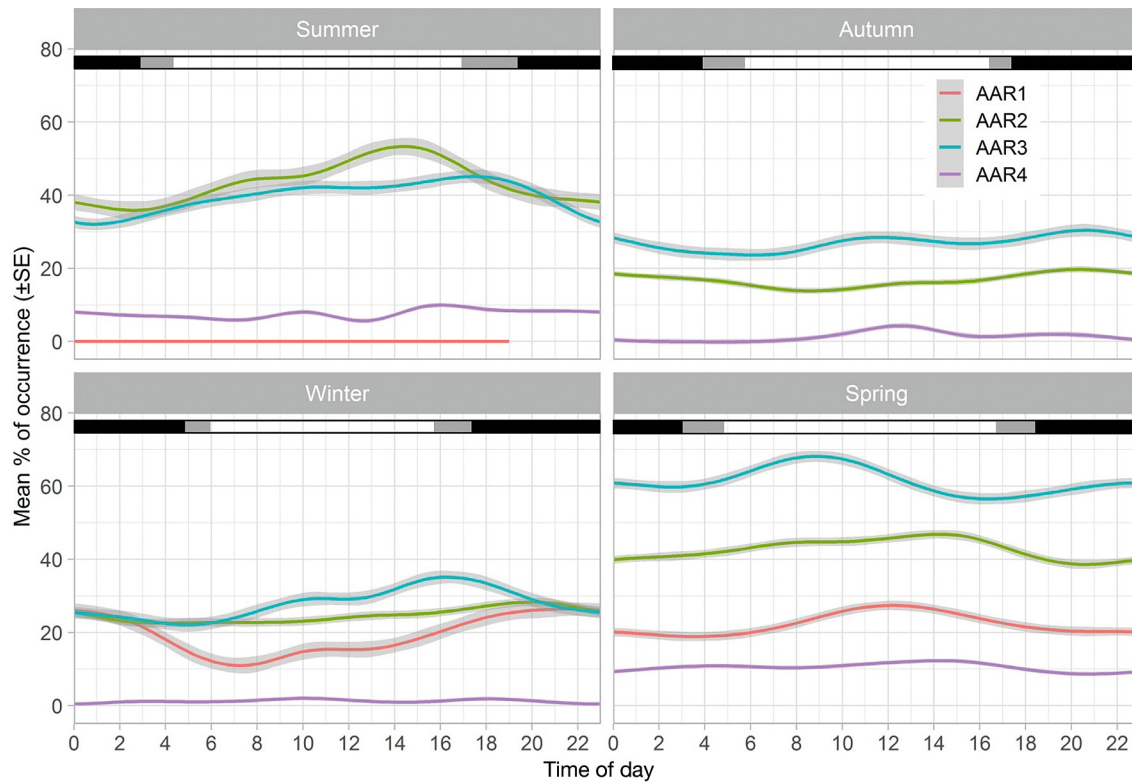
Fig. 7. Circular smoothed diel percentage of acoustic occurrence of sperm whales according to season off the west coast of South Africa. At autonomous acoustic recorder (AAR) 1, no acoustic data were recorded in autumn and not all hours of the day were recorded in summer, thus a flat line. Horizontal diel bar shading: (black) average nighttime hours; (grey) average twilight hours; (white) average daytime hours. Grey shading indicates SE. Coordinated Universal Time was used

above 0.14 m, wind speeds below 6 m s^{-1} and relatively warm SSTA around 1.2°C had the highest effect for AAR3 (Fig. 8m–r). Dawn and dusk, low lchl *a* levels (below −1 mg m^{-3}), wind speed around 5 m s^{-1}, SSHA around 0 m and negative SSTA around −1°C had the highest effect for AAR4 (Fig. 8s–w).

For AAR1 sperm whale acoustic occurrence, SSHA and daylight regime were the most important predictors, SSTA and wind speed were moderately important, and lchl *a* and month of the year were the least important (Fig. 8x). For AAR2, month of the year was the most important predictor, lchl *a*, daylight regime and wind speed were moderately important, and SSTA and SSHA were the least important (Fig. 8y). For AAR3 sperm whale acoustic occurrence, month of the year was the most important predictor, daylight and lchl *a* were moderately important, and SSHA, wind speed and SSTA were the least important (Fig. 8z). Daylight regime was the most important predictor for AAR4, lchl *a* and wind speed were moderately important, and SSHA and SSTA were the least important (Fig. 8aa). All predictor variables were significantly important (Fig. 8x–aa), indicating that they are informative.

