Marine Pollution Bulletin 214 (2025) 117742

Contents lists available at ScienceDirect

Marine Pollution Bulletin

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

Distance matters to sperm whales: Behavioural disturbance in response to both sonar received level and source distance

Paul J. Wensveen [a,b,*], Saana Isojunno [c], Petter H. Kvadsheim [d], Frans-Peter A. Lam [e], Charlotte Curé [f], Alexander M. von Benda-Beckmann [e], Patrick J.O. Miller [c]

[a] Westman Islands Research Centre, Institute of Research Centres, University of Iceland, Ægisgata 2, 900 Vestmannaeyjar, Iceland
[b] Faculty of Life and Environmental Sciences, School of Engineering and Natural Sciences, University of Iceland, Sturlugata 7, 102 Reykjavík, Iceland
[c] Sea Mammal Research Unit, Scottish Oceans Institute, School of Biology, University of St Andrews, St Andrews, Fife KY16 8LB, Scotland, United Kingdom
[d] Sensor and Surveillance Systems, Norwegian Defence Research Establishment (FFI), 3191 Horten, Norway
[e] Acoustic Sensor & Sonar Systems Group, Netherlands Institute for Applied Scientific Research (TNO), Oude Waalsdorperweg 63, 2597 AK, The Hague, the Netherlands
[f] Cerema-University Gustave Eiffel, UMRAE, Strasbourg F-67035, France

ARTICLE INFO

Keywords:
Anthropogenic noise
Naval sonar
DTAG
Biologging
Deep diving odontocete
Behavioural response
Noise pollution

ABSTRACT

Understanding the main drivers of behavioural disturbance in deep-diving cetaceans would improve predictions of anthropogenic noise effects on individual animals, habitats and populations. To investigate the potential roles of received level and source distance on behavioural disturbance, we tagged 14 sperm whales in northern Norway with multi-sensor data loggers and conducted dose-escalation experiments. Each experiment included 1 to 4 tagged individuals and involved multiple vessel passes ('exposure sessions', $n = 25$ total) by a navy frigate or research vessel towing a naval sonar, at different starting distances and maximum source levels. We analysed behaviour state series and proxies for locomotor activity and foraging success with generalized additive mixed models. The probability of occurrence of non-foraging active behaviour was affected by received level, source distance and session order, with decreased foraging effort at higher received levels and shorter distances, and during subsequent sessions (indicating short-term sensitisation). Prey capture attempts decreased with increasing received level when whales kept foraging. Similar to what has been suggested for some populations of blue whales and beaked whales regularly exposed to navy sonar, but unlike northern bottlenose whales in more pristine waters, source distance affected sperm whale behavioural responses on a high-latitude foraging ground.

1. Introduction

Marine animals face a myriad of anthropogenic stressors in today's oceans (Tyack et al., 2022). Cetaceans are challenging to study due to their fully aquatic lifestyles and therefore effects of stressors, be it from changes in anthropogenic or natural sources, can easily go unnoticed. Anthropogenic noise is one stressor that has raised the interest and concern of the scientific and regulatory communities (Merchant et al., 2022; Southall et al., 2019a). Key factors are the importance of sound to marine taxa (e.g., (Miller et al., 2022; Kunc et al., 2016)) and the pervasiveness of noise in the marine environment due to the increasing human footprint coupled with efficient physical propagation.

The sperm whale (*Physeter macrocephalus*), globally designated as "vulnerable" on the IUCN Red List (Taylor et al., 2019), is a long-living species that was heavily hunted in the 18th–20th century. Large-scale commercial whaling ceased around 1987, by which time its global population size had declined by roughly two-thirds (Whitehead and Shin, 2022). Sperm whales are a large deep-sea predator with a circumglobal distribution that feed mainly on cephalopods, in addition to fish, beneath the photic zone (Whitehead, 2003; Santos et al., 1999; Similä et al., 2022). They are thought to play an important role in bringing nutrients to the surface and enhance primary production, especially in nutrient-limited waters (Lavery et al., 2010; Roman et al., 2014). While female sperm whales tend to stay in groups at lower latitudes, the larger males make long-distance migrations to high-latitude foraging grounds where they are sighted mostly solitary (Whitehead, 2003). In these habitats, sperm whales overlap with areas with regular naval sonar operations and other noise-generating activities. For example, sperm whales in the Nordic Seas and North Atlantic are often found in slope waters near two critical naval chokepoints, the GIUK Gap

---

* Corresponding author at: Westman Islands Research Centre, Institute of Research Centres, University of Iceland, Ægisgata 2, 900 Vestmannaeyjar, Iceland.
E-mail address: pjw@hi.is (P.J. Wensveen).

https://doi.org/10.1016/j.marpolbul.2025.117742
Received 29 November 2024; Received in revised form 21 February 2025; Accepted 22 February 2025

and Bear Island Gap, which are important focus points for anti-submarine warfare exercises (Ramirez-Martinez et al., 2024; Depledge, 2020; Åtland et al., 2022). Given their vulnerable status and potentially important mesopredator roles in marine ecosystems, such as the nutrient turnover, it is important to understand the risks of multiple stressors, including underwater noise.

