Vol. 746: 121–140, 2024
https://doi.org/10.3354/meps14693

MARINE ECOLOGY PROGRESS SERIES
Mar Ecol Prog Ser

Published October 8

Check for updates

OPEN ACCESS CC BY

Accounting for sperm whale population demographics in density estimation using passive acoustic monitoring

Alba Solsona-Berga[1,*], Kaitlin E. Frasier[1], Natalie Posdaljian[1], Simone Baumann-Pickering[1], Sean Wiggins[1], Melissa Soldevilla[2], Lance Garrison[2], John A. Hildebrand[1]

[1]Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA 92037, USA
[2]NOAA Southeast Fisheries Science Center, Miami, FL 33149, USA

ABSTRACT: Sperm whales *Physeter macrocephalus* are highly sexually dimorphic, with adult males having larger bodies, more powerful echolocation clicks, and slower echolocation clicking rates compared to females. This study introduces methods for estimating sperm whale population densities in the Gulf of Mexico (GoMex) by accounting for the population demographics using passive acoustic monitoring and reveals that ignoring the differences between demographic segments can introduce bias in density estimates. Weekly densities were estimated per 3 demographic segments: social groups consisting of adult females and their offspring, mid-size animals, and adult males. Analysis revealed that the GoMex sperm whale population is primarily composed of social groups, which account for 92 to 98% of the overall population. Mid-size animals and adult males made up a small proportion of the population and were only intermittently present. Our 7 yr GoMex density estimates, including the 2010 *Deepwater Horizon* (DWH) oil spill period and subsequent years, revealed demographic-specific trends. Declines found at 2 north-central GoMex sites, coupled with increases at a southeastern site, may indicate population movements and potential impacts from the 2010 DWH oil spill and elevated noise levels from anthropogenic activities in the north-central GoMex.

KEY WORDS: *Physeter macrocephalus* · Sperm whale · Population demographics · Density estimation · Passive acoustic monitoring

1. INTRODUCTION

Human pressures, including underwater noise, ship collisions, fisheries interactions, pollution, and climate change, hinder the recovery of sperm whale *Physeter macrocephalus* populations (National Academies of Sciences, Engineering, and Medicine 2017, Sousa et al. 2019). Historical commercial whaling severely reduced global abundance by 57% (Whitehead & Shin 2022), leading to endangered status under the US Endangered Species Act, depleted under the US Marine Mammal Protection Act, and Vulnerable under the IUCN Red List of Threatened Species (Taylor et al. 2019). Sperm whales in the northern Gulf of Mexico (GoMex) are managed as a separate stock under the Marine Mammal Protection Act with an estimated population of 1180 animals and an average density of 1.7 animals 1000 km^{-2} (2017–2018 period; Garrison et al. 2020, Hayes et al. 2022). While still recovering from historical commercial whaling (Townsend 1935, Reeves et al. 2011), this species faces contemporary challenges related to habitat degradation. The GoMex region experiences exceptionally high levels of noise pollution from seismic surveys for oil and gas exploration and trafficked shipping lanes (Wiggins et al. 2016, Estabrook et al. 2016). Additionally, the lingering ef-

*Corresponding author: asolsonaberga@ucsd.edu

© A.Solsona-Berga, K.E. Frasier, N. Posdaljian, S. Baumann-Pickering, S. Wiggins, J.A. Hildebrand and outside the USA the US Government 2024. Open Access under Creative Commons by Attribution Licence. Use, distribution and reproduction are unrestricted. Authors and original publication must be credited.

Publisher: Inter-Research · www.int-res.com

fects of the 2010 *Deepwater Horizon* (DWH) oil spill — the largest in US history (Ramseur 2010, Levy & Gopalakrishnan 2010) — pose a unique challenge to this population. The consequences of these stressors remain poorly understood (Farmer et al. 2018), highlighting the need to monitor changes and inform management and conservation efforts.

Reliable estimates of population abundance or density trends are crucial indicators of species' status in the wild but remain challenging, particularly for pelagic species like sperm whales. Passive acoustic monitoring (PAM) is an attractive approach for remote long-term data collection, population estimation (Frasier et al. 2016, von Benda-Beckmann et al. 2018), and trend analysis (Hildebrand et al. 2015). Assessment of animal density and abundance from autonomous PAM recorders relies on the detection and counting of individual or group vocalizations in a given area and time period (Marques et al. 2013). Animal vocalization rates, source levels, and group sizes are key scalar parameters required for accurate density estimation from such acoustic recorders (Marques et al. 2009, Hildebrand et al. 2015), and to date, these parameters are assumed to apply at the species or population level. These assumptions may introduce bias in density estimates, as sperm whales have complex population demographics including social and sexual maturity segregation across different latitudinal ranges for most populations (Best 1979, Lyrholm et al. 1999), as well as sexual dimorphism in body size, which has been linked to differences in echolocation click characteristics and diving behaviors (Gordon 1991, Watwood et al. 2006, Growcott et al. 2011, Solsona-Berga et al. 2022).

Adult females and immature animals form social groups in low and mid-latitudes (Best 1979, Rice 1989). In contrast, maturing males form bachelor groups of similar-aged animals and become increasingly solitary, moving to higher latitudes as they mature (Best 1979, Whitehead 2003). Males transit to lower-latitude breeding grounds, but the timing remains poorly understood (Rice 1989). The GoMex population is mostly comprised of adult females and immature animals, with smaller body and group sizes than in other ocean basins (Jaquet & Gendron 2009). Adult male movements and breeding times in this regional population are still unknown. The little knowledge we have, derived from studies of genetic diversity and occasional visual observations showing both sexes moving across the GoMex basin, with some males breeding in different ocean basins (Lyrholm et al. 1999, Alexander et al. 2016), is complemented by findings from a recent long-term PAM study (Solsona-Berga et al. 2022). This study revealed spatial and temporal variability of the northern GoMex population demographics, including seasonal patterns and possible male presence year-round.

Adult males, which have larger bodies, produce more powerful echolocation clicks at a slower rate than adult females and juveniles (Goold & Jones 1995, Solsona-Berga et al. 2022). Larger heads of adult males may allow the buildup and discharge of greater volumes/pressures of air during sound production, resulting in higher source level emissions (Goold & Jones 1995). Differences in clicking rates between sexes may be related to the maximum detection range for prey (Jensen et al. 2018). With a larger echolocation range, adult males may slow down the interval between clicks to wait for more distant echoes. Differences in clicking rates, detectability, and group sizes among sex/age groups (Douglas et al. 2005) may significantly impact acoustic density estimation of sperm whales. Calculating demographic-specific acoustic density estimates could reduce error introduced by averaging parameters that are extremely different across demographic segments, and facilitate analyses of demographic-specific population trends.

Using long-term PAM, we present a framework for estimating demographic-specific sperm whale densities using 2 approaches, cue and group counting, accounting for differing clicking rates, group sizes, and detectability among demographic segments. Our 7 yr GoMex density estimates, including the 2010 DWH oil spill period and subsequent years, revealed demographic-specific trends. Declines found at 2 north-central GoMex sites and increases at a southeastern site could be linked to population movements, potential DWH oil spill impacts, and anthropogenic activities.

2. MATERIALS AND METHODS

Sperm whales were monitored at 3 northern GoMex locations during and following the DWH oil spill (2010–2017). These monitoring sites (Fig. 1) included one near Mississippi Canyon (MC) within 15 km of the DWH wellhead, another near Green Canyon (GC), located outside and northwest of the DWH surface oil footprint, and a third near Dry Tortugas (DT), outside and southeast of the oil footprint (Table S1 in the Supplement at www.int-res.com/articles/suppl/m746p121_supp.pdf).

At each site, a High-frequency Acoustic Recording Package (Wiggins & Hildebrand 2007) recorded

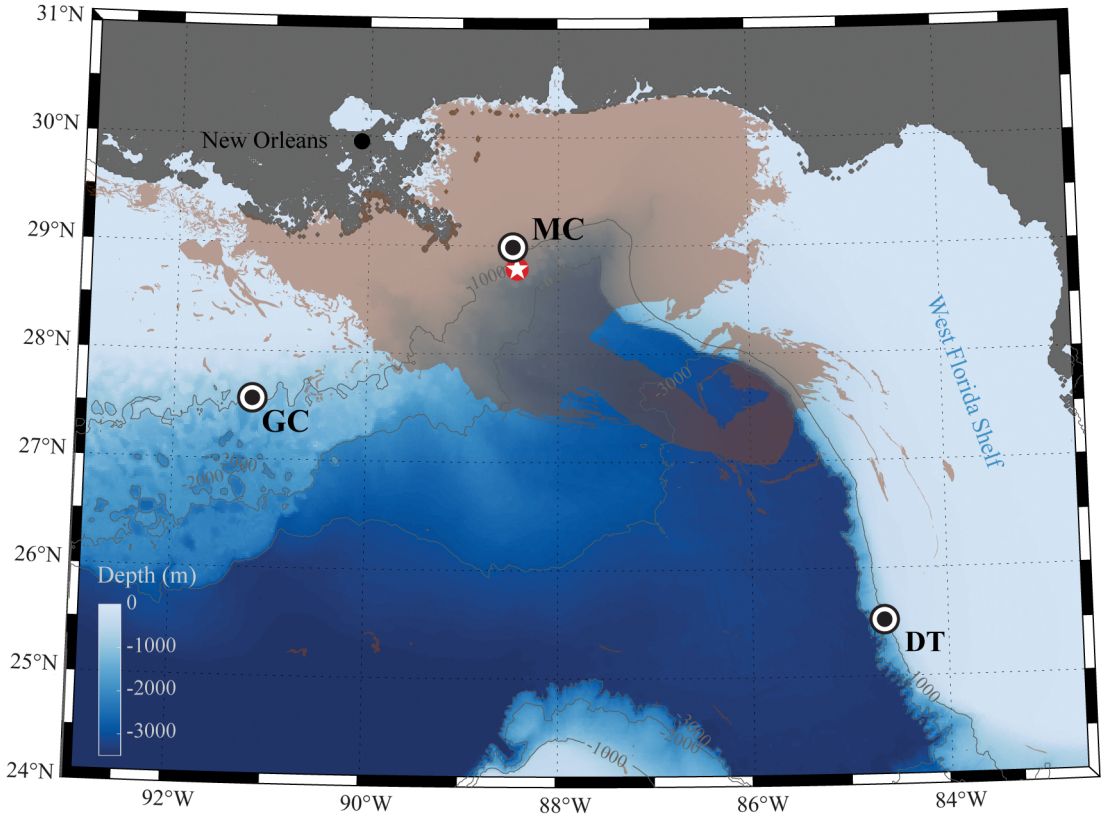

Fig. 1. Acoustic monitoring sites in the northern Gulf of Mexico from 2010 to 2017, named for nearby oceanographic features: Green Canyon (GC), Mississippi Canyon (MC), and Dry Tortugas (DT). *Deepwater Horizon* wellhead is shown by a star, with cumulative surface oil footprint (Kobara 2019) in brown. Bathymetric contours at 1000 m depth increments are illustrated

sound nearly continuously at a sampling rate of 200 kHz. Sperm whale echolocation clicks were automatically detected and classified for use as cues for estimating weekly densities. All analyses, including signal detection, classification, and subsequent density and trend analyses were carried out in MATLAB R2016b (Mathworks). Signal detection and classification was performed using an automated multi-step approach described by Solsona-Berga et al. (2022). This approach implemented a band-pass filter between 5 and 95 kHz to reduce background noise. Detections were filtered out at a peak-to-peak received sound pressure level (RL) of 135 dB_{pp} re: 1 µPa to establish a consistent detection threshold. Other marine mammals like beaked whales or delphinids were automatically filtered out using spectral click characteristics such as peak frequency and spectral shape (Solsona-Berga et al. 2022). Times during ship passages were also excluded because sperm whale clicks were often indistinguishable from ship-related noise. An automated vessel detector, Triton Ship-Detector (Supplement of Solsona-Berga et al. 2020), identified ship passages as times of increased noise in 3 specific frequency bands: 1–5, 5–10, and 10–50 kHz (see detector settings in https://github.com/MarineBioAcousticsRC/Triton/wiki/Ship-Detector). Reduced recording effort resulting from the removal of these noise-dominated time periods was accounted for in density estimates by adjusting recording effort parameters to reflect the duration of data suitable for sperm whale detection. Sperm whale detections were grouped into encounters and manually validated in DetEdit software (Solsona-Berga et al. 2020) by 2 analysts (A.S.B., N.P.). While this process ensured classification accuracy of encounters, not every individual detection was evaluated. Sperm whale social vocalizations (codas, creaks, and slow clicks) and creaks for foraging were not included in this analysis. The high RL threshold of the detector prevented the detection of codas and creaks, which have lower source levels compared to regular echolocation clicks (usual clicks). Although slow clicks were detectable with these settings, they were not found during the DetEdit validation process.