4. DISCUSSION

The year-round acoustic occurrence of sperm whales at most acoustic stations off the west coast of South Africa suggests that the high biological productivity (linked to environmental conditions such as lchl *a*, SSTA, SSHA and wind speed) associated with the Benguela ecosystem upwelling regime can support high biomasses of prey to sustain the food requirements of these animals throughout the year. Best (1969) hypothesized that the lack of seasonal trend in sperm whale movement could be due to favourable oceanographic conditions on the west coast. Low acoustic occurrence of sperm whales in autumn could have potentially coincided with the seasonal rapid decline in biological productivity (characterized by low lchl *a*, SSHA, SSTA and wind speed) off the South African west coast at this time of the year (Brown 1992, Shabangu et al. 2019, this study), which could have resulted in low prey biomasses. According to the RF models, these different environmental conditions differently but significantly influence the acoustic occurrence of whales at different AAR sites, indicating that environmental
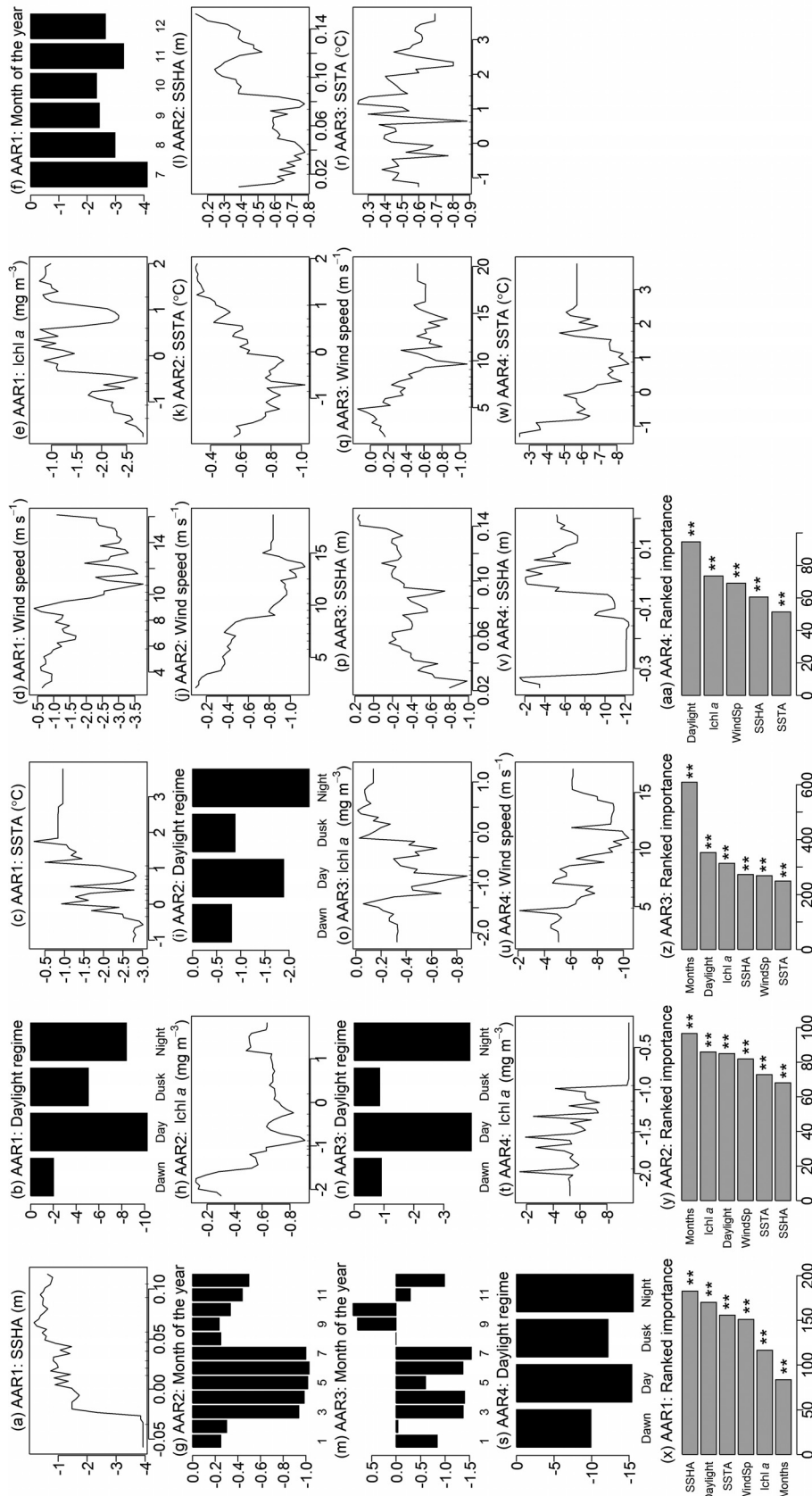
Fig. 8. Relative effects and ranked relative importance of predictor variables on sperm whale acoustic occurrence from autonomous acoustic recorder (AAR) 1 (a–f,x), AAR2 (g–l,y), AAR3 (m–r,z) and AAR4 (s–x,aa) estimated by RF models. See Table 2 for definitions of variable abbreviations. WindSp: wind speed; lchl a: log-transformed chl a. (a–w) y-axes are the partial effect of each predictor on the acoustic occurrence in logit scale; (x–aa) x-axes are the ranked relative importance of each predictor variable on call occurrence. **predictor variable with significant (p < 0.05) importance

variables significantly drive the seasonal occurrence of whales.