As an echolocating toothed whale, sperm whales are highly sound-dependent creatures that produce clicks for foraging, finding mates, navigation, and social cohesion (Miller et al., 2004; Hersh et al., 2022; Fais et al., 2015). Sperm whales may be affected by anthropogenic noise pollution through the mechanisms of masking, distraction or misleading (Dominoni et al., 2020). Some studies have suggested high sensitivity to noise disturbance (acoustic tomography, (Bowles et al., 1994); vessel passes, (Azzara et al., 2013); naval sonar, (Isojunno et al., 2016)), whilst others have found no or minor effects (seismic surveys, (Madsen et al., 2002; Miller et al., 2009)). Specific reasons for these differences may vary, as the disturbance responses of wildlife can depend on many context variables relating to the exposure as well as intrinsic and extrinsic factors (Tablado and Jenni, 2017; Booth et al., 2022).

Since the 1990s, large-scale naval exercises have been linked to strandings of cetaceans, particularly beaked whales (Cox et al., 2006; Bernaldo de Quirós et al., 2019), which led to increased research effort on sonar effects (e.g., (Hooker et al., 2012; Southall et al., 2016; Harris et al., 2018)). Studies have also targeted marine mammals other than beaked whales, as sub-lethal effects may have population-level consequences when repeated often and over long periods (Pirotta et al., 2018; Farmer et al., 2018). Evidence on the scale of disturbance effects of naval exercises has started to emerge for species including goose-beaked whales (*Ziphius cavirostris*), Blainville's beaked whales (*Mesoplodon densirostris*), and sperm whales (Stanistreet et al., 2022; Tyack et al., 2011; Joyce et al., 2019). Behavioural response studies using experimental approaches have associated features of the acoustic signal, such as source level or frequency, and exposure context to an increasing array of response metrics (e.g., (Miller et al., 2012; Wensveen et al., 2019; Southall et al., 2019b; Wensveen et al., 2017)). Such experimental studies have found that male sperm whales at high latitudes are medium-sensitive compared with other cetaceans (Miller et al., 2022; Harris et al., 2015). After controlling for several context variables, one study (Isojunno et al., 2020) observed minor but matching foraging reductions between high-power pulsed active sonar and medium-power continuous active sonar in this area, leading them to conclude that received sound exposure level (SEL) better predicted those responses than received sound pressure level. Non-controlled, incidental sonar exposures were also reported by that study (Isojunno et al., 2020) and an earlier study on sperm whales in northern Norway (Isojunno et al., 2016), but neither observed responses to those incidental signals.

Behavioural response studies on goose-beaked whales and blue whales (*Balaenoptera musculus*) in areas with regular naval activity suggested that received level and/or distance mediated behavioural responsiveness (Southall et al., 2016; Southall et al., 2019b; DeRuiter et al., 2013; Falcone et al., 2017), but these studies were not designed to distinguish between the two potential drivers. Responses of northern bottlenose whales (*Hyperoodon ampullatus*), also a beaked whale, in more pristine Arctic waters were associated with received level and not distance (Wensveen et al., 2019). This lack of a 'distance effect' could be due to a higher perceived level of risk being associated with unfamiliar or relatively novel signals, a phenomenon widespread across taxa (Crane and Ferrari, 2017), although further data on *Hyperoodon ampullatus* from areas with more regular sonar use or on other beaked whales from more acoustically pristine areas are needed to test this hypothesis. For example, Arctic cetaceans including beluga (*Delphinapterus leucas*) and narwhal (*Monodon monoceros*) generally exhibit further and stronger responses to anthropogenic noise in their relatively pristine habitats compared with cetaceans in areas with more human activities (e.g., (Heide-jørgensen et al., 2021; Williams et al., 2022; Martin et al., 2022)).

This study aimed to test specific hypotheses about what exposure factors can influence behavioural disturbance from naval sonar in male sperm whales in northern Norway: acoustic received level, source distance, or an interaction of both. We conducted dose-escalation experiments during which individuals were exposed to received levels previously associated with behavioural responses in this species, but at a wider range of source distances than in previous experiments. Physical spreading of sound away from the source results in an intrinsic relationship between acoustic received level and distance if source level were to be kept constant. An important aim of our study design was therefore to use a set of experimental procedures to break up this inherent ambiguity by exposing subject whale to: 1) *close* and *distant* vessel pass starting distances at identical transmission source levels, 2) ramp-up protocols that stepwise-increased the source level during the first half of the session, and 3) three different maximum source levels, including one that could only be produced by an operational naval frigate, to expose whales to similar received levels at different distances. We aimed to tag multiple individuals per experiment to increase data coverage at the longest distances and to reduce the total number of sonar transmissions required in the study area.