2.1. Population demographic classes

Sperm whale demographics were incorporated into density estimates to reduce potential sex/age class bias by considering variability in clicking rates, detectability, and group sizes among size classes. Echolocation click repetition rate, measured as the interclick interval (ICI), served as a proxy for body length to infer demographic structure (Solsona-Berga et al. 2022, Posdaljian et al. 2024, Westell et al. 2024). This method transformed ICIs into time series of modal ICI distributions within 5 min windows, associating dominant ICI patterns to body size classes. In this study, we used the same time series of size classes found for GoMex sperm whales by Solsona-Berga et al. (2022). Sperm whale detections were categorized into 3 age/sex groups (referred to hereafter as demographic classes) in 5 min bins based on ICI distributions: (1) small animals (<12 m) with ICIs of 0.44–0.70 s, presumed to be social groups of adult females and their offspring; (2) mid-size animals (11–13 m) with ICIs of 0.70–0.85 s, presumed to be adult females or subadult males; and (3) large animals (13–15 m) with ICIs of 0.85–1.0 s, presumed to be adult males (Solsona-Berga et al. 2022).

2.2. Density estimation

Animal densities were estimated using 2 point-transect distance sampling approaches: cue and group counting (Buckland et al. 2001, Buckland 2006). Cue counting used individual click detections as the fundamental unit for density estimation. Group counting used presence or absence of click detections in discrete time windows as the unit of analysis (e.g. Hildebrand et al. 2015). Both approaches require knowledge of detector performance, cue properties, and cue detectability around the sensor (Marques et al. 2009).

Cue counting converted the number of detected clicks, $n_{kt,}$ at site k during week t into animal density $\hat{D}_{kt}$, based on an estimator of animal density using acoustic cues recorded from fixed single sensors (Marques et al. 2009):

$$\hat{D}_{kt} = \frac{n_{kt}\left(1-\hat{c}_{pk}\right)\left(1+\hat{c}_{nk}\right)}{\pi\, w^2\, \hat{P}_k\; T_{kt}\, \hat{r}_k} \tag{1}$$

where $\hat{c}_{pk}$ is the proportion of false positive detections at site k, $\hat{c}_{nk}$ is the proportion of false negative detections at site k, $\hat{P}_k$ is the probability of detecting a click within a horizontal radius of size w (beyond which no detections are assumed possible) at site k, T_{kt} represents the total time monitored at site k during week t, and $\hat{r}_k$ is the estimated click production rate at site k.

Group counting required detection of animal presence within time windows (Hildebrand et al. 2015), along with knowledge of the detectability of clicks produced by the observed group of animals. It relies on knowledge of group size and group vocalization behavior. Using group counting, the estimated density D_{kt} at site k, during week t is:

$$\hat{D}_{kt} = \frac{n_{kt}\left(1-\hat{c}_{pk}\right)\left(1+\hat{c}_{nk}\right)\hat{s}}{\pi\, w^2\, \hat{P}_k\;\; \hat{P}_v\;\; T_{kt}} \tag{2}$$

where n_{kt}, c_{pk}, c_{nk}, and P_k are similar to those in Eq. (1) but refer here in particular to time bins (5 min) with group detections rather than individual click detections. $\hat{s}$ represents the mean group size, and $\hat{P}_v$ is the probability of a group being vocally active in a bin.

The coefficient of variation of estimated densities, CV(D_{kt}), was obtained using the delta method approximation (Seber 1982), to incorporate uncertainty in detection performance, probability of detection, and click rate for cue counting:

$$\mathrm{CV}\left(\hat{D}_{kt}\right) = \left\{\hat{D}_{kt}^{\,2} \times \left[\mathrm{CV}\left(\hat{c}_{pk}\right)^2 + \mathrm{CV}\left(\hat{c}_{nk}\right)^2 + \mathrm{CV}\left(\hat{P}_k\right)^2 + \mathrm{CV}\left(\hat{r}_k\right)^2\right]\right\}^{1/2} \tag{3}$$

and uncertainty in group size and vocalization probabilities for group counting:

$$\mathrm{CV}\left(\hat{D}_{kt}\right) = \left\{\hat{D}_{kt}^{\,2} \times \left[\mathrm{CV}\left(\hat{c}_{pk}\right)^2 + \mathrm{CV}\left(\hat{c}_{nk}\right)^2 + \mathrm{CV}\left(\hat{P}_k\right)^2 + \mathrm{CV}\left(\hat{s}\right)^2 + \mathrm{CV}\left(\hat{P}_v\right)^2\right]\right\}^{1/2} \tag{4}$$

Weekly densities were estimated per demographic class identified in the data set, considering density parameters specific to each class (social groups, mid-size animals, and adult males). Inclusive densities, estimated as the summed densities of the different population segments accounting for demographic differences (e.g. click production rate, group size, probability of detection, probability of vocal activity), were compared with generalized densities, which were estimated without consideration of demographic differences.

2.2.1. Detections and detector performance

To estimate the total number of detected clicks per demographic class each week (n_{kt}), the class given to

each time bin was applied to all detected clicks within that bin. If a bin was assigned to more than 1 class (n = 280 216; 0.3% of the total bins with sperm whale presence), the clicks in that bin were proportionally assigned a class based on the proportion of classes for that week. Some bins were not candidates for class categorization due to low numbers of clicks or lack of neighboring time bins to inform categorization (less than 9% of bins). To estimate the proportion of false positive clicks at each site ($\hat{c}_{pk}$), a random subsample of clicks (n = 11 000) that was evenly distributed across the entire data set was selected and assessed using DetEdit's evaluation tool (Solsona-Berga et al. 2020). The proportion of false positive 5 min bins ($\hat{c}_{pk}$) was calculated by evaluating the bin containing the randomly selected clicks. False negatives ($\hat{c}_{nk}$) were estimated using a simulation method incorporating the distribution of received click amplitudes, including those that did not pass the criteria of the detector spectral click characteristic, to account for all missed detections of clicks with RLs ⩾135 dB_{pp} re: 1 µPa (described in Section 2.2.2 below).

2.2.2. Detection probability

A Monte Carlo simulation estimated the probability of detecting an individual click or any clicks from a group of echolocating animals in 5 min bins ($\hat{P}_k$) over a horizontal distance range around each sensor (Frasier et al. 2016). This simulation aimed to determine received levels of expected clicks, accounting for signal characteristics, diving behavior, acoustic environmental propagation, and detector performance. This simulation method selected random samples (over 500 iterations) from a uniform distribution of animal positions and orientations (n = 100 000 iteration^{-1}) with respect to the hydrophone (Table 1). A mean detection probability and its variance was computed as the ratio of detected clicks or groups to the total number of simulations (Frasier et al. 2016). Unlike previously implemented Monte Carlo simulations from fixed single sensors (Hildebrand et al. 2015, 2019, Frasier et al. 2016), this approach used: (1) a bimodal diving pattern (Fig. S1) rather than one unique diving pattern to resemble tag data diving behavior from GoMex sperm whales, where animals featured search and foraging acoustic behavior in the mid-water column and seafloor transits referred here as benthic dives (Mate et al. 2017), and (2) a 3-dimensional forward-beam pattern which was allowed to vary in the backward direction based on a circular piston model, as shown in sperm whales when echolocating (Zimmer et al. 2005).

To estimate demographic-specific detection probabilities, 6 scenarios were simulated per site (1 for each of the 3 demographic classes for both the cue and group counting methods) to account for differences in echolocation click characteristics and dive behavior based on body size (Gordon 1991, Watwood et al. 2006, Growcott et al. 2011, Solsona-Berga et al. 2022) and the different bathymetric features and depths of each site (Fig. 1; Table S1). Click peak frequencies were obtained from the data for each site and class (Table 1). Frequency-dependent propagation loss volumes at each site at the typical peak frequency per class were simulated using the beam-tracing algorithm 'Bellhop' (Porter & Bucker 1987), with site-specific environmental and physical parameters (see settings, databases, and versions in Frasier et al. 2016) drawn from the Oceanographic and Atmospheric Master Library (OAML) using ESME Workbench (Mountain et al. 2012). Simulations used previously published diving behavior parameters of the GoMex population (e.g. start clicking depth and foraging phase body angle; Table 1); however, dive parameters have only been documented for mid-water column dives with unknown relative bottom depth (Watwood et al. 2006). Proportion of time spent in each bimodal dive is unknown for the GoMex population, as are the female source levels and beam patterns, which have only been measured for adult males in the Mediterranean, Bahamas, and Norway (Møhl et al. 2003, Zimmer et al. 2005, Nosal & Frazer 2007). For the unknown parameters for specific demographics or only documented for other populations, a grid-search was used to obtain a distribution of values from which to sample the mean and variability for the simulations (Hildebrand et al. 2019).

The grid-search approach, as described by Hildebrand et al. (2019), optimized parameter values by assessing a range of potential values, pre-selected based on literature for other populations or demographics (Table S2). Simulations were run with all other parameters held constant while the distribution of the parameter of interest was inspected for best goodness-of-fit by comparing model results with actual data. All possible combinations of parameters for optimization within the pre-selected ranges were tested. Optimized parameter values best fitting the model output were determined by minimizing the sum of the squared misfit for clicks or bins of any amplitude with RLs ⩾135 dB_{pp} re: 1 µPa (Table 1). Final values were used in 6 scenario simulations per site to estimate the probability of detection ($\hat{P}_k$). As expected, more clicks were detected at lower received levels as the distance from the hydrophone

Table 1. Monte Carlo detectability simulation parameters, along with their modeled sensitivity. Unique demographic class, site, or counting method parameters are listed separately. Cue counting uses individual click detections, while group counting uses 5 min time bins with group detections as the unit of analysis. RL: received sound pressure level; TL: transmission loss; MC: Mississippi Canyon; GC: Green Canyon; DT: Dry Tortugas