Variations in the acoustic occurrence of sperm whales between AARs might indicate variation in the spatio-temporal occurrence of sperm whales off the west coast of South Africa, where seasonal acoustic occurrence varied with the water depth at AAR deployment, month of the year, environmental conditions and seasonal changes in AAR detection ranges. For example, AARs 3 and 4 had the same sampling protocol and were deployed within the same period but at different water depths (1118 m for AAR3 versus 4481 m for AAR4); however, AAR3 produced more hours with sperm whale clicks than AAR4. Purdon et al. (2020b) found water depths between 750 and 1500 m and a distance to shore of approximately 130 km to have the highest effect on sperm whale distribution in the southern African region. Moreover, AAR3 had a higher detection range than AAR4 due to variations in ambient noise between the 2 positions. The same interpretation is not applicable to sperm whale click detection at AARs 1 and 2, since the 2 AARs were closely spaced on the shelf edge with a few hundred meters difference in water depth.

The high acoustic occurrence of sperm whales obtained from AARs 2 and 3 deployed at 1118 m water depth suggests that this water depth is one of the preferred and more important habitats, as indicated by whale catches where more whales were caught inhabiting water depths of 1000 m than 4000 m (Best 1999). AAR4 deployed at 4481 m water depth produced the lowest percentage of sperm whale acoustic occurrence, indicating that a small proportion of this species occupies this deep sea area, as similarly indicated by whale catches (Best 1999) and species distribution modelling (Purdon et al. 2020b). Furthermore, sperm whale clicks produced at great depths could have been deflected by the thermocline (an acoustic barrier) and prevented from reaching AARs 1 and 4, as these 2 recorders were sometimes positioned above or below the thermocline depth (~200 m depending on season; Shabangu et al. 2020c). On the other hand, AARs 2 and 3 were well positioned to detect sperm clicks as they were well below the thermocline depth throughout the year.

Although AARs 1 and 2 recorded at half the sampling rate of AARs 3 and 4, they produced higher sperm whale acoustic occurrence. This result indicates that sperm whales can be studied successfully using low sampling rates (such as 4096 Hz), which are implemented to study the low frequency sounds of baleen whales (Shabangu et al. 2019, 2020a,c) and also to preserve the battery life of AARs. However, if possible, the use of higher sampling rates is recommended in order to capture the whole frequency range and the directional component of sperm whale clicks. AAR2 had the shortest sampling protocol (single block of 20 min h^{-1}) of all the AARs (Table 1) but it produced the second highest average percentage of acoustic occurrence of sperm whales, even when it was recording concurrently with AAR1 (Fig. 5), which had the longest sampling protocol (single block of 30 min h^{-1}). This result suggests that the sampling protocol of an AAR did not appreciatively reduce the probability of sperm whale click detection, as equally observed by Stanistreet et al. (2018). Even though AARs 1 and 2 were closely positioned (4.8 km apart) and were recording concurrently for 3 months, they differed during the time of overlap (Fig. 5) in that AAR2 recorded 22% more acoustic occurrence than AAR1. This result indicates that deeper deployment (AAR2 was deployed in deeper waters than AAR1) allowed the detection of more clicks, since sperm whales produce more clicks in deeper waters (Watwood et al. 2006). Although AARs 3 and 4 had the same sampling protocols, AAR3 detected more clicks than AAR4, further endorsing a 1118 m water depth as crucial for this species. It also likely that a 300 m (AAR3) deployment depth is better than 200 m (AAR4), allowing more clicks to be recorded, since this is below the thermocline depth (Shabangu et al. 2020c).

The absence of sperm whale clicks from the AAR deployed off the Maud Rise (Shabangu & Charif 2020, Shabangu et al. 2020a,b) was somewhat unexpected as sperm whales have been sighted (Findlay et al. 2014) and acoustically recorded south of 65° to near the sea ice edge in the eastern Weddell Sea (Shabangu et al. in press). This acoustic absence of sperm whales off the Maud Rise might indicate that this area is not an important habitat for this species as it is covered by sea ice for most of the year from May through September (Shabangu & Charif 2020, Shabangu et al. 2020a,b). Sperm whales are known to avoid sea ice-infested waters (Jefferson et al. 1993, 2015). Our RF model results indicated that sperm whale acoustic occurrence was high for recorders (AARs 1–3) in environments characterized by warm SSTA compared to a recorder (AAR4) in an environment characterized by cooler SSTA (Fig. 8). Correspondingly, Purdon et al. (2020b) found sperm whale distribution to be influenced by warm SST. Results of Purdon et al. (2020) and the present study indicate that this species might be vulnerable to climate warming, not only in the high latitudes, but also in low latitude habitats such as the west coast of South Africa.