2. Material and methods

2.1. Data collection

The field methods are described in detail in (Kvadsheim et al., 2020; Kvadsheim et al., 2021). Data were collected in August and September of 2019 in waters off Andenes, northern Norway (Table 1). Whales were located visually by naked eye or binoculars, or acoustically by monitoring their echolocation clicks using a towed hydrophone array from the 55-m research vessel H.U. Sverdrup II. A water jet propulsion Man Over Board boat (the 'tag boat') was launched from the research vessel when sperm whales were in the general area. Before each experiment, we deployed suction cup-attached data loggers on individuals using a 15-m pole that was cantilever-mounted to the bow of the tag boat. We targeted for tagging any individual that surfaced sufficiently close to the tag boat. The tag boat team used a directional hydrophone, as well as information from the visual and acoustic observers on the main research vessel relayed over VHF radio, to position the tag boat as close as possible to a surfacing whale. The data loggers known as mixed-DTAGs (Kleivane et al., 2022) consisted of a DTAG multi-sensor core unit (version 3; University of Michigan), a snapshot GPS logger (F5G 234C, Lotek Wireless Inc.), VHF transmitter (F1835B, Advanced Telemetry Systems Inc.) and Argos transmitter (SPOT-386, Wildlife Computers Inc.) placed inside a robust purpose-built housing. The DTAG core unit recorded continuous sound at 120 kHz using two hydrophones (16 bits and 12 dB or 0 dB gain per channel), as well as three-axial acceleration, three-axial magnetic field, and pressure at between 250 Hz and 625 Hz (decimated to 50 Hz). When the tag was dry, one Argos transmission was sent every 45 s (no hourly limit; transmission delay 0.5 s) and one GPS snapshot was logged every 30 s (the highest rate that could be set). VHF signals were 25 ms long and transmitted twice per second in the 215–221 MHz band. Once multiple animals were tagged, one of them was designated as the focal individual for tracking. The focal whale was tracked visually and by VHF until the tag released from the animal at a pre-programmed time.

Table 1

Summary of the collected data. Duration refers to the analysed data and excludes tagging periods. Each experiment was conducted with the research vessel (RV) or naval frigate vessel (FV) as the source vessel. Exposure sessions are listed in order of presentation and include only those that could be analysed. PAS, pulsed active sonar; MPAS-C, medium-level PAS-*close*; HPAS-D, high-level PAS-*distant*; HPAS-C, high-level PAS-*close*; XHPAS-D, extra high-level PAS-*distant*; XHPAS-C, extra high-level PAS-*close*; NS, no-sonar control; *BF_2h*, blackfish detected within the last 2 h; *UPAS_2h*, unidentified pulsed active sonar detected within the last 2 h; *TS*, test sonar pulses from the frigate (number of 1-min time-bins). Sex and age of the tagged whales are unknown, but they were most likely physically mature males based on their size and the location of research (Whitehead, 2003).

Deployment id	Focal whale	Tag on (UTC)	Duration (h)	Experiment id	Exposure sessions	*BF_2h* (h)	*UPAS_2h* (h)	*TS* (# bins)
sw19_241a	Y	29 August 08:32	21.0	1 (RV)	HPAS-C, HPAS-D, MPAS-C	9.7	0.0	5
sw19_241b	–	29 August 09:43	19.7	1 (RV)	HPAS-C, HPAS-D, MPAS-C	6.8	0.0	6
sw19_243a	Y	31 August 07:32	20.5	2 (RV)	HPAS-D, HPAS-C	17.0	0.0	0
sw19_244a	–	01 September 19:20	31.5	3 (FV)	XHPAS-D, XHPAS-C	21.1	2.0	10
sw19_245a	Y	02 September 19:22	12.5	3 (FV)	XHPAS-D, XHPAS-C	3.0	0.0	10
sw19_248b	Y	05 September 06:23	24.1	4 (RV)	NS, HPAS-D, MPAS-C, HPAS-C	12.9	4.1	0
sw19_250a	Y	07 September 11:32	18.5	5 (FV)	XHPAS-C, XHPAS-D	14.7	6.5	6
sw19_253bc[1]	Y	10 September 08:07	25.7	6 (FV)	NS, XHPAS-D, XHPAS-C	18.3	0.0	21
sw19_254a	Y	11 September 15:04	20.7	7 (FV)	XHPAS-C, HPAS-C, XHPAS-D	1.3	0.0	0
sw19_255b	Y	12 September 07:11	17.3	8 (FV)	XHPAS-D, XHPAS-C, NS	14.9	6.8	4
sw19_255c	–	12 September 08:03	8.2	8 (FV)	XHPAS-D	8.1	1.2	0
sw19_255d	–	12 September 08:34	3.9	8 (FV)	[2]	2.4	0.0	0
sw19_259a	–	16 September 09:25	4.0	9 (FV)		4.0	0.0	0
sw19_259b	Y	16 September 10:28	21.6	9 (FV)	NS, XHPAS-C, XHPAS-D	2.2	0.0	0

[1] Tag sw19_253c was redeployed on the same individual after it prematurely released during the no-sonar session.

[2] Tag sw19_255d suffered a depth sensor failure during the baseline period and therefore no sonar period data were available.

2.2. *Experimental design*

Tagging was followed by a baseline pre-exposure phase for contrasting animal behaviour during sonar exposure. This baseline phase started once the research vessel had recovered the tag boat (Isojunno and Miller, 2015) and lasted a minimum of 4 h. The subsequent experimental phase consisted of up to four 40-min vessel passes ('exposure sessions'; Table 1) to the focal whale, with a minimum of 1 h 20 min between sessions. The source vessel during the exposure sessions was either the research vessel towing the SOCRATES source (TNO, the Netherlands) or the 134-m Norwegian Navy frigate KNM Otto Sverdrup towing a CAPTAS-Mk2 source (Thales Group, France). Each sonar source was towed ~485 m behind the source vessel at 120 m depth. The aim was to start the session with the vessel at either 7.4 km (*close*) or 14.8 km (*distant*) from the focal whale and to approach the whale at an angle of ~45° to its expected path. Source vessel course and speed (4.1 m s^{-1}) were kept constant during each exposure session. The research vessel aimed to sail 4–6 km wide square boxes around the focal whale to facilitate visual, VHF, and acoustic tracking using target motion analysis. This movement pattern was changed only when the research vessel, and not the frigate, conducted the vessel passes to the focal whale. Visual tracking was typically lost between, and in the beginning and end of, such vessel passes with the research vessel and the focal whale was tracked by VHF during these periods.