Parameter		Distribution sample mean	SD	Reference	$\hat{P}_k$ cue sensitivity	$\hat{P}_k$ group sensitivity
Max. detection range (km)		35	–	–		
Min. click RL (dB_{pp})		135	–	–		
Signal characteristics (demographic-specific)						
Source level (dB_{pp})	Social group	237 ± 5[a]	2–3	Grid search	17% dB^{-1}	6% dB^{-1}
	Mid-size	238 ± 5[a]		(Møhl et al. 2003,	19% dB^{-1}	5% dB^{-1}
	Adult male	242 ± 5[a]		Zimmer et al. 2005)	17% dB^{-1}	6% dB^{-1}
Peak frequency (kHz)	Social group	10	–	Data	–	–
	Mid-size	9			–	–
	Adult male	9			–	–
Directivity (dB)	Social group	30.5 ± 2.5	–	Grid search	−9% dB^{-1}	–
	Mid-size	32.5 ± 2.5		(Møhl et al. 2003,	−4% dB^{-1}	–
	Adult male	32.5 ± 2.5		Zimmer et al. 2005, Nosal & Frazer 2007)	−5% dB^{-1}	–
90° off-axis TL (dB)		37.5 ± 2.5	–	Zimmer et al. (2005), Nosal & Frazer (2007)	−5% dB^{-1}	–
180° off-axis TL (dB)		27.5 ± 2.5	–	Zimmer et al. (2005)	−5% dB^{-1}	–
Minimum off-axis TL (dB)		35.5 ± 2.5	2–3	Grid search (Zimmer et al. 2005)	–	−1% dB^{-1}
Diving behavior (site-specific)						
Water column dive altitude (m)	MC	450 ± 50	10–30	Grid search	0% $50m^{-1}$	−3% $50m^{-1}$
	GC	650 ± 50		(Watwood et al. 2006)	0% $50m^{-1}$	−2% $50m^{-1}$
	DT	850 ± 50			2% $50m^{-1}$	−1% $50m^{-1}$
Water column dive fraction (%)	MC	55 ± 5	1–2	Grid search	6% $10\%^{-1}$	−3% $10\%^{-1}$
	GC	55 ± 5		(Watwood et al. 2006)	3% $10\%^{-1}$	−5% $10\%^{-1}$
	DT	95 ± 5			4% $10\%^{-1}$	−1% $10\%^{-1}$
Benthic dive altitude (m)		10	10–30	Mate et al. (2017)	–	–
Benthic dive fraction (%)	MC	45 ± 5	1–2	Grid search	−3% $10\%^{-1}$	5% $10\%^{-1}$
	GC	45 ± 5		(Mate et al. 2017,	−4% $10\%^{-1}$	4% $10\%^{-1}$
	DT	5 ± 5		Irvine et al. 2017)	−5% $10\%^{-1}$	3% $10\%^{-1}$
Start clicking depth (m)		210 ± 10	65–75	Watwood et al. (2006)	1% $50m^{-1}$	0% $50m^{-1}$
Foraging phase	Cue	0[b]	0–50	Watwood et al. (2006)	1% deg^{-1}	–
Vertical body angle (deg)	Group	0[b]	0–65	Watwood et al. (2006)	–	1% deg^{-1}
Foraging phase horizontal body angle (deg)		0 ± 360	–	–	–	–

[a]Site DT estimated source levels were 2 dB lower than the other sites (social group: 235 ± 5 dB_{pp}, mid-size animals: 236 ±5 dB_{pp}, and adult males: 240 ±5 dB_{pp})
[b]Normal distribution, left truncated at 0 degrees

increased, due to the expanded observable area. The misfit of the number of clicks at lower received levels between the model and the actual data suggested missed detections, providing an opportunity to quantify them. False negatives ($\hat{c}_{nk}$) were estimated as the sum of the misfit between the final model and the measured data for the received level clicks or bins immediately above the threshold (135–140 dB_{pp} re: 1 µPa).

2.2.3. Vocal activity

Click production rate ($\hat{r}_k$), which is required for cue counting, may vary among regions, locations, age groups, and sexes (Table S3). To account for this variability, click production rate was calculated as:

$$\hat{r}_k = \frac{\hat{P}_{cyc}}{\Sigma\widehat{\mathrm{ICI}}_k} \quad (5)$$

where $\hat{P}_{cyc}$ is the mean proportion of time spent clicking, and $\widehat{\mathrm{ICI}}_k$ is the ICI at each site k, and of each demographic class when estimating demographic-specific estimates. Proportion of time spent clicking ($\hat{P}_{cyc}$) was computed using behavioral data reported from acoustic tag records in the GoMex (Watwood et al. 2006), multiplying the percentage of time spent in search phase producing regular clicks during a foraging dive (81 ± 5% SD) by the total time of a foraging dive, including resting and socializing periods at the surface without echolocation (72 ± 33%). Tag data in the GoMex only included female and immature animals, but populations in sub-tropical and temperate latitudes, including both sexes, spend equivalent amounts of time clicking (Watwood et al. 2006, Ward et al. 2012), hence the same mean proportion of time clicking was used for all classes.

ICI was computed per class and site using exclusive clicks from single demographic bins. ICI distribution per class was fit with a Gaussian distribution to estimate mean and coefficient of variation. A relatively constant click rate is maintained during a dive (Madsen et al. 2002b, Zimmer et al. 2003, Douglas et al. 2005) but variability exists (e.g. during terminal buzzes, descent, and ascent phases; Madsen et al. 2002a,b, Thode et al. 2002, Teloni et al. 2008). Only dominant click rate per class was estimated using this approach, restricted between 0.3 and 1.4 s based on the demographic classification method, as not all dive variability could be equally quantified with these data. This restriction excluded the tail of the distributions, acknowledging the limitations in accounting for all dive variability.

Assuming that the probability of detecting a group of animals during a given time window increases with group size and the synchrony in their clicking behavior (i.e. the extent to which the click sequences of animals within the group overlap), we estimated the required probability of a group being vocally active ($\hat{P}_v$) for group counting using the method of Hildebrand et al. (2019):

$$P_v = \left\{\hat{P}_{cyc} \times [s - (s-1) \times o]\right\} \quad (6)$$

where $\hat{P}_{cyc}$ is the mean proportion of time spent clicking by one animal, s is the group size, and o is the synchrony of clicks among group members in a bin. Each animal added to the group is assumed to contribute both overlapped (synchronous) and non-overlapped (asynchronous) echolocation time to the bin. Adult females and their young form social groups with an average of 6.1 ± 4.8 animals in the GoMex (Jochens et al. 2008). Group sizes of adult males in the GoMex remain unknown but were expected to be 1–2 animals, considering their typical solitary behavior and limited interactions with social groups (Best 1979, Gero et al. 2014, Cantor et al. 2019). Sub-adult males form bachelor groups in which the group size reduces with age (Gaskin 1970, Ohsumi 1971, Best 1979), and aggregation of males have been documented in various latitudes dispersed over ~10–30 km (e.g. Gillespie 1979, Leaper & Scheidat 1998). An ad hoc test directly counting the number of interleaved received click sequences as a proxy for group size when animals were detected near the sensor (Hildebrand et al. 2015) supported the assumptions of group sizes of social groups and adult males (Text S1), and group sizes of 1–2 mid-size animals were estimated to be detected simultaneously, similar to adult males.

Three GoMex sperm whales from the same group were simultaneously tagged in a previous study (Jochens et al. 2008). This offered a rare opportunity to estimate synchrony of clicks among group members. Synchrony (o) and variation were estimated by measuring proportion of time, defined in 5 min bins, during which each pairwise combination of the 3 animals overlapped (Fig. S2). This showed a synchrony of 77 ± 4% among pairs of whales within the group.

2.3. Long-term trend estimation

Site trends were estimated by removing seasonal variability from weekly densities to characterize trends independently of the significant seasonality present among demographic classes of the population in the northern GoMex (Solsona-Berga et al. 2022). Seasonality was removed using a monthly seasonal pattern decomposition procedure (Cleveland et al. 1990), and linear trend and rate of change per 1000 km^2 yr^{-1} was estimated using the Theil-Sen slope estimator (Sen 1968). This estimator computed the slope between each pair of points in a time series, using the true time difference, and estimated the overall slope as the mean across all pairs. This ensured a robust fit insensitive to outliers and short-term inter-annual variability. Means and associated 95% confidence intervals were generated using a parametric bootstrap with 100 iterations, each with 500 points selected to compute the associated slope.

3. RESULTS

Sperm whales were present year-round in 31% of 5 min bins at MC and 14% at GC over a cumulative monitoring effort of 5.4 yr per site, after removal of ship

noise-dominated time periods. Cumulative monitoring effort at the southern site (DT) was 4.7 yr, and sperm whales were detected in 5% of bins during that period (Table 2). Ship passage periods removed from detection effort varied by site and over time (Fig. S3). Northern sites had the most ship passages, with MC having the highest noise levels during the oil spill (April–August 2010) with more than 20 h wk^{-1} removed. In subsequent years, less than 10 h wk^{-1} were removed at this site. GC had its peak noise level between 2012 and 2016, removing 20 h wk^{-1}, compared to 5 h wk^{-1} removed at other times. The southeastern site (DT) had the fewest ship passages, removing less than 5 h wk^{-1}, with 2 short periods in 2012 and 2016 reaching 20 h wk^{-1}.

Most bins with sperm whale presence were assigned into 1 or more demographic classes (97% at MC, 96% at GC, 91% at DT; Table 2). All 3 classes were detected across all sites, with putative social groups being the most common (92% of the positive detection bins at MC, 87% at GC, and 74% at DT). Mid-size animals were moderately common (4% of detection bins at MC, 7% at GC, 14% at DT), and adult males were rare (1% of detection bins at MC, 3% at GC, 4% at DT). Presence of 2 classes in a bin were infrequent (0.3% of detections bins at MC, 0.1% at GC, 0.075% at DT), and cases where all 3 classes were detected simultaneously occurred only in 1 bin at MC (Table 2).

3.1. Demographic class differences in ICI and peak frequency

ICI varied by recording site within each class (Fig. 2, Table 3). ICI for social groups was consistent between the northern sites, and slightly larger at DT, resulting in an overall mean ICI of 482 ms with low variability between sites (SD = 55 ms, CV = 11%). Mid-size animals had consistent ICI between sites, with an overall mean ICI of 670 ms and small variability (SD = 63 ms, CV = 9%), while adult males exhibited greater variability between the northern sites, resulting in an overall mean ICI of 812 ms with slightly higher but still low variability between sites (SD = 96 ms, CV = 12%). Click rates were estimated using the ICI per class and site (Table 4), which resulted in a higher clicking rate for social groups than for adult males and mid-size animals.

Click peak frequencies varied by class, with social groups having higher peak frequencies than adult males or mid-size animals (Fig. 3). Mean peak frequencies of social groups were 10.2–10.5 kHz, those of mid-size animals were 9.0–9.4 kHz, and those of adult males were 8.8–9.8 kHz (Table 3). Estimation of mean peak frequencies at MC was complicated by a bimodal distribution (Fig. 3), with a low-frequency peak at ~7 kHz in addition to the primary peaks at >~10 kHz (social groups) and ~9–10 kHz (mid-size animals).