Given the observed seasonal patterns for the diel percentage of sperm whale acoustic occurrence for all AARs, the RF models classified daylight regime (dawn and dusk) either as the most or as a moderately important variable for all AARs, likely indicating elevated whale foraging activities during these transition periods in relation to diurnal migration patterns of prey. For some AARs, there was no variation in diel percentage of sperm whale acoustic occurrence for certain seasons, which could indicate that this whale species was foraging at all times of the day during those seasons (e.g. Best 1999, Davis et al. 2007). Variations in diel-vocalizing patterns of sperm whales per AAR could be due to the localized availability of prey and likely changes in the size of whales detected. Additionally, such diel and seasonal variations could be due to seasonal changes in NLs (Fig. 2) and oceanographic conditions affecting the position of the thermocline (Shabangu et al. 2020c). Similarly, Merkens et al. (2019) observed no single diel pattern across locations in the Pacific Ocean.

The BELLHOP model results indicated that AAR1 had a shorter average detection range of 15 km in winter, whilst AAR4 had the longest average modelled detection range of 78 km in summer. The maximum modelled detection range of 83 km for sperm whale clicks at the 1100 m animal vocalizing depth estimated for AAR3 in winter is more than twice that of the previous maximum detection range of 35 km reported by Mathias et al. (2013). Such big differences in the detection ranges of sperm whale clicks between our study and previous studies (e.g. Mathias et al. 2013, André et al. 2017) could be due to varying TL, sea state conditions (affecting NLs), SLs, recorder types, recorder depths, sound propagation models used and bathymetric properties of different regions. For example, Shabangu et al. (2020b) attributed the difference in detection ranges between their study and another study that used the same detection range estimation approach to the depth of the hydrophone in the water column, which resulted in different sound attenuation.

The seasonal variability in peaks of sperm whale acoustic occurrence coincided with changes in environmental variables, where summer and spring were characterized by high SSHA and wind speed but low SSTA. The RF model results showed that August and September had the most effect on sperm whale acoustic occurrence for AAR1; these months coincide with the increased primary productivity in the Benguela ecosystem in late winter and early spring (Andrews & Hutchings 1980, Brown 1992, Hagen et al. 2001). September and October were indicated by

RF models to have more effect on sperm whale acoustic occurrence for AARs 2 and 3, likely due to the high primary productivity (high SSHA and wind speed but low SSTA) associated with spring in the Benguela ecosystem (Andrews & Hutchings 1980, Brown 1992, Hagen et al. 2001, Shabangu et al. 2019). Although month of the year was eliminated as a predictor from the RF model for AAR4 due to multicollinearity, September through December had the highest percentage of sperm whale acoustic occurrence. These months coincide with the known increased primary productivity in the Benguela ecosystem in spring and summer (Andrews & Hutchings 1980, Brown 1992, Hagen et al. 2001).

The increased sperm whale acoustic occurrence in summer and late winter through spring is comparable to the high sperm whale catches made in autumn and late winter through spring during the whaling era in this region off the west coast of South Africa (Best 1969). Sperm whale seasonal occurrence is likely to have changed slightly over time, as the first occurrence peak in summer is slightly earlier compared to that in autumn during the whaling era. Additionally, these observed variations in the seasonal acoustic occurrence of sperm whales between deployment sites could indicate the seasonal distribution change (i.e. migration) of males and females, as shown by whale catches off the west coast of South Africa (Best 1969, 1974, 2007). This pattern of 2 peaks in acoustic occurrence is different from those of Antarctic blue and fin whales (Shabangu et al. 2019), but comparable to that of Antarctic minke whales that likely dive to 300 m given the strong harmonics extending to 2 kHz detected by AAR2 positioned at 300 m (Shabangu et al. 2020a). Furthermore, the 2 peaks observed in our study might indicate northward migration in summer and southward migration in late winter through spring, as shown by seasonal whale abundance patterns off the Donkergat whaling station that depend on whale size and sex (Best 1969). These migration directions are different from those of baleen whales (e.g. Shabangu et al. 2019, 2020a,c), likely because the migration of these whales to the high latitudes is not dictated by sea ice conditions as they do not prey on the sea ice-dependent Antarctic krill *Euphausia superba*.