The source was deployed but not transmitting during no-sonar control sessions (always *close*; Table 1). Five types of exposure sessions with pulsed active sonar (PAS) transmission were presented in rotating order: the extra high-level PAS-*close* (XHPAS-C) or *distant* (XHPAS-D) with maximum source level of ≥220 dB re 1 μPa^2 m^2, high-level PAS-*close* (HPAS-C) or *distant* (HPAS-D) with maximum source level of 214 dB re 1 μPa^2 m^2, and medium-level PAS-*close* (MPAS-C) with maximum source level of 206 dB re 1 μPa^2 m^2. XHPAS and MPAS exposures were only conducted by the frigate and research vessel, respectively. These different types of exposure sessions were used to maximise data coverage across the bivariate space of received level and source distance. The study design was not factorial, so each combination of session type and presentation order was not considered a separate treatment. Each sonar pulse was a 1-s duration hyperbolic frequency upsweep from 1.28 to 1.92 kHz, transmitted once every 20 s (SOCRATES) or 21–24 s (CAPTAS). Source level was increased stepwise over the first 20 min of the session according to a pre-defined MPAS, HPAS or XHPAS ramp-up protocol (Kvadsheim et al., 2020; Kvadsheim et al., 2021).

To reduce the chance of exposing the same individuals to significant levels in multiple experiments, sonar exposure experiments were conducted at least 37 (HPAS) and 56 km (XHPAS) away from previous experiment locations within 24 h. Photo-identification was used to ensure that the same individuals were not tagged more than once. A 100 or 200 m mitigation zone around the source, depending on transmitted source level, was monitored visually during exposure sessions. Visual observation was aided by infrared night vision equipment during low-light conditions. If any marine mammals entered the mitigation zone, active transmission would be ceased immediately. The experiments were conducted under permits from the Norwegian Animal Research Authority (permit no 18/126201), and experimental procedures were approved by the Animal Welfare Ethics Committee of the University of St Andrews.

2.3. *Data processing*

Standard procedures were used to convert DTAG-recorded pressure to depth and derive body orientation (pitch, roll, heading) from tri-axial acceleration and magnetic field strength (Johnson and Tyack, 2003). Fluke strokes were detected using an automated detector based upon cyclic variations in pitch (Tyack et al., 2006), with detection parameters determined manually for each tag record by inspecting the magnitude of the fluke stroke signals within the pitch series. An important input to the detector was the fundamental stroke frequency of the whale, which varied between 0.11 and 0.16 Hz across the tag deployments. The geographical position of the whale throughout the tag record was estimated at 1 Hz from its body orientation, forward speed (based on flow noise in the 66–94 Hz band), and visual and GPS position fixes using a state-space model fitted in a Bayesian framework (Wensveen et al., 2015). The horizontal source distance was determined using the estimated whale track and an approximate source track for each sonar transmission. Based on the assumption that the source closely followed the trail of the ship, the source track was approximated by adding 120 s (the time for the source to cover ~485 m at 4.1 m s^{-1}) to the times of the vessel's GPS track.

We followed the data-processing and statistical procedures described in (Isojunno et al., 2020), with any differences noted below. Echolocation click trains of the tagged animal were manually identified by inspecting acoustic data (96 kHz) from the tag, both aurally and visually using spectrograms. The start of each terminal echolocation 'buzz' (i.e., rapid increase in click rate indicative of a prey capture attempt; (Miller

Table 2
List of data variables used in statistical models.

Variable	Description
Response:	
pNFA	Posterior probability of the non-foraging active state
buzz	Buzz start presence/absence: 1 when a buzz start was present in the 1-min time-bin, 0 otherwise
fluke	Fluke stroke rate: number of strokes per minute
Predictor:	
state	Most likely categorical behaviour state (only in *fluke* models; Isojunno et al., 2020).
solarnoon	Time since solar noon (h) at the start of each 1-min time-bin
depth	Whale depth (m) at the start of the 1-min time-bin (only in *buzz* models; Miller et al., 2004)
depth_bathy	Bathymetric depth (m) at the start of the tag deployment
BF_2h	Blackfish (pilot or killer whale): 1 when detected acoustically or visually in the last 2 h, 0 otherwise
UPAS_2h	Unidentified pulsed active sonar: 1 when detected acoustically in the last 2 h, 0 otherwise
TS	Test sonar, i.e., incidental sonar exposure: 1 when test pulses were transmitted by the frigate, 0 otherwise
NS	No-sonar: 1 during no-sonar control sessions, 0 otherwise
SELmax	Maximum single-pulse sound exposure level (dB re 1 μPa^2 s) since sonar exposure start, for each 1-min time-bin
DISTmin	Minimum source distance (km) since sonar exposure start, for each 1-min time-bin
DISTmin_prev	Minimum source distance (km) of previous sonar exposure sessions, during exposure (i.e., order effect)
DISTmin_post	Minimum source distance (km) of sonar exposure session, during post-exposure
BF_2h_dur	Same as predictor variable *BF_2h*, but 0 outside sonar exposure sessions

et al., 2004)) was judged from both change in amplitude and inter-click interval, relative to regular click trains. The presence of sounds produced by killer whales, *Orcinus orca*, or long-finned pilot whales, *Globicephala melas*, (hereafter, "blackfish") and the presence of distant sonar signals that were not transmitted as part of the experiments were also annotated. Visual sightings of blackfish species, automatic detections on the acoustic array, and CAPTAS transmission logs also informed the presence/absence variables for blackfish species, incidental sonar, and unidentified sonar (Table 2).