3.2. Grid-search parameter optimization

The grid-search approximated sperm whale diving behavior parameters and signal characteristics unknown for a particular demographic or only known from other populations. Table 1 contains the values that provided the best fit between modeled received levels and observed distributions per site and class for both cue and group counting (Fig. 4). Based on the grid-search optimization for signal characteristics, estimated echolocation source levels and directivity varied by class but not between sites. The one excep-

Table 2. Number of 5 min bins assigned to a single demographic class, to multiple classes, and where assignment was not possible per site. Total bins with sperm whale presence and monitoring effort as number of bins are indicated in *italics*. The percentage of relative presence of each class is reported by the total bins with positive presence at each site. MC: Mississippi Canyon; GC: Green Canyon; DT: Dry Tortugas

Bin category	MC		GC		DT	
	Bins (n)	%	Bins (n)	%	Bins (n)	%
Social group	162595	91.8	68471	86.5	17691	74.1
Mid-size animals	6687	3.8	5501	6.9	3249	13.6
Adult males	1499	0.8	2129	2.7	850	3.6
Social group and mid-size	417	0.2	54	0.1	14	0.1
Social group and adult males	144	0.1	34	<0.1	2	<0.1
Mid-size and adult males	45	<0.1	2	<0.1	0	<0.1
Social group and mid-size and adult males	1	<0.1	0	<0.1	0	<0.1
Unassigned	5775	3.3	2992	3.8	2064	8.6
Total bins positive presence	*177163*		*79183*		*23870*	
Total effort bins	*563593*		*562812*		*489990*	

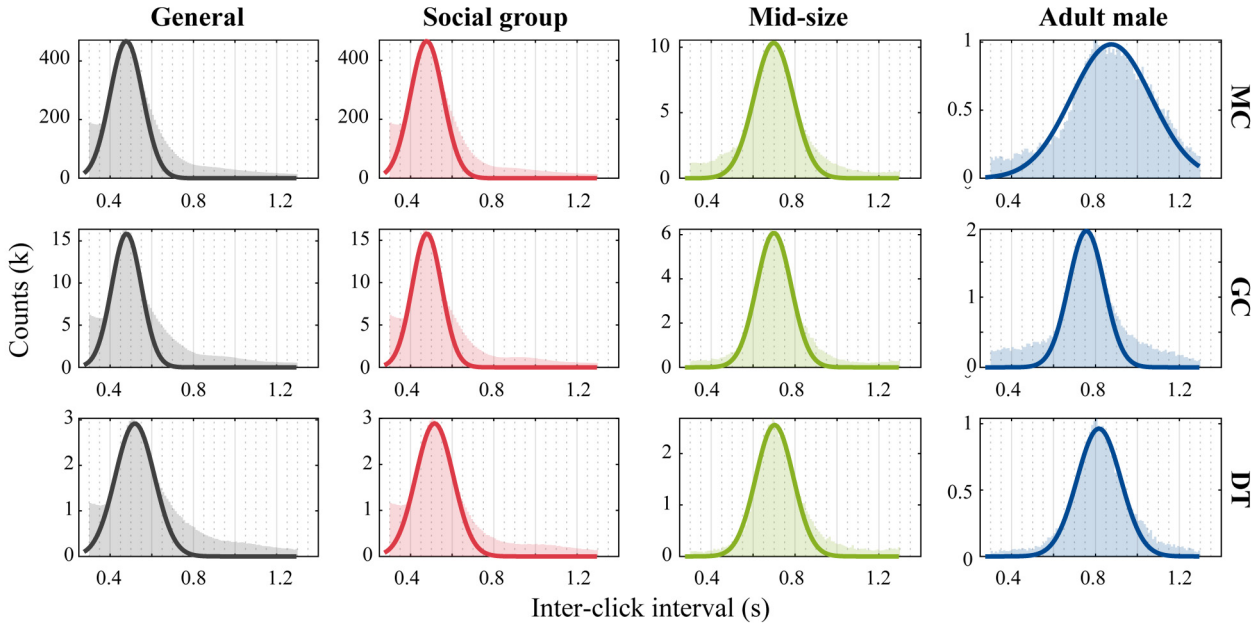

Fig. 2. Distribution of inter-click intervals (ICIs) for sperm whales by site, demographic class, and in general (all detections regardless of class), based on a histogram (shaded area) of ICIs with a bin width of 10 ms. A Gaussian distribution (line) omitting long tails was fit to obtain the mean and coefficient of variation. Refer to Fig. 1 for site abbreviations

Table 3. Inter-click interval (ICI) and peak frequency for sperm whale demographic classes per site. Mean and coefficient of variation (CV) of Gaussian fit for ICI and peak frequency distribution were estimated per class using only data from bins with a single class. MC: Mississippi Canyon; GC: Green Canyon; DT: Dry Tortugas

Site	Social group		Mid-size		Adult males		Generalized	
	Mean	CV	Mean	CV	Mean	CV	Mean	CV
ICI (ms)								
MC	478	0.117	698	0.095	876	0.154	483	0.116
GC	479	0.104	699	0.083	757	0.079	481	0.107
DT	515	0.124	702	0.090	817	0.091	524	0.127
Peak frequency (kHz)								
MC	10.2	0.106	9.0	0.097	9.8	0.101	10.1	0.111
GC	10.4	0.083	9.4	0.077	8.8	0.074	10.3	0.090
DT	10.5	0.110	9.3	0.088	9.4	0.083	10.2	0.120

tion was DT, where best fits resulted from source levels 2 dB lower for all classes (Table 1). Estimated source levels for social groups (237 ± 5 dB_{pp} re: 1 µPa @ 1m) and mid-size animals (238 ± 5 dB_{pp} re: 1 µPa @ 1m) were similar but lower than those estimated for adult males (242 ± 5 dB_{pp} re: 1 µPa @ 1m). Estimated directivity was the same for mid-size animals and adult males (32.5 ± 2.5 dB) and lower for social groups (30.5 ± 2.5 dB).

Optimized diving behavior parameters using grid-search resulted in different proportions of time spent on mid-water or benthic dives, as well as dive altitude across all sites (Table 1). Estimated mid-water column dive altitude was different at each site (MC = 450 ± 50 m; GC = 650 ± 50 m; DT = 850 ± 50 m). However, based on the varying bathymetric depths at each site, these altitudes suggest that the mid-water-column foraging depth was consistently near 500 m at all sites. A bimodal dive pattern with similar proportion of time spent in the mid-water (55 ± 5%) and benthic dives (45 ± 5%) was estimated at MC and GC. In contrast, foraging dives at DT were almost exclusively estimated to occur in mid-water (95 ± 5%).

Received click levels predicted by the final simulations for the 3 classes and sites were in good agreement with measured received levels in the range of 135–160 dB_{pp} re: 1 µPa (Fig. 4). Group counting, however, resulted in a lower overall fit. Above 160 dB_{pp} re: 1 µPa, fewer detections were measured than predicted using both counting methods for all classes and sites, likely as a result of too few recorded high-amplitude clicks. Near the detection received level threshold, fewer detections were measured than predicted, particularly using group counting. Greater numbers of low received level clicks are always expected relative to higher received level clicks because of a larger area for animals

Table 4. Sperm whale weekly mean densities and coefficient of variation (CV) derived from cue counting by site and demographic class. A maximum horizontal detection range (w) of 35 km was used to estimate densities. Estimated density trends regardless of class are referred to as generalized. MC: Mississippi Canyon; GC: Green Canyon; DT: Dry Tortugas

Class	Site	$\hat{D}_{kt}$ Mean density (animals 1000 km^{-2})	CV ($\hat{D}_{kt}$)	$n_{kt}\, T_{kt}^{-1}$ (n s^{-1})	$\hat{c}_{pk}$ False positive (% clicks)	CV ($\hat{c}_{pk}$)	$\hat{c}_{nk}$ False negative (% clicks)	CV ($\hat{c}_{nk}$) rate	$\hat{r}_k$ Click (n sec^{-1})	CV ($\hat{r}_k$)	$\hat{P}_k$ Probability of detection	CV ($\hat{P}_k$)
Demographic-specific												
Social group	MC	2.259	0.014	0.098	5.1	0.02	2.8	0.01	1.220	0.19	0.009	0.022
	GC	0.825	0.018	0.033	2.7	0.03	2.2	0.01	1.218	0.19	0.008	0.023
	DT	0.345	0.030	0.009	8.2	0.07	0.1	0.01	1.132	0.19	0.006	0.025
Mid-size	MC	0.048	0.029	0.002	5.1	0.02	2.7	0.01	0.836	0.19	0.012	0.022
	GC	0.032	0.029	0.001	2.7	0.03	5.7	0.01	0.834	0.19	0.011	0.023
	DT	0.023	0.042	0.001	8.2	0.07	1.3	0.01	0.831	0.19	0.008	0.026
Adult male	MC	0.007	0.030	0.0004	5.1	0.02	0.0	0.01	0.666	0.19	0.022	0.020
	GC	0.009	0.135	0.0005	2.7	0.03	4.0	0.01	0.770	0.19	0.020	0.020
	DT	0.006	0.036	0.0002	8.2	0.07	0.0	0.01	0.714	0.19	0.015	0.023
Generalized												
	MC	2.319	0.014	0.100	5.1	0.02	2.8	0.01	1.208	0.19	0.009	0.022
	GC	0.868	0.017	0.034	2.7	0.03	2.2	0.01	1.213	0.19	0.008	0.023
	DT	0.379	0.028	0.010	8.2	0.07	0.1	0.01	1.113	0.19	0.006	0.025

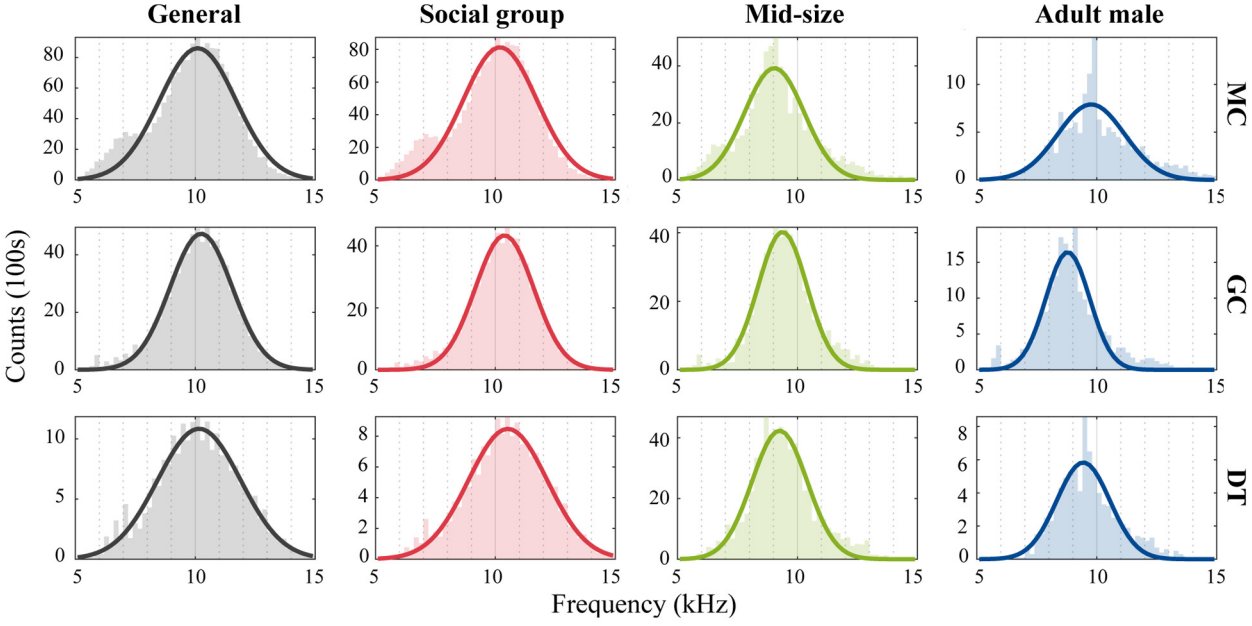

Fig. 3. Distribution of peak frequencies for sperm whale clicks by site, demographic class, and in general (all detections regardless of class). A Gaussian distribution (line), omitting long tails, was fit to obtain the mean and coefficient of variation. Refer to Fig. 1 for site abbreviations

to occupy as distance from the sensor increases. Since simulations selected random samples from a uniform distribution of animal positions and orientations in the horizontal plane around the hydrophone, the models indicate that groups were missed near the detection received level threshold, particularly for mid-size animals and adult males (Fig. 4, see Tables 4 & 5).