The RF models enabled us to explicitly interpret the relationship between the environmental conditions of the Benguela ecosystem and sperm whale acoustic occurrence. Such effects might indicate that sperm whales are dependent on the environmental conditions within the Benguela ecosystem. Changes in

these environmental conditions might lead to localized changes in seasonal abundances and distribution of sperm whale prey, which might in turn lead to changes in sperm whale occurrence as observed, for example, off the Galapagos Islands, Ecuador (Whitehead 1996) and in the Gulf of Alaska (Diogou et al. 2019).

5. CONCLUSIONS

The results of this research corroborate the effectiveness and benefits of using passive acoustic monitoring to provide information on the recent acoustic occurrence and behaviour of seldom-sighted marine mammals such as sperm whales. The observed year-round acoustic presence of sperm whales in South African waters indicates residency to a certain extent, and such occurrence is likely due to the high primary productivity in the west coast bioregion associated with the upwelling Benguela Current. The RF model results show that environmental conditions (lchl a, SSHA, SSTA, wind speed, daylight regime and month of the year) significantly influenced the seasonal acoustic occurrence of sperm whales. This work indicates that sperm whales were acoustically present in high numbers in summer and late winter through spring, whereas diel-vocalizing patterns were detected in winter, spring and summer, updating and extending existing knowledge of seasonal occurrence and behaviour obtained from whaling statistics. These whales showed some preference for water depth around 1118 m, probably in relation to their prey distribution on the shelf edge and favourable environmental conditions characterizing this important habitat. Seasonal detection ranges varied between AARs, due to varying ambient NLs at different AAR deployment locations. Our highest maximum modelled detection range of 83 km is considerably farther than the 35 km previously reported in the literature. This is the first study to illustrate the seasonal acoustic occurrence and diel-vocalizing patterns of sperm whales in relation to environmental conditions off the west coast of South Africa based on empirical data, making it useful for the conservation and protection of the species in this region. The identified important habitats are fundamental for the conservation and protection of sperm whales from potential threats off the west coast of South Africa.

LITERATURE CITED

Altmann A, Toloşi L, Sander O, Lengauer T (2010) Permutation importance: a corrected feature importance measure. Bioinformatics 26:1340–1347

André M, Caballé A, van der Schaar M, Solsona A and others (2017) Sperm whale long-range echolocation sounds revealed by ANTARES, a deep-sea neutrino telescope. Sci Rep 7:45517

Andrews WRH, Hutchings L (1980) Upwelling in the southern Benguela Current. Prog Oceanogr 9:1–8

Antunes R, Schulz T, Gero S, Whitehead H, Gordon J, Rendell L (2011) Individually distinctive acoustic features in sperm whale codas. Anim Behav 81:723–730

Backus RH, Schevill WE (1966) Physeter clicks. In: Norris KS (ed) Whales, dolphins, and porpoises. University of California Press, Berkekey, p 510–528

Best PB (1969) The sperm whale (*Physeter catodon*) off the west coast of South Africa. 4. Distribution and movements. Investl Rep Div Sea Fish S Afr 78:1–12

Best PB (1974) Status of the whale populations off the west coast of South Africa, and current research. In: Schevill WE (ed) The whale problem: a status report. Harvard University Press, Cambridge, p 53–81

Best PB (1999) Food and feeding of sperm whales *Physeter macrocephalus* off the west coast of South Africa. S Afr J Mar Sci 21:393–413

Best PB (2007) Whales and dolphins of the Southern African subregion. Cambridge University Press, Cambridge

Bioacoustics Research Program (2017) Raven Pro: interactive sound analysis software (Version 1.5). The Cornell Lab of Ornithology, Ithaca, NY. www.birds.cornell.edu/raven

Boyer TP, Antonov JI, Baranova OK, Coleman C and others (2013) World Ocean Database 2013. NOAA Atlas NESDIS 72. NOAA Printing Office, Silver Spring, MD

Branch TA, Butterworth DS (2001) Estimates of abundance south of 60°S for cetacean species sighted frequently on the 1978/79 to 1997/98 IWC/IDCR-SOWER sighting surveys. J Cetacean Res Manag 3:251–270

Breiman L (2001) Random forests. Mach Learn 45:5–32

Brown PC (1992) Spatial and seasonal variation in chlorophyll distribution in the upper 30 m of the photic zone in the southern Benguela/Agulhas ecosystem. S Afr J Mar Sci l2:515–525

Chiquet RA, Ma B, Ackleh AS, Pal N, Sidorovskaia N (2013) Demographic analysis of sperm whales using matrix population models. Ecol Model 248:71–79