Received SEL (dB re 1 μPa^2 s) was measured for each sonar pulse identified in the acoustic recordings using a semi-automated procedure similar as in (Miller et al., 2012; Miller et al., 2011). The approach used a matched filter to identify the start of the first arrival of the sonar pulses. Acoustic pressures were measured in three frequency bands (1.12–1.41 kHz, 1.41–1.78 kHz and 1.78–2.24 kHz) that overlapped with the transmitted signal by applying 6th-order Butterworth bandpass filters to the unfiltered waveform. Pulse duration was defined as the time during which the time-weighted (10-ms) sound pressure level exceeded a threshold of 20 dB below its maximum (in a few cases this threshold was not reached and were the pulse start/end times marked manually). Received SEL was integrated across the three frequency bands and over the duration of the single pulse. We favoured single-pulse SEL over sound pressure level as the received level metric predicting behavioural responses in sperm whales (Isojunno et al., 2020) but note that these metrics are within a few dB for 1-s duration signals.

Each tag deployment was assigned behaviour states at 1-min resolution based on acoustic and movement parameters using a Bayesian state-switching model (Isojunno and Miller, 2015). This unsupervised classification method estimated the posterior probability of being in one of six states: layer-restricted search behaviour, descending to a deeper depth, ascending to a shallower depth, surfacing/logging at the surface, resting at and below the surface, and a non-foraging active state. The non-foraging active state can be considered a potential 'response state', because sperm whales in this area are mostly either diving and foraging, or resting (Isojunno et al., 2020). To maximise the amount of data informing the behaviour state classification, the state classification model was fitted to the non-exposure (baseline) control periods from 28 previous tag records (Isojunno et al., 2016; Isojunno et al., 2020) in addition to the 14 new tag records that included sonar exposure (Table 1).

2.4. Statistical analysis

Generalized additive mixed-effects models (GAMMs) were fitted for three response variables of interest, each modelled at the resolution of the 1-min time-bins, with tag id specified as a random effect to account for repeated measures of individuals. The analysis was performed in the statistical computing software R using the mgcv package (version 1.8.40; (Wood, 2008)). Variable names in the text are italicized for clarity. The three response variables were: *pNFA*, the posterior mean probability of the tagged whale being in the non-foraging active state associated with disturbance; *buzz*, the presence/absence of one or more buzz starts, a proxy for prey capture attempts (Miller et al., 2004); and *fluke*, the fluke stroke rate, a proxy of locomotor activity. These variables were carefully selected to avoid, as much as possible, multiple testing of the same behavioural response (Isojunno et al., 2020). Specifically, models for *buzz* only included data from the foraging states (descent, layer-restricted search, ascent). Models for *fluke* were controlled for behaviour state effects and did not include data from the surface state for which fluke rate measurements can be unreliable. Within-individual temporal residual autocorrelation was reduced to acceptable levels using an ARMA(1,1) structure for *pNFA* and *buzz* and AR(1) structure for *fluke* (Harrison, 2021). The ARMA parameters were estimated for the full models and fixed for reduced models to decrease computational time (Harrison, 2021). Response variables *pNFA* and *buzz* were modelled as binomial variables with logit link, whereas *fluke* was specified a Poisson distribution with log link (all with "quasi" option).

Candidate predictors included baseline covariates such as time since solar noon (*solarnoon*) and exposure covariates such as the maximum single-pulse sound exposure level (*SELmax*) and minimum source distance (*DISTmin*) since the start of the exposure session (see Table 2 for all variables and descriptions). We included the interaction between *SELmax* and *DISTmin*, as well as terms for session order and post-exposure that were defined by the previous *DISTmin* (Table 2). *SELmax* was transformed before model fitting by subtracting 60 dB re 1 μPa^2 s (a value below its minimum observed value). The *DISTmin* covariates were transformed by subtracting 25 km (a value above its maximum observed value) and flipping the sign. We assumed that none of the animals would respond to the sonar at these two offsets, but note that the exact values are not important because we used additive models. Each 1-min time-

Table 3
Terms in the full baseline model and exposure model. The baseline model was fitted to the time-bins in the baseline control series and the exposure model was fitted to all time-bins excluding tagging periods. Terms indicated with *s()* are smooth covariates and *ti()* is a tensor product interaction; others are factor covariates. *B* represents the covariates of the selected (reduced) baseline model.

Model	Specification
Baseline	*s(solarnoon) + s(depth_bathy) + BF_2h + UPAS_2h + TS + s(depth)*$^{1)}$ *+ state*$^{2)}$
Exposure	*B + NS + s(SELmax) + s(DISTmin) + ti(SELmax,DISTmin) + BF_2h_dur + s(DISTmin_post) + s(DISTmin_prev)*

[1] Only the full baseline model for *buzz* included whale *depth*.