3.3. Detection probability

Probability of detecting sperm whales over a horizontal range (35 km) varied between demographic class and site for both counting methods (Fig. 5). Adult males were detectable at greater distances than social groups and mid-size animals, influenced main-

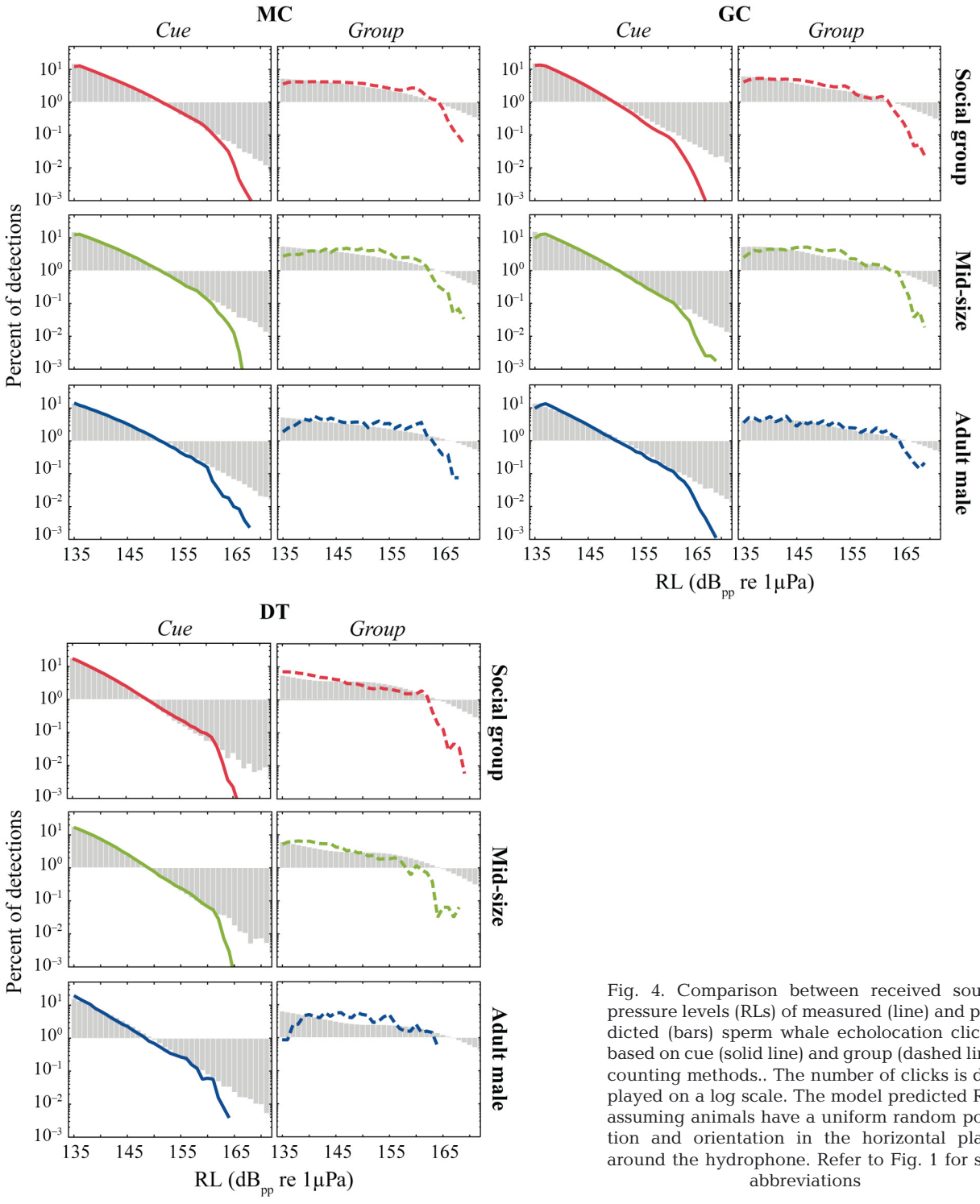

Fig. 4. Comparison between received sound pressure levels (RLs) of measured (line) and predicted (bars) sperm whale echolocation clicks, based on cue (solid line) and group (dashed line) counting methods.. The number of clicks is displayed on a log scale. The model predicted RLs assuming animals have a uniform random position and orientation in the horizontal plane around the hydrophone. Refer to Fig. 1 for site abbreviations

ly by higher mean source levels. A 1 dB_{pp} re: 1 µPa increase in mean source level led to a 17–19% increase in the mean detection probability of individual clicks and 5–6% increased detectability of a group of whales (Table 1). Mid-size animals had a slightly longer detection range than social groups, with marginally higher mean source levels and directivity. For all classes, individual clicks were rapidly less detectable between 2 and 5 km from the sensor at all sites, with the exception of DT, which was on average 10% lower than MC and GC. Differences in the detectability of individual clicks were likely driven by the steep

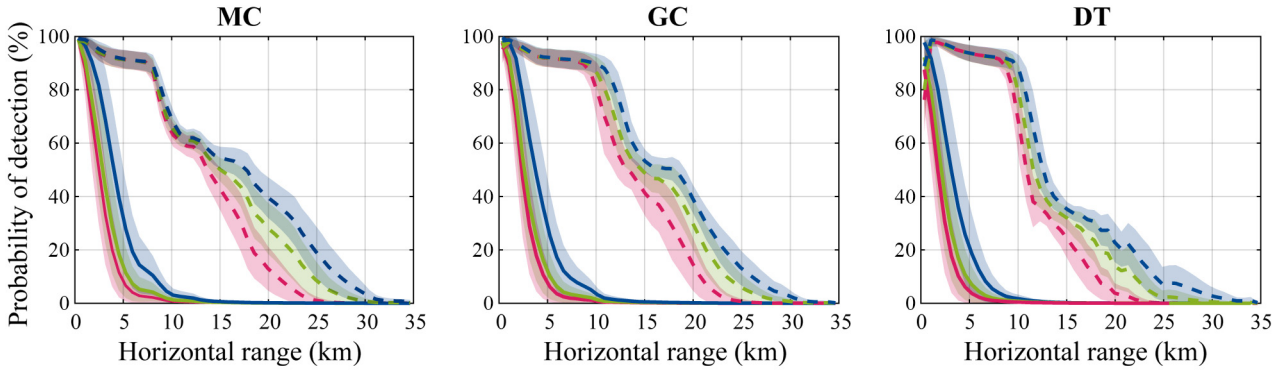

Fig. 5. Estimated mean detection probability for individual sperm whale clicks (solid lines) and groups (dashed lines) in 5 min bins for social groups (red), mid-size animals (green), and adult males (blue). Shaded areas show ± 1 SD. Refer to Fig. 1 for site abbreviations

bathymetry of this site. Clicks from a group of whales were detectable 70–90% on average at 8 km and declined rapidly after 8–15 km. Detectability of a group was also lower at greater distances at DT, with a probability of detecting 50% of a group at a distance below 15 km compared to 15–20 km at MC and GC.

3.4. Densities and trends in sperm whale demographic segments

Mean weekly densities per demographic class and site were estimated between 2010 and 2017 incorporating class-specific parameters using cue (Table 4) and group counting (Table 5). Social groups dominated across all sites, accounting for 92–98% of the overall densities, followed by mid-size animals representing 1–6% and adult males only 0.2–1.4%. Social groups occurred in high densities at MC, near the wellhead (2.3 and 2.7 animals 1000 km^{-2} for cue and group counting, respectively) and in lower densities at GC (cue and group: 0.8 and 1.1 animals 1000 km^{-2}). At these 2 northern sites, social groups were present year-round, but densities gradually decreased over the 7 yr period (Fig. 6, Table 6), with a greater annual decline at MC (cue and group: −3.4 and −7.4%) than at GC (cue and group: −1.6 and −2.4%). The southern site, DT, had the lowest den-

Table 5. Sperm whale weekly mean densities and coefficients of variation (CV) derived from group counting by site and demographic class. A maximum horizontal detection range (w) of 35 km was used to estimate densities. Estimated density trends regardless of class are referred to as generalized. MC: Mississippi Canyon; GC: Green Canyon; DT: Dry Tortugas

Class	Site	$\hat{D}_{kt}$ Mean density(animals 1000 km^{-2})	CV ($\hat{D}_{kt}$)	$n_{kt}\ T_{kt}^{-1}$ (n s^{-1})	$\hat{c}_{pk}$ False positive (% bins)	CV ($\hat{c}_{pk}$)	$\hat{c}_{pk}$ False negative (% bins)	CV ($\hat{c}_{pk}$)	$\hat{s}$ Mean group size (animals)	CV ($\hat{s}$)	$\hat{P}_v$ Prob group vocal	CV ($\hat{P}_v$)	$\hat{P}_k$ Probability of detection	CV ($\hat{P}_k$)
Demographic-specific														
Social group	MC	2.736	0.013	0.290	0.1	0.01	3.5	0.01	6.1	0.18	1.00	0.06	0.174	0.007
	GC	1.102	0.014	0.122	0.1	0.01	3.5	0.01	6.1	0.18	1.00	0.06	0.181	0.008
	DT	0.437	0.024	0.036	0.1	0.01	0.5	0.01	6.1	0.18	1.00	0.06	0.132	0.007
Mid-size	MC	0.037	0.035	0.013	0.1	0.01	9.5	0.01	1.5	0.25	0.65	0.04	0.228	0.008
	GC	0.027	0.033	0.010	0.1	0.01	7.5	0.01	1.5	0.25	0.65	0.04	0.233	0.007
	DT	0.024	0.038	0.007	0.1	0.01	5.0	0.01	1.5	0.25	0.65	0.04	0.173	0.009
Adult male	MC	0.007	0.044	0.003	0.1	0.01	8.0	0.01	1.5	0.25	0.65	0.04	0.283	0.007
	GC	0.008	0.083	0.004	0.1	0.01	2.0	0.01	1.5	0.25	0.65	0.04	0.279	0.006
	DT	0.006	0.052	0.002	0.1	0.01	18.0	0.01	1.5	0.25	0.65	0.04	0.218	0.009
Generalized														
	MC	2.971	0.012	0.314	0.1	0.01	3.5	0.01	6.1	0.18	1.00	0.06	0.174	0.007
	GC	1.273	0.014	0.141	0.1	0.01	3.5	0.01	6.1	0.18	1.00	0.06	0.181	0.008
	DT	0.589	0.021	0.049	0.1	0.01	0.5	0.01	6.1	0.18	1.00	0.06	0.132	0.007