Clapham P, Baker CS (2009) Modern whaling. In: Perrin WF, Würsig B, Thewissen JGM (eds) Encyclopedia of marine mammals, 2nd edn. Academic Press, San Diego, CA, p 1239–1243

Cleveland WS, Grosse E, Shyu WM (1992) Local regression models. In: Chambers JM, Hastie TJ (eds) Statistical models in S. Chapman & Hall/CRC, Boca Raton, FL

Davis RW, Jaquet N, Gendron D, Markaida U, Bazzino G, Gilly W (2007) Diving behavior of sperm whales in relation to behavior of a major prey species, the jumbo squid, in the Gulf of California, Mexico. Mar Ecol Prog Ser 333: 291–302

DeLong ER, Delong DM, Clarke-Pearson DL (1988) Comparing areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 44:837–845

Department of Environmental Affairs (2017) Marine Living Resources Act 1998 (Act No. 18 of 1998): draft policy on boat-based whale and dolphin watching. South African Government Gazette 40626:145

Diogou N, Palacios DM, Nystuen JA, Papathanassiou E, Katsanevakis S, Klinck H (2019) Sperm whale (*Physeter macrocephalus*) acoustic ecology at Ocean Station PAPA in the Gulf of Alaska—Part 2: Oceanographic drivers of interannual variability. Deep Sea Res I 150: 103044

Donlon CJ, Martin M, Stark J, Roberts-Jones J, Fiedler E, Wimmer W (2012) The operational sea surface temperature and sea ice analysis (OSTIA) system. Remote Sens Environ 116:140–158

Elith J, Leathwick JR, Hastie T (2008) A working guide to boosted regression trees. J Anim Ecol 77:802–813

Elwen S, Findlay K, Meÿer M, Oosthuizen H, Plön S (2016) A conservation assessment of *Physeter macrocephalus*. In: Child MF, Roxburgh L, Do Linh San E, Raimondo D, Davies-Mostert HT (eds) The red list of mammals of South Africa, Swaziland and Lesotho. South African National Biodiversity Institute and Endangered Wildlife Trust, South Africa

Fais A, Johnson M, Wilson M, Aguilar Soto N, Madsen PT (2016) Sperm whale predator–prey interactions involve chasing and buzzing, but no acoustic stunning. Sci Rep 6: 28562

Figueroa HK (2006) XBAT: extensible bioacoustic tool. http://xbat.org (accessed 22 June 2015)

Findlay KP, Best PB (2016) Distribution and seasonal abundance of large cetaceans in the Durban whaling grounds off KwaZulu-Natal, South Africa, 1972–1975. Afr J Mar Sci 38:249–262

Findlay K, Thornton M, Shabangu F, Venter K, Thompson I, Fabriciussen O (2014) Report of the 2013/14 South African Antarctic blue whale survey, 000° – 020° E. International Whaling Commission Paper SC/65b/SH01

Focke KC, Mitchell SK, Horton CW Sr (1982) Analysis of deep ocean sound attenuation at very low frequencies. J Acoust Soc Am 71:1438–1444

Fox J, Monette G (1992) Generalized collinearity diagnostics. J Am Stat Assoc 87:178–183

Fox J, Weisberg S (2019) An R companion to applied regression, 3rd edn. SAGE Publications, Thousand Oaks, CA

Friedman J, Hastie T, Tibshirani R (2000) Additive logistic regression: a statistical view of boosting. Ann Stat 28: 337–407

Gambell R (1967) Seasonal movements of sperm whales. Symp Zool Soc Lond 19:237–254

Gordon JCD (1987) The behaviour and ecology of sperm whales off Sri Lanka. PhD thesis, University of Cambridge

Gordon JCD (1991) Evaluation of a method for determining the length of sperm whales (*Physeter catodon*) from their vocalizations. J Zool Lond 224:301–314

Growcott A, Miller B, Sirguey P, Slooten E, Dawson S (2011) Measuring body length of male sperm whales from their clicks: the relationship between inter-pulse intervals and photogrammetrically measured lengths. J Acoust Soc Am 130:568–573

Guisan A, Edwards TC Jr, Hastie T (2002) Generalized linear and generalized additive models in studies of species distributions: setting the scene. Ecol Model 157: 89–100

Hagen E, Feistel R, Agenbag JJ, Ohde T (2001) Seasonal and interannual changes in Intense Benguela Upwelling (1982–1999). Oceanol Acta 24:557–568

Halpern BS, Frazier M, Afflerbach J, Lowndes JS and others (2019) Recent pace of change in human impact on the world's ocean. Sci Rep 9:11609

Hastie TJ, Tibshirani R, Friedman J (2009) The elements of statistical learning. Springer, New York, NY

Ho TK (1995) Random decision forests. Proc 3rd Int Conf Document Analysis and Recognition, Montreal, QC, 14–16 August 1995, p 278–282