[2] Only the full baseline model for *fluke* included behaviour *state*.

bin in the ‘baseline’ control series (i.e., the combined pre-exposure baseline and post-exposure data, excluding 20 min following each exposure session) was assigned a transformed value of 0 for *SELmax* and *DISTmin*. Continuous variables were included as thin plate spline smooth covariates and binary variables as factor covariates. *SELmax* and *DISTmin* were expected to induce monotonic changes in the response variables so the flexibility of their smooths was limited by setting $k = 5$ maximum degrees of freedom. Variable *solarnoon* was fitted as a cyclic penalised cubic regression spline smooth (with $k = 7$ and bs=”cc”).

Covariate selection was a two-step process. A selection of baseline covariates was carried out first using only the time-bins in the baseline control series. The best ‘baseline model’ was then carried forward for the selection of the ‘exposure model’ covariates using all time-bins in the dataset excluding tagging periods (Table 3). The thin-plate smooths had a shrinkage spline base which included a penalty to reduce model complexity by excluding smooth terms from the model (bs=”ts”). Binary variables were backwards-eliminated based on *p*-values in ANOVA until all were below 0.05. Smooths for baseline covariates with p-values above 0.05 in Wald-like tests were removed only after the stepwise baseline model selection. The exposure model for *fluke* did not converge and was therefore refitted with a linear interaction instead of the tensor product interaction. Data from both focal and non-focal individuals were included in the analysis.

3. Results

3.1. Data set

We analysed a total of 249 h of data (14,953 time-bins of 1 min) from 9 experiments with 14 tagged individuals (Table 1), including 291 h of data in the non-exposure control series of which 116 h were pre-exposure baseline data. The duration of the analysed data varied from 3.9 h (235 time-bins of 1 min) to 31.5 h (1,890 time-bins of 1 min) among the individuals (Table 1). All tagged sperm whales were presumably male based upon their body size and the location of the study site. Eight individuals were presented with XHPAS-C, XHPAS-D or HPAS-C exposure sessions by the frigate, 4 individuals with HPAS-C, HPAS-D or MPAS-C sessions by the research vessel, and for 2 individuals only pre-exposure data were analysed (Table 1). Four no-sonar control vessel approaches were conducted, of which 3 were the first session in the experiment (Table 1). The behavioural time series of one tagged sperm whale is illustrated (Fig. 1).

Blackfish detections were generally more prevalent in the data set than detections of unidentified and incidental sonar. Blackfish were detected in all tag records, with a total of 136 h (55 %) of analysed data having at least one blackfish detection in the preceding 2 h (variable *BF_2h*; Table 1). Twenty-one hours (8 %) of analysed data had at least one detection of unidentified sonar in the preceding 2 h (*UPAS_2h*) and 62 time-bins of 1 min contained incidental test sonar from the navy frigate (*TS*) (Table 1). Testing for the effects of blackfish presence and unidentified and incidental sonar exposures on the whales’ behaviour was conducted as part of the GAMM analysis (see section 3.3).

3.2. Time allocation, buzz rate and stroke rate

As expected for the study site, the tagged sperm whales allocated most of their time to foraging (Fig. 2), with the descent, layer-restricted search, ascent, and surface state accounting for 22 %, 36 %, 19 % and 17 % (total of 93 %) of the individual-average time budget during baseline, respectively. The non-foraging active state averaged 4.3 % of the time budget during baseline, 4.1 % during no-sonar control sessions, and 8.3 % during sonar exposure (Fig. 2a). The time spent in this state by the