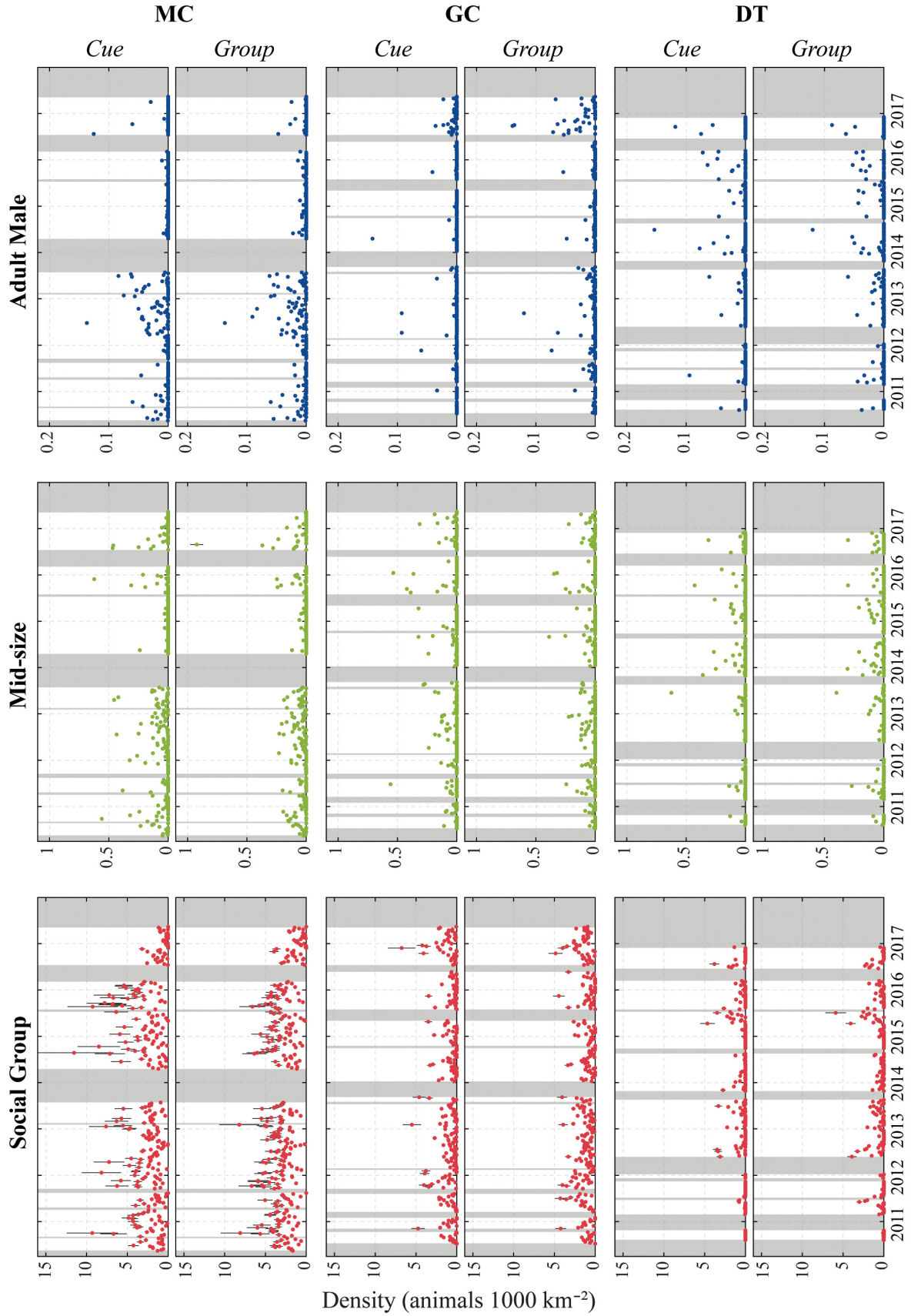

Fig. 6. Weekly site density estimates for sperm whale social groups, mid-size animals, and adult males based on click and group counting. Circles denote estimates and vertical lines show ± 1 SD. Shaded areas lack recording effort. Density scale varies for each demographic class. Refer to Fig. 1 for site abbreviations

Table 6. Mean density trends as an annual percent change and 95% confidence interval (CI) by site, demographic class, and counting method (cue, group) from 2010 to 2017. Estimated density trends regardless of class are referred to as generalized. MC: Mississippi Canyon; GC: Green Canyon; DT: Dry Tortugas

Site	Class	Cue Annual % change Mean	Cue Annual % change 95% CI	Group Annual % change Mean	Group Annual % change 95% CI	Trend
MC	Generalized	−3.77	[−6.07, −1.20]	−6.51	[−8.32, −5.03]	▼
	Demographic-specific					
	Social group	−3.41	[−6.31, −0.78]	−7.41	[−9.47, −5.31]	▼
	Mid-size	−6.33	[−11.93, −1.62]	−5.12	[−9.52, −1.43]	▼
	Adult male	−7.43	[−14.21, −1.59]	−10.07	[−15.45, −5.56]	▼
GC	Generalized	−0.25	[−4.01, 2.95]	−0.51	[−3.16, 1.51]	▼
	Demographic-specific					
	Social group	−1.61	[−5.36, 1.46]	−2.36	[−5.24, 0.96]	▼
	Mid-size	4.47	[0.86, 10.27]	3.39	[−0.07, 8.93]	▲
	Adult male	2.03	[1.09, 3.82]	14.36	[9.64, 18.33]	▲
DT	Generalized	16.27	[10.90, 25.02]	7.90	[5.02, 11.91]	▲
	Demographic-specific					
	Social group	15.87	[7.17, 21.80]	8.08	[4.13, 11.94]	▲
	Mid-size	9.47	[4.24, 14.49]	9.16	[3.68, 14.97]	▲
	Adult male	16.24	[9.11, 25.39]	11.25	[5.61, 17.88]	▲

sities of social groups (cue and group: 0.3 and 0.4 animals 1000 km^{-2}) where presence was most variable, consistently increasing in spring/summer (Fig. 6). In contrast with northern sites, social group densities increased over this period at DT (Table 6; cue and group: 16 and 8% annual change).

Mid-size animal densities were similar across sites, with mean weekly densities marginally higher at MC (cue and group: 0.05 and 0.04 animals 1000 km^{-2}) than GC and DT (0.03 and 0.02 animals 1000 km^{-2}, respectively; Tables 4 & 5). Mean densities of mid-size animal across sites were an order of magnitude smaller than the lowest mean densities of social groups at DT. Mid-size animal densities fluctuated among the 7 years (Fig. 6), and different trends were observed across sites (Table 6). Densities declined at MC (cue and group: −6 and −5% annual change) but increases were observed at GC (cue and group: 5 and 3% annual change) and at moderately higher rates at DT (cue and group: 10 and 9% annual change).

Adult males had the lowest densities among classes, being relatively rare at all sites with mean densities of 0.007 and 0.006 animals 1000 km^{-2} at MC and DT, respectively, and marginally higher at GC (0.008 animals 1000 km^{-2}; Tables 4 & 5, Fig. 6). Similar to mid-size animals, adult male densities declined at MC (cue and group: −7 and −10% annual change) but increases were observed at GC (cue and group: 2 and 14% annual change) and at moderately higher rates at DT (cue and group: 16 and 11% annual change).

3.5. Densities with and without demographic considerations

Generalized mean densities per site were estimated for the 7 yr period using parameters treating the entire population as a single entity for both cue (Table 4) and group counting (Table 5). A comparison between generalized mean densities and inclusive mean densities, which incorporated the sum of the 3 demographic class densities, showed that the bias from ignoring demographics in the GoMex was small (0.2–2%) for cue counting, but larger (7–23%) for group counting, consistently leaning towards overestimation (Table 7). The bias was lowest at MC and highest at DT. Discrepancies in counting methods improved at all sites from 25% for generalized densities to 18% for inclusive densities at MC, 38 to 27% at GC, and 43 to 23% at DT.

4. DISCUSSION

Accurate estimates of animal densities and trends are fundamental for effective wildlife management and conservation. PAM distance sampling methods require multiple assumptions when estimating parameters to convert acoustic detections into absolute numbers of animals. We illustrate that applying generalized assumptions about animal behavior at the species or population level can introduce bias in density estimates for sperm whales. This species exhibits

Table 7. Comparison of density estimates with and without demographics differences. 'Generalized' ignores demographic differences, whereas 'inclusive' adds the estimated density sum of each demographic segment: social group, mid-size animals, and adult males. Difference estimated as the density change between methods divided by average densities of methods per site. MC: Mississippi Canyon; GC: Green Canyon; DT: Dry Tortugas

Counting method	Site	Generalized mean density (animals 1000 km^{-2})	Inclusive mean density (animals 1000 km^{-2})	Generalized-inclusive difference (%)
Cue	MC	2.319	2.314	0.2
	GC	0.868	0.866	0.3
	DT	0.379	0.372	2.0
Group	MC	2.971	2.780	6.7
	GC	1.273	1.138	11.2
	DT	0.589	0.467	23.1
Cue-group difference	MC	24.6%	18.3%	
	GC	37.8%	27.2%	
	DT	43.3%	22.7%	

sexual dimorphism in adult whales, characterized by variations in click characteristics and diving behavior, group sizes, and stratified spatial distribution, resulting in non-conformity to assumptions of equal detection probability between sex/age groups. Monitoring demographic structure in the northern GoMex using dominant ICI distributions as a proxy for size classes (Solsona-Berga et al. 2022) delineated acoustic detections of 3 demographic segments. Demographic-specific parameters were considered when estimating densities and allowed for separate analysis of demographic trends.

Misdetection and classification of sperm whales can lead to biased density estimates. Accounting for imperfect detection required estimating the proportion of false positives and negatives. False positives were manually calculated from a random subset of the 7 yr data set at each site. False positive detections, attributed to other marine mammals and noise sources, were estimated regardless of class, and so differences in false positives between demographic classes are unknown but assumed to be relatively small given manual categorization by 2 analysts. This dual-analyst approach helped identify and resolve discrepancies, enhancing classification accuracy. However, manual classification still involves some subjectivity. This underscores the need for ongoing improvements and the potential integration of automated methods in future studies to ensure consistency. Deviation between predicted and measured clicks at low amplitude provided a means to estimate percentage of false negatives per site, class, and counting method and illustrated lower sensitivity of detecting clicks from a group of whales near the amplitude detection threshold (135 dB_{pp} re: 1 µPa), particularly in mid-size animals and adult males. This reduced sensitivity likely stems from the difficulty of classifying demographics with only a few clicks. Although the number of unassigned time bins was low, those with a small number of clicks are inherently less reliable for classification. Distant and smaller groups of mid-size animals and adult males may only be detected through a few clicks, increasing the likelihood of these groups being unassigned and missed.

Cue counting involves estimating clicking rate, which varied slightly between sites for each class driven by the measured ICIs. Differences in ICI between sites were minimal for each demographic class (SD <100 ms), resulting in minor differences in average click rate estimates. Some differences could be explained by differences in the total number of clicks available for analysis, with far larger numbers of clicks at the northern sites. The proportion of time spent clicking was computed using behavioral data from acoustic tag records of females and juveniles in the GoMex (Watwood et al. 2006). Although populations in subtropical and temperate latitudes involving both sexes spend equivalent amounts of time clicking (Watwood et al. 2006, Ward et al. 2012), future studies could validate this information by placing acoustic tags on GoMex adult males.

Group counting requires an estimate of group size and synchrony. While social group size is well documented in the northern GoMex based on extensive fieldwork (Jochens et al. 2008) and synchrony was estimated from a unique instance of simultaneously tagged whales from the same group (Jochens et al. 2008), information on adult male and mid-size animal group sizes is limited. Expected small group sizes for adult males and mid-size animals were supported by an ad hoc test directly counting simultaneously received click sequences as a proxy for group size (Hildebrand et al. 2015) when animals were detected near the sensor. Future studies could validate group size estimates and minimize uncertainty by acoustically tracking several animals from these demographic segments.

Parameters unknown for certain demographics or only known for other populations were approximated with a grid-search approach (Hildebrand et al. 2019) and provided information on differences in diving behavior between sites and signal characteristics be-

tween demographics. However, while grid-search optimization tested all parameter combinations within pre-selected ranges, it was computationally intensive and may have missed complex parameter interactions, leading to convergence on local minima rather than the global optimum. This is because it evaluates parameters independently and systematically. Implementing techniques like Bayesian optimization could address these issues, capturing parameter dependencies and providing more accurate parameter optimization.