James G, Witten D, Hastie T (2013) An introduction to statistical learning: with applications in R. Springer, New York, NY

Jaquet N, Dawson S, Douglas L (2001) Vocal behavior of male sperm whales: Why do they click? J Acoust Soc Am 109:2254–2259

Jefferson TA, Leatherwood S, Webber MA (1993) FAO species identification guide. Marine mammals of the world. FAO, Rome

Jefferson TA, Webber MA, Pitman RL (2015) Marine mammals of the world: a comprehensive guide to their identification. Academic Press, San Diego, CA

Kahru M, DiLorenzo E, Manzanpo-Sarabia M, Mitchell BG (2012) Spatial and temporal statistics of sea surface temperature and chlorophyll fronts in the California Current. J Plankton Res 34:749–760

Kane MJ, Price N, Scotch M, Rabinowitz P (2014) Comparison of ARIMA and Random Forest time series models for prediction of avian influenza H5N1 outbreaks. BMC Bioinformatics 15:276

Lavery TJ, Roudnew B, Gill P, Seymour J and others (2010) Iron defecation by sperm whales stimulates carbon export in the Southern Ocean. Proc R Soc B 277: 3527–3531

Liaw A, Wiener M (2002) Classification and regression by randomForest. R News 2/3:18–22

Maritorena S, Fanton d'Andon OH, Mangin A, Siegel DA (2010) Merged satellite ocean color data products using a bio-optical model: characteristics, benefits and issues. Remote Sens Environ 114:1791–1804

Mathias D, Thode AM, Straley J, Andrews RD (2013) Acoustic tracking of sperm whales in the Gulf of Alaska using a two-element vertical array and tags. J Acoust Soc Am 134:2446–2461

Mellinger DK, Stafford KM, Fox CG (2004) Seasonal occurrence of sperm whale (*Physeter macrocephalus*) sounds in the Gulf of Alaska 1999–2001. Mar Mammal Sci 20: 48–62

Merkens KP, Simonis AE, Oleson EM (2019) Geographic and temporal patterns in the acoustic detection of sperm whales *Physeter macrocephalus* in the central and western North Pacific Ocean. Endang Species Res 39: 115–133

Miller GA, Heise GA, Lichten W (1951) The intelligibility of speech as a function of the context of the test materials. J Exp Psychol Gen 41:329–335

Møhl B, Wahlberg M, Madsen PT, Miller LA, Surlykke A (2000) Sperm whale clicks: directionality and source level revisited. J Acoust Soc Am 107:638−648

Møhl B, Wahlberg M, Madsen PT, Heerfordt A, Lund A (2003) The monopulsed nature of sperm whale clicks, J Acoust Soc Am 114:1143−1154

Monahan AH (2012) The temporal autocorrelation structure of sea surface winds. J Climate 25:6684−6700

Oliveira C, Wahlberg M, Silva MA, Johnson M and others (2016) Sperm whale codas may encode individuality as well as clan identity. J Acoust Soc Am 139: 2860−2869

Pavan G, Hayward TJ, Borsani JF, Priano M, Manghi M, Fossati C, Gordon J (2000) Time patterns of sperm whale codas recorded in the Mediterranean Sea 1985−1996. J Acoust Soc Am 107:3487−3495

Porter MB (2011) The BELLHOP manual and user's guide. Preliminary draft. Heat, Light and Sound Research, La Jolla, CA

Purdon J, Shabangu FW, Pienaar M, Somers MJ, Findlay K (2020a) 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

Purdon J, Shabangu FW, Yemane D, Pienaar M, Somers MJ, Findlay K (2020b) Species distribution modelling of Bryde's whales, humpback whales, southern right whales, and sperm whales in the southern African region to inform their conservation in expanding economies. PeerJ 8:e9997

R Core Team (2019) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Rendell L, Whitehead H (2004) Do sperm whales share coda vocalizations? Insights into coda usage from acoustic size measurement. Anim Behav 67:865−874

Shabangu FW, Charif RA (2020) Short moan call reveals seasonal occurrence and diel-calling pattern of crabeater seals in the Weddell Sea, Antarctica. Bioacoustics, https://doi.org/10.1080/09524622.2020.1819877

Shabangu FW, Yemane D, Stafford KM, Ensor P, Findlay KP (2017) Modelling the effects of environmental conditions on the acoustic occurrence and behaviour of Antarctic blue whales. PLOS ONE 12:e0172705

Shabangu FW, Findlay KP, Yemane D, Stafford KM, van den Berg M, Blows B, Andrew RK (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 FW, Findlay K, Stafford KM (2020a) 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 Mamm Sci 36:658−675