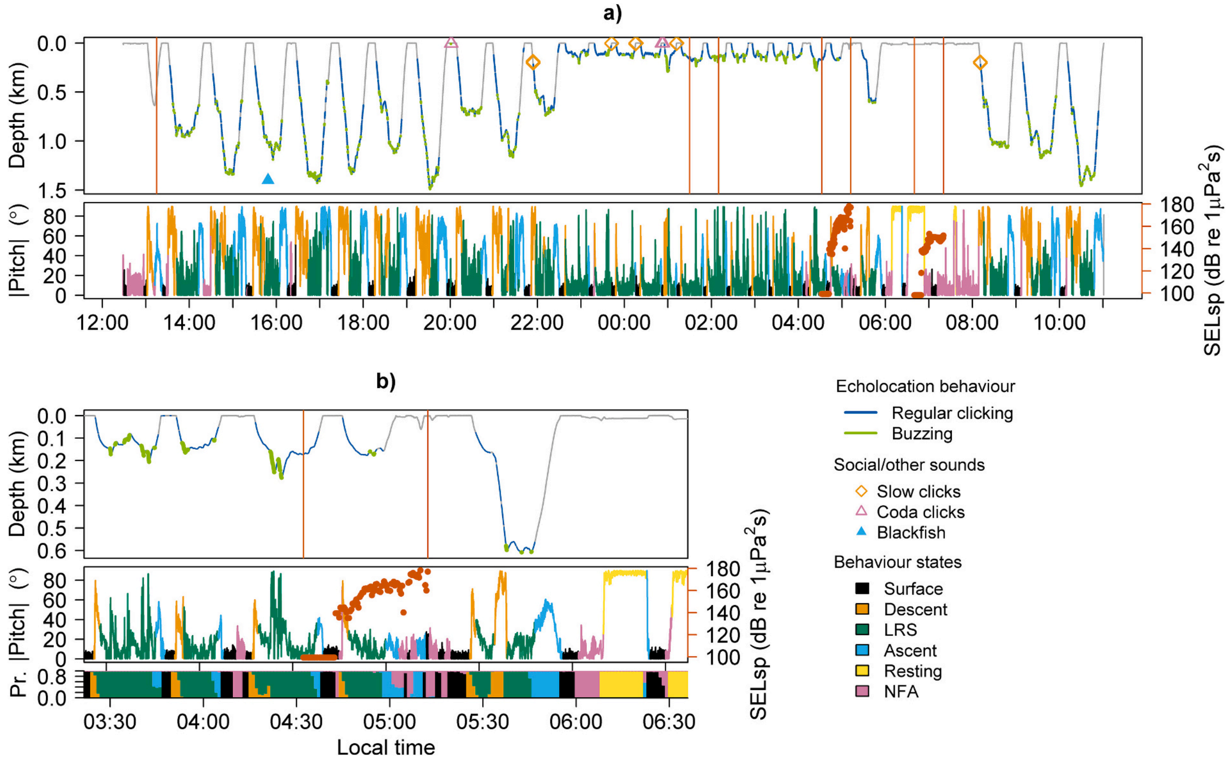

Fig. 1. Example tag data. a) Full tag record of sw19_259b with the depth and pitch time series shown at 5 Hz sample rate. Echolocation clicks and other click types produced by the tagged whale are shown in the upper panel. Blackfish detection is shown below the dive profile. Orange vertical lines in the upper panel indicate the start of the pre-exposure baseline and start/end of the consecutive no-sonar control, extra high-level PAS-*close* (XHPAS-C) and extra high-level PAS-*distant* (XHPAS-D) sessions. Behaviour state allocations at 1-min time resolution are shown in the lower panel. Orange dots on the lower panel’s second y-axis represent the single-pulse sound exposure levels on the tag. b) Detailed view of the period before, during and after the XHPAS-C sonar session with posterior probabilities of behaviour *state* shown in the bottom panel. LRS, layer-restricted search; NFA, non-foraging active.

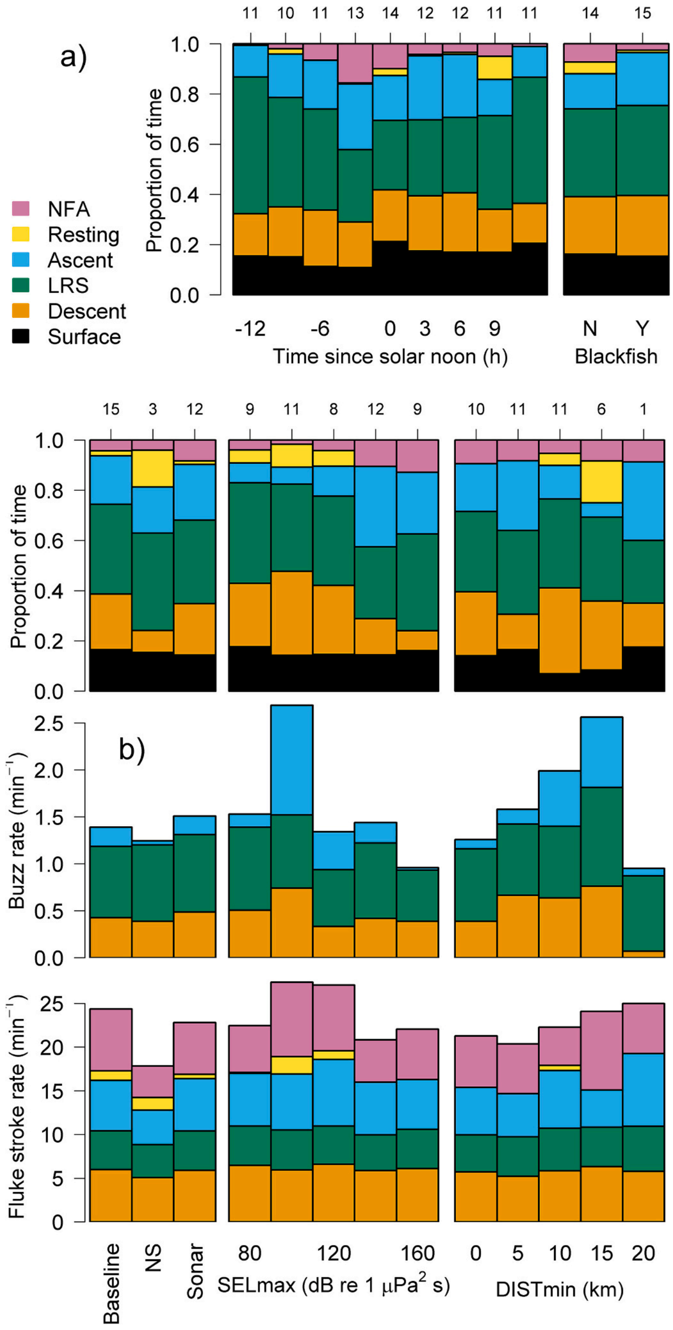

Fig. 2. Individual-average time budgets, buzz rates and stroke rates. a) Time-activity budgets calculated from the posterior probabilities of the categorical behaviour *state* and b) numbers of buzz starts and fluke strokes per 1 min time-bin, coloured by the posterior estimate of behaviour *state*. Data are shown as a function of context variables *solarnoon* and *BF_2h*, experimental phase, and sonar dose terms *SELmax* and *DISTmin*. Numbers on the top axis are sample sizes in individuals. NFA, non-foraging active; LRS, layer-restricted search; NS, no-sonar control. Buzz rates during baseline averaged 0.8 min^{-1} in layer-restricted search, 0.4 min^{-1} during descent, and 0.2 min^{-1} during ascent. Fluke stroke rates during baseline averaged 4.4 min^{-1} in layer-restricted search, 6.0 min^{-1} during descent, 5.8 min^{-1} during ascent, 1.1 min^{-1} during resting, and with 7.0 min^{-1} were highest during the non-foraging active state.