Grid-search indicated a consistent water-column foraging depth of approximately 500 m across all sites, in line with tag data diving behavior from the northern GoMex (Watwood et al. 2006). At northern sites, modeling revealed that 40–50% of foraging dives occur near the seafloor (800–1000 m), showing a bimodal diving pattern with mid-water and benthic dives, consistent with tag data collected in the northern GoMex (Mate et al. 2017). In contrast, modeling suggested virtually exclusive mid-water column foraging at the southern site with deeper bottom depth (1300 m). Tagging animals near the southern site would be valuable for validating these potential foraging differences. While more complex models could explore dive patterns based on time of day (e.g. determining if benthic dives occur at specific times) or depth of prey fields, tracking or tagging studies could address these potential foraging differences. Grid-search estimated source levels of 237 ± 5 dB_{pp} re: 1 µPa for social groups, 238 ± 5 dB_{pp} re: 1 µPa for mid-size animals, and 242 ± 5 dB_{pp} re: 1 µPa for adult males. Although source levels of females and juveniles have not been directly measured in previous studies, they are expected to be lower than those of adult males given their smaller body size (Jensen et al. 2018). Our findings align with measured source levels for adult males, with apparent source levels up to 229 dB_{peak} (Zimmer et al. 2005) and 235 dB_{rms} re: 1 µPa (Møhl et al. 2003), which, assuming a sinusoidal wave, are equivalent to 235–244 dB_{pp} re: 1 µPa. Notably, at DT, the model-derived source levels were 2 dB lower across all classes compared to the northern sites. These lower source levels were necessary to achieve a good fit between the predicted and measured received level distributions. The reason for this difference, and its relation to whales foraging exclusively at mid-water depths at DT, remains unclear. These differences may potentially be related to our model simulations assuming a uniform distribution of sperm whales around the sensor to facilitate detection probability estimation. Actual sperm whale distribution may be more complex, particularly in sites with varying bathymetry, such as the sharp slope of site DT. Grid-search estimated that directivity was consistent for mid-size animals and adult males (32.5 ± 2.5 dB) but lower for social groups (30 ± 2.5 dB). These values are slightly higher than the previously measured directivity of 27 dB for adult males (Møhl et al. 2003, Zimmer et al. 2005, Nosal & Frazer 2007). Our estimates of directivity for mid-size animals and adult males may have slightly decreased the detectability of these classes, as small changes in this input parameter resulted in a maximum 4–5% change in overall detectability predictions.

Modeling detection probabilities revealed substantial differences in the detection range of each demographic class and site. Adult males were detectable at larger distances than social groups and mid-size animals. The distance at which individual clicks of adult males could be detected was approximately 5 km farther than the other classes, and 10 km farther for the detection of a group. Differences in detectability of individual clicks of social groups and mid-size animals were marginal but more prominent when detecting a group of mid-size animals, which were detectable approximately 2–5 km farther away than social groups. Overall, individual clicks became less detectable at shorter distances (2–5 km) than for a group (20–25 km) under the same environmental conditions. This is because the probability of an individual click being on-axis is very low, but higher for the group during a 5 min window (Frasier et al. 2016). Larger detection ranges of groups were expected since on-axis clicks are detectable at relatively far ranges. Differences in detectability between the southern site (DT) and the northern sites (MC and GC) were observed, likely driven by the steep bathymetry of the southern site and our model simulations assuming a uniform distribution of sperm whales around the sensor. Detection probability models are also sensitive to assumptions about signal characteristics and diving behaviors. Source level, beam directivity, and bimodal dive fraction had the most influence on detectability predictions, with source level having the most substantial impact. Future studies would benefit from tagging or long-term acoustic tracking data to understand the non-uniform distribution of sperm whales and signal characteristics or diving behavior of the different demographics, providing more realistic simulation models and improving density estimation accuracy. In the absence of such data, grid-search optimization has provided approximate estimates despite its limitations in capturing complex parameter interactions.

4.1. Impact of demographic differences in density estimation

Ignoring population demographics biases sperm whale density estimation from passive acoustics. Larger differences between generalized and inclusive methods occurred in group counting, while small differences were found with cue counting. Variation among sites highlights non-uniform bias, with the lowest differences at MC and the highest at DT. Dominant presence of social groups at MC may explain the relatively modest bias when demographics were ignored, contrasting with the highest bias at DT with the greatest demographic variability. These findings may be important in ocean basins with less skewed demographic composition and greater variations in spatial and temporal presence, where neglecting demographic differences may lead to larger biases in densities.

Accuracy of each counting method relied on meeting assumptions and estimating parameters reliably. Large differences between cue and group counting occurred with a generalized approach. Inclusive density estimation improved agreement between counting methods, particularly at sites with larger discrepancies, attributed to a more accurate consideration of model assumptions regarding differences in detectability, click rates, and diving behavior associated with demographics.

4.2. Density and trends

We chose a weekly time scale to balance the need for fine-scale estimates and to meet our simulation assumptions about animal positions relative to the hydrophone. Weekly densities can capture relevant patterns in animal movements and population changes, which we plan to analyze alongside oceanographic variations for future habitat model predictions. We estimated a sperm whale mean density of 1.3 animals 1000 km^{-2} over the 3 sites (1.4 animals 1000 km^{-2} when not considering population demographic differences). Compared with earlier large-scale vessel-based population estimates in the northern GoMex of 1180 animals (CV = 0.22) for the 2017–2018 period, with an average density of 1.7 animals 1000 km^{-2} (Garrison et al. 2020, Hayes et al. 2022), our estimate aligns reasonably well given the limited spatial sampling with PAM recorders. Data collection at additional sites will provide better spatial coverage and improve application of these methods to the broader region.

Northern sites showed higher densities than the southern site, and different trends were observed among the 3 sites over the 7 yr period. MC in the north-central GoMex exhibited the highest densities among all sites, aligning with previous findings suggesting it is a core area for sperm whales (Jochens et al. 2008). In contrast, the southeast site, DT, had the lowest densities, potentially due to its location near the continental slope. Sperm whales have been sighted at high rates in this area, but most sightings occurred in waters of the abyssal plain with bottom depths greater than 2000 m (Garrison & Dias 2020). Social group presence at DT was most variable, consistently increasing in spring and summer (Solsona-Berga et al. 2022), potentially explained by inshore/offshore movements across deeper and shallower isobaths. The impact of the GoMex Loop Current on the West Florida continental shelf, particularly near DT, where it meets shallow isobaths, can induce upwelling-favorable motion across the entire slope (Sorinas et al. 2023) and could drive variation in sperm whale presence.

Declines in sperm whale density at the northern sites and increases at the southern site suggest movements in the GoMex region. Trends aligned between the generalized and demographic-specific methods at MC and DT but varied at GC. While the generalized method documented declines at GC, only social groups exhibited such trends, highlighting the importance of demographic consideration to capture relevant shifts in the population. Declines may be linked to elevated noise from anthropogenic activities in the north-central GoMex (Wiggins et al. 2016, Estabrook et al. 2016), where we detected the most ship passages, and the potential lasting consequences of the DWH oil spill, particularly at MC near the wellhead, which had the most pronounced decline. While the full extent of DWH damage on deepwater ecosystems is uncertain, estimates suggest a significant impact in the northeastern quarter of the deepwater GoMex (Murawski et al. 2020), including benthic contamination (Brooks et al. 2015). Approximately 16.1% of the GoMex sperm whale stock was exposed to high oil concentrations (Aichinger Dias et al. 2017). The availability of limited pre-spill ecosystem data complicates damage assessment, relying heavily on post-spill time series (Frasier et al. 2020). Population-level declines in intermediate trophic level organisms (Sutton et al. 2022), preyed on by sperm whales, may contribute to observed sperm whale declines in the northern GoMex as a result of the oil spill.

To interpret site-specific trends, a comprehensive understanding of sperm whale responses to local con-

ditions, including both natural and anthropogenic processes, is crucial. Continued monitoring, combining long-term acoustic data with monitoring of the GoMex deep ecosystem (Cook et al. 2020, Woodstock et al. 2021), can aid in identifying the drivers influencing these trends.

5. CONCLUSIONS

This study highlights the importance of accounting for demographic differences in passive acoustics for accurate sperm whale density estimates. Considering differences in detectability, click rates, group sizes, and diving behavior improved the agreement between cue and group counting methods, allowing monitoring of demographic-specific density trends. Small biases for cue counting and moderate biases for group counting resulted at sites with skewed demographic composition, particularly in the GoMex population composed predominantly of social groups. In regions with balanced sex/age group presence, ignoring demographics could introduce higher bias in density estimates. Estimated long-term density trends aligned between generalized and demographic-specific methods and indicate an increase in the southeastern site and a decline in the northern site near the oil spill wellhead. Generalized estimates indicate declines at the northwestern site, whereas demographic-specific density estimates at this site indicate that only social groups exhibited such trends. These findings warrant further research on the drivers behind these changes and highlight the importance of monitoring demographic-specific changes.

Acknowledgements. Funding for data collection and analysis was provided by the Natural Resource Damage Assessment partners (20105138), the US Marine Mammal Commission (20104755/E4061753), the Center for the Integrated Modeling and Analysis of the Gulf Ecosystem (CIMAGE) Consortium of the Gulf of Mexico Research Initiative (SA 12-10/GoMexRI-007), and NOAA's RESTORE Science Program's Grant No. NOAA-NOS-NCCOS-2019-2005608 through the Gulf Coast Restoration Trust Fund to the NOAA Southeast Fisheries Science Center. We thank Steve Murawski and Sheryl Gilbert for help and encouragement in conducting this project; Len Thomas, Tiago Marques, and Danielle Harris for assistance with statistical methods; and HARP technicians Bruce J. Thayre, John P. Hurwitz, Tim Christianson, Brent Hurley, Chris Garsha, Joshua Jones, and Ryan Griswold. Data processors and analysts included Hannah R. Bassett and Erin O'Neil. Vessels and key cruise leads who made this work possible include Tony Martinez, Dr. Jesse Wicker, Dr. Keith Mullin, and the crews of RVs 'Gordon Gunter' and 'Pisces', as well as Ocean Alliance and the RV 'Odyssey'. The present work was part of A.S.B.'s PhD thesis, available at http://hdl.handle.net/2117/129269. The scientific results and conclusions, as well as any views or opinions expressed herein, are those of the authors and do not necessarily reflect those of NOAA or the Department of Commerce.

LITERATURE CITED

Aichinger Dias L, Litz J, Garrison L, Martinez A, Barry K, Speakman T (2017) Exposure of cetaceans to petroleum products following the *Deepwater Horizon* oil spill in the Gulf of Mexico. Endang Species Res 33:119–125

Alexander A, Steel D, Hoekzema K, Mesnick SL and others (2016) What influences the worldwide genetic structure of sperm whales (*Physeter macrocephalus*)? Mol Ecol 25: 2754–2772

Best PB (1979) Social organization in sperm whales, *Physeter macrocephalus*. In: Winn HE, Olla BL (eds) Behavior of marine animals. Springer US, Boston, MA, p 227–289

Brooks GR, Larson RA, Schwing PT, Romero I and others (2015) Sedimentation pulse in the NE Gulf of Mexico following the 2010 DWH blowout. PLOS ONE 10:e0132341

Buckland ST (2006) Point-transect surveys for songbirds: robust methodologies. Auk 123:345–357

Buckland ST, Anderson DR, Burnham KP, Laake JL, Borchers DL, Thomas L (2001) Introduction to distance sampling: estimating abundance of biological populations. Oxford University Press, Oxford

Cantor M, Gero S, Whitehead H, Rendell L (2019) Sperm whale: the largest toothed creature on Earth. In: Würsig B (ed) Ethology and behavioral ecology of odontocetes. Springer International Publishing, Cham, p 261–280

Cleveland RB, Cleveland WS, McRae JE, Terpenning I (1990) STL: a seasonal-trend decomposition procedure based on loess. J Off Stat 6:3–73

Cook AB, Bernard AM, Boswell KM, Bracken-Grissom H and others (2020) A multidisciplinary approach to investigate deep-pelagic ecosystem dynamics in the Gulf of Mexico following *Deepwater Horizon*. Front Mar Sci 7:548880

Douglas LA, Dawson SM, Jaquet N (2005) Click rates and silences of sperm whales at Kaikoura, New Zealand. J Acoust Soc Am 118:523–529

Estabrook BJ, Ponirakis DW, Clark CW, Rice AN (2016) Widespread spatial and temporal extent of anthropogenic noise across the northeastern Gulf of Mexico shelf ecosystem. Endang Species Res 30:267–282