Shabangu FW, Andrew RK, Yemane D, Findlay KP (2020b) Acoustic seasonality, behaviour and detection ranges of Antarctic blue and fin whales under different sea ice conditions off Antarctica. Endang Species Res 43:21−37

Shabangu FW, Andrew RK, Findlay K (2020c) Acoustic occurrence, diel-vocalizing pattern and detection ranges of southern right whale gunshot sounds off South Africa's west coast. Mar Mamm Sci, https://doi.org/10.1111/mms.12760

Shabangu FW, Stafford KM, Findlay KP, Rankin S and others (in press) Overview of the IWC SOWER cruise circumpolar acoustic survey data and analyses of Antarctic blue whale calls. J Cetacean Res Manag Spec Issue

Shannon V (2006) A plan comes together. In: Shannon V, Hempel G, Malanotte-Rizzoli P, Moloney C, Woods J (eds) Benguela: predicting a large marine ecosystem. Elsevier, Amsterdam, p 3−10

Shannon V (2009) Benguela Current. In: Steele JH, Thorpe SA, Turekian KK (eds) Encyclopedia of ocean sciences, 2nd edn. Academic Press, London, p 316−327

Signorell A, Aho K, Alfons A, Anderegg N and others (2020) DescTools: tools for descriptive statistics. R package version 0.99.36. https://cran.r-project.org/package= DescTools

Smith WHF, Sandwell DT (1997) Global seafloor topography from satellite altimetry and ship depth soundings. Science 277:1956−1962

Stanistreet JE, Nowacek DP, Bell JT, Cholewiak DM and others (2018) Spatial and seasonal patterns in acoustic detections of sperm whales Physeter macrocephalus along the continental slope in the western North Atlantic Ocean. Endang Species Res 35:1−13

Taburet G, Sanchez-Roman A, Ballarotta M, Pujol MI and others (2019) DUACS DT2018: 25 years of reprocessed sea level altimeter products. Ocean Sci 15:1207−1224

Taylor BL, Baird R, Barlow J, Dawson SM and others (2008) Physeter macrocephalus. The IUCN Red List of Threatened Species 2008:e.T41755A10554884 (downloaded on 23 July 2018) https://dx.doi.org/10.2305/IUCN.UK.2008. RLTS.T41755A160983555.en

Thode A, Mellinger DK, Stienessen S, Martinez A, Mullin K (2002) Depth dependent acoustic features of diving sperm whales (Physeter macrocephalus) in the Gulf of Mexico. J Acoust Soc Am 112:308−321

Thorp WH (1967) Analytic description of the low-frequency attenuation coefficient. J Acoust Soc Am 42:270

van Wyk JA (2015) Defining the blue economy as a South African strategic priority: toward a sustainable 10th province? J Indian Ocean Reg 11:153−169

Watkins WA, Schevill WE (1977) Sperm whale codas. J Acoust Soc Am 62:1485−1490

Watwood SL, Miller PJO, Johnson M, Madsen PT, Tyack PL (2006) Deep-diving foraging behaviour of sperm whales (Physeter macrocephalus). J Anim Ecol 75: 814−825

Weilgart LS, Whitehead H (1988) Distinctive vocalizations from mature male sperm whales Physeter macrocephalus. Can J Zool 66:1931−1937

Weilgart L, Whitehead H (1993) Coda communications by sperm whales (Physeter macrocephalus) off the Galápagos Islands. Can J Zool 71:744−752

Whitehead H (1993) The behaviour of mature male sperm whales on the Galápagos Islands breeding grounds. Can J Zool 71:689−699

Whitehead H (1996) Variation in the feeding success of sperm whales: temporal scale, spatial scale and relationship to migrations. J Anim Ecol 65:429−438

Whitehead H (2002) Estimates of the current global population size and historical trajectory for sperm whales. Mar Ecol Prog Ser 242:295−304

Whitehead H, Weilgart L (1991) Patterns of visually observable behaviour and vocalizations in groups of female sperm whales. Behaviour 118:275−296

Wood SN (2017) P-splines with derivative based penalties and tensor product smoothing of unevenly distributed data. Stat Comput 27:985–989

Wright MN, Ziegler A (2017) ranger: a fast implementation of random forests for high dimensional data in C++ and R. J Stat Softw 77:1–17

Zhang HM, Reynolds RW, Bates JJ (2006) Blended and gridded high resolution global sea surface wind speed and climatology from multiple satellites: 1987–present. American Meteorological Society 14th Conference on Satellite Meteorology and Oceanography , 29 January – 2 February 2006, Atlanta, GA, P2.23

Zimmer WMX, Tyack PL, Johnson MP, Madsen PT (2005) Three-dimensional beam pattern of regular sperm whale clicks confirms bent-horn hypothesis. J Acoust Soc Am 117:1473–1485