Farmer NA, Baker K, Zeddies DG, Denes SL and others (2018) Population consequences of disturbance by offshore oil and gas activity for endangered sperm whales (*Physeter macrocephalus*). Biol Conserv 227:189–204

Frasier KE, Wiggins SM, Harris D, Marques TA, Thomas L, Hildebrand JA (2016) Delphinid echolocation click detection probability on near-seafloor sensors. J Acoust Soc Am 140:1918–1930

Frasier KE, Solsona-Berga A, Stokes L, Hildebrand JA (2020) Impacts of the Deepwater Horizon oil spill on marine mammals and sea turtles. In: Murawski SA, Ainsworth CH, Gilbert S, Hollander DJ, Paris CB, Schlüter M, Wetzel DL (eds) Deep oil spills: facts, fate, and effects. Springer International Publishing, Cham, p 431–462

Garrison LP, Dias LA (2020) Distribution and abundance of cetaceans in the northern Gulf of Mexico. NOAA Tech Memo NMFS-SEFSC-747

Garrison LP, Ortega-Ortiz J, Rappuci G (2020) Abundance of marine mammals in waters of the U.S. Gulf of Mexico

during the summers of 2017 and 2018. Southeast Fish Sci Cent Ref Doc PRBD 2020-07
Gaskin DE (1970) Composition of schools of sperm whales *Physeter catodon* Linn. east of New Zealand. NZ J Mar Freshw Res 4:456–471
Gero S, Milligan M, Rinaldi C, Francis P and others (2014) Behavior and social structure of the sperm whales of Dominica, West Indies. Mar Mamm Sci 30:905–922
Gillespie D (1997) An acoustic survey for sperm whales in the Southern Ocean sanctuary conducted from the RSV Aurora Australis. Rep Int Whal Comm 47:897–907
Goold JC, Jones SE (1995) Time and frequency domain characteristics of sperm whale clicks. J Acoust Soc Am 98: 1279–1291
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
Hayes S, Josephson E, Maze-Foley K, Rosel PE, Wallace J (2022) U.S. Atlantic and Gulf of Mexico marine mammal stock assessments 2021. NOAA Tech Memo NMFS-NE-288:380
Hildebrand JA, Baumann-Pickering S, Frasier KE, Trickey JS and others (2015) Passive acoustic monitoring of beaked whale densities in the Gulf of Mexico. Sci Rep 5:16343
Hildebrand JA, Frasier KE, Baumann-Pickering S, Wiggins SM and others (2019) Assessing seasonality and density from passive acoustic monitoring of signals presumed to be from pygmy and dwarf sperm whales in the Gulf of Mexico. Front Mar Sci 6:66
Irvine L, Palacios DM, Urbán J, Mate B (2017) Sperm whale dive behavior characteristics derived from intermediate-duration archival tag data. Ecol Evol 7:7822–7837
Jaquet N, Gendron D (2009) The social organization of sperm whales in the Gulf of California and comparisons with other populations. J Mar Biol Assoc UK 89: 975–983
Jensen FH, Johnson M, Ladegaard M, Wisniewska DM, Madsen PT (2018) Narrow acoustic field of view drives frequency scaling in toothed whale biosonar. Curr Biol 28:3878–3885.e3
Jochens A, Biggs D, Benoit-Bird K, Engelhaupt D and others (2008) Sperm whale seismic study in the Gulf of Mexico: synthesis report. OCS Study MMS 2008-006. US Department of the Interior, Minerals Management Service, Gulf of Mexico OCS Region, New Orleans, LA
Kobara S (2019) Deepwater Horizon Oil spill dataset compiled for Gulf of Mexico Ocean Observation (GCOOS). https://gisdata.gcoos.org/datasets/afafd2255f9d43bd8a5531de7e98c9a5/explore?layer=1&location=26.534849%2C-88.258099%2C6.26
Leaper R, Scheidat M (1998) An acoustic survey for cetaceans in the Southern Ocean Sanctuary conducted from the German government research vessel Polarstern. Rep Int Whal Comm 48:431–437
Levy JK, Gopalakrishnan C (2010) Promoting ecological sustainability and community resilience in the US Gulf Coast after the 2010 *Deepwater Horizon* oil spill. J Nat Resour Policy Res 2:297–315
Lyrholm T, Leimar O, Johanneson B, Gyllensten U (1999) Sex-biased dispersal in sperm whales: contrasting mitochondrial and nuclear genetic structure of global populations. Proc R Soc B 266:347–354
Madsen P, Wahlberg M, Møhl B (2002a) Male sperm whale (*Physeter macrocephalus*) acoustics in a high-latitude habitat: implications for echolocation and communication. Behav Ecol Sociobiol 53:31–41
Madsen PT, Payne R, Kristiansen NU, Wahlberg M, Kerr I, Møhl B (2002b) Sperm whale sound production studied with ultrasound time/depth-recording tags. J Exp Biol 205:1899–1906
Marques TA, Thomas L, Ward J, DiMarzio N, Tyack PL (2009) Estimating cetacean population density using fixed passive acoustic sensors: an example with Blainville's beaked whales. J Acoust Soc Am 125:1982–1994
Marques TA, Thomas L, Martin SW, Mellinger DK and others (2013) Estimating animal population density using passive acoustics. Biol Rev Camb Philos Soc 88:287–309
Mate BR, Irvine LM, Palacios DM (2017) The development of an intermediate-duration tag to characterize the diving behavior of large whales. Ecol Evol 7:585–595
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
Mountain DC, Anderson D, Brughera A, Cross M and others (2012) The ESME Workbench: simulating the impact of anthropogenic sound on marine mammals. In: Popper AN, Hawkins A (eds) The effects of noise on aquatic life. Springer, New York, NY, p 217–219
Murawski SA, Ainsworth CH, Gilbert S, Hollander DJ, Paris CB, Schlüter M, Wetzel DL (eds) (2020) Scenarios and responses to future deep oil spills: fighting the next war. Springer International Publishing, Cham
National Academies of Sciences, Engineering, and Medicine (2017) Approaches to understanding the cumulative effects of stressors on marine mammals. National Academies Press, Washington, DC
Nosal EM, Frazer LN (2007) Sperm whale three-dimensional track, swim orientation, beam pattern, and click levels observed on bottom-mounted hydrophones. J Acoust Soc Am 122:1969–1978
Ohsumi S (1971) Some investigations on the school structure of sperm whale. Sci Rep Whales Res Inst Tokyo 23:1–25
Porter MB, Bucker HP (1987) Gaussian beam tracing for computing ocean acoustic fields. J Acoust Soc Am 82: 1349–1359
Posdaljian N, Solsona-Berga A, Hildebrand JA, Soderstjerna C, Wiggins SM, Lenssen K, Baumann-Pickering S (2024) Sperm whale demographics in the Gulf of Alaska and Bering Sea/Aleutian Islands: an overlooked female habitat. PLOS ONE 19:e0285068
Ramseur JL (2010) Deepwater Horizon oil spill: the fate of the oil. Cong Res Serv R41531. https://crsreports.congress.gov/product/pdf/R/R41531
Reeves RR, Lund JN, Smith TD, Josephson EA (2011) Insights from whaling logbooks on whales, dolphins, and whaling in the Gulf of Mexico. Gulf Mex Sci 29:41–67
Rice DW (1989) Sperm whale *Physeter macrocephalus* Linnaeus. In: Ridgeway SH, Harrison R (eds) Handbook of marine mammals. Academic Press, San Diego, CA, p 177–233
Seber GAF (1982) The estimation of animal abundance and related parameters, 2nd edn. Charles Griffin, London
Sen PK (1968) Estimates of the regression coefficient based on Kendall's tau. J Am Stat Assoc 63:1379–1389
Solsona-Berga A, Frasier KE, Baumann-Pickering S, Wiggins SM, Hildebrand JA (2020) DetEdit: a graphical user inter-

face for annotating and editing events detected in long-term acoustic monitoring data. PLOS Comput Biol 16: e1007598

Solsona-Berga A, Posdaljian N, Hildebrand JA, Baumann-Pickering S (2022) Echolocation repetition rate as a proxy to monitor population structure and dynamics of sperm whales. Remote Sens Ecol Conserv 8:827–840

Sorinas L, Weisberg RH, Liu Y, Law J (2023) Ocean–atmosphere heat exchange seasonal cycle on the West Florida Shelf derived from long term moored data. Deep Sea Res II 212:105341

Sousa A, Alves F, Dinis A, Bentz J, Cruz MJ, Nunes JP (2019) How vulnerable are cetaceans to climate change? Developing and testing a new index. Ecol Indic 98:9–18

Sutton TT, Milligan RJ, Daly K, Boswell KM and others (2022) The open-ocean Gulf of Mexico after *Deepwater Horizon*: synthesis of a decade of research. Front Mar Sci 9:753391

Taylor BL, Baird R, Barlow J, Dawson SM, Ford J, Mead JG, Notarbartolo di Sciara G, Wade P, Pitman RL (2019) *Physeter macrocephalus* (amended version of 2008 assessment). The IUCN Red List of Threatened Species 2019: e.T41755A160983555. https://dx.doi.org/10.2305/IUCN.UK.2008.RLTS.T41755A160983555.en. Accessed 25 September 2024

Teloni V, Mark JP, Patrick MJO, Peter MT (2008) Shallow food for deep divers: dynamic foraging behavior of male sperm whales in a high latitude habitat. J Exp Mar Biol Ecol 354:119–131

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

Townsend CH (1935) The distribution of certain whales as shown by the logbook records of American whaleships. Zoologica (NY) 19:3–50

von Benda-Beckmann AM, Thomas L, Tyack PL, Ainslie MA (2018) Modelling the broadband propagation of marine mammal echolocation clicks for click-based population density estimates. J Acoust Soc Am 143:954–967

Ward JA, Thomas L, Jarvis S, DiMarzio N and others (2012) Passive acoustic density estimation of sperm whales in the Tongue of the Ocean, Bahamas. Mar Mammal Sci 28: E444–E455

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

Westell A, Rowell TJ, Posdaljian N, Solsona-Berga A, Van Parijs SM, DeAngelis AI (2024) Acoustic presence and demographics of sperm whales (*Physeter macrocephalus*) off southern New England and near a US offshore wind energy area. ICES J Mar Sci (in press)

Whitehead H (2003) Sperm whales: social evolution in the ocean. University of Chicago Press, Chicago, IL

Whitehead H, Shin M (2022) Current global population size, post-whaling trend and historical trajectory of sperm whales. Sci Rep 12:19468

Wiggins SM, Hildebrand JA (2007) High-frequency Acoustic Recording Package (HARP) for broad-band, long-term marine mammal monitoring. International Symposium on Underwater Technology and Workshop on Scientific Use of Submarine Cables and Related Technologies. University of Tokyo, p 551–557

Wiggins SM, Hall JM, Thayre BJ, Hildebrand JA (2016) Gulf of Mexico low-frequency ocean soundscape impacted by airguns. J Acoust Soc Am 140:176–183

Woodstock MS, Sutton TT, Frank T, Zhang Y (2021) An early warning sign: trophic structure changes in the oceanic Gulf of Mexico from 2011—2018. Ecol Modell 445:109509

Zimmer WMX, Johnson MP, D'Amico A, Tyack PL (2003) Combining data from a multisensor tag and passive sonar to determine the diving behavior of a sperm whale (*Physeter macrocephalus*). IEEE J Oceanic Eng 28:13–28

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

Editorial responsibility: Craig Radford, Warkworth, New Zealand
Reviewed by: M. Ferrari, P. Cauchy, B. S. Miller

Submitted: January 3, 2024
Accepted: August 29, 2024
Proofs received from author(s): September 26, 2024