whales was correlated with the received *SELmax* of the sonar signals (Fig. 2a) and the non-foraging active state was allocated more often when the received single-pulse SEL was near its session maximum (Figs. 1, S1). Resting was the least common behaviour state; with an estimated 2.0 % of the individual-average time budget during baseline and 1.4 % during sonar exposure (Fig. 2a). The individual-average buzz rate and fluke stroke rate varied across behaviour states but, within states, were similar between baseline and sonar exposure (Fig. 2b).

3.3. *Behavioural response to sonar*

Sonar exposure effects were supported in the models for response variables *pNFA* and *buzz*, but not *fluke* (Table 4; see Table 2 for variable definitions). The GAMM analysis also supported effects of one or two baseline covariates for each response variable, highlighting the importance of controlling for natural fluctuations in variables of interest (Table 4). Effects of bathymetric depth (*depth_bathy*), no-sonar control (*NS*), unidentified sonar (*UPAS_2h*), incidental sonar (*TS*), or the presence of blackfish altering the sonar response itself (*BF_2h_dur*), on the whales' behaviour were not supported. The predictive performance of the models and how well they met model assumptions was considered good for *buzz* (Fig. 4) and acceptable for *pNFA* and *fluke* (Figs. S2, S4).

For the probability of non-foraging active state (*pNFA*), there was strong evidence for effects of time since *solarnoon* ($F = 9.3$, $p < 0.001$) and blackfish presence within the last 2 h (*BF_2h*; $F = 21.6$, $p < 0.001$) (Table 4). The percentage of time whales spent in the non-foraging active state at solar noon was predicted to be 5.6× higher than 12 h later, reflecting strong baseline fluctuations in the occurrence of this behaviour state at different times of day (Fig. S2c).

The analysis also supported an interactive sonar effect between received *SELmax* and *DISTmin* ($F = 0.2$, $p = 0.02$) and an effect of repeat exposures (*DISTmin_prev*; $F = 0.8$, $p = 0.03$) on response variable *pNFA* (Table 4). The time that sperm whales spent in the non-foraging active state during the first sonar exposure session was predicted to increase from a typical baseline probability of 2.9 % (95 % CI 0.8–5.0 %; low/high baseline probabilities 0.6/8.1 %) to 8.5 % (−1.8–19 %; 1.9/21 %) at *SELmax* 180 dB re 1 μPa^2 s and *DISTmin* 1 km, representing a 1.9× (−1.6–5.5; 1.4/3.2) increase (Figs. 3, S3). Based on these model predictions, we are 95 % confident that the probability of the non-foraging active state decreased back to baseline levels for *DISTmin* 13 km or greater depending on *SELmax* (dashed contours in Figs. 3 and S3). The supported interaction of *SELmax* with *DISTmin* indicated the existence of two response thresholds at which previously unexposed animals

Table 4
Model statistics for the selected exposure models. The analysed response variables were the probability of non-foraging active state (*pNFA*), presence/absence of buzz starts (*buzz*), and fluke stroke rate (*fluke*). E.d.f., effective degrees of freedom; (ref.) d.f., (reference) degrees of freedom. Statistics for factor and smooth covariates are based on ANOVA and Wald-like tests, respectively. Only covariates with a *p*-value below 0.05 were interpreted.

Response variable	Predictor variable	e.d.f.	(ref.) d.f.	F	p
pNFA	*BF_2h*	–	1	21.6	<0.001
	s(solarnoon)	3.4	5	9.3	<0.001
	s(SELmax)	0.0	4	0.0	0.54
	s(DISTmin)	0.0	4	0.0	0.75
	ti(SELmax,DISTmin)	1.0	16	0.2	0.02
	s(DISTmin_prev)	0.8	4	0.8	0.03
	s(DISTmin_post)	0.7	4	0.6	0.06
buzz	*s(depth)*	7.9	9	128.5	<0.001
	s(solarnoon)	4.3	5	10.8	<0.001
	s(SELmax)	0.9	4	1.5	0.005
	s(DISTmin)	0.0	4	0.0	0.37
	ti(SELmax,DISTmin)	0.0	16	0.0	0.68
	s(DISTmin_prev)	0.3	4	0.1	0.22
	s(DISTmin_post)	0.0	4	0.0	0.60
fluke	*state*	–	4	254.0	<0.001
	BF_2h	–	1	19.5	<0.001
	SELmax:DISTmin	–	1	0.4	0.51
	s(SELmax)	0.0	4	0.0	0.06
	s(DISTmin)	0.0	4	0.0	0.92
	s(DISTmin_prev)	0.0	4	0.0	0.55
	s(DISTmin_post)	0.9	4	3.0	<0.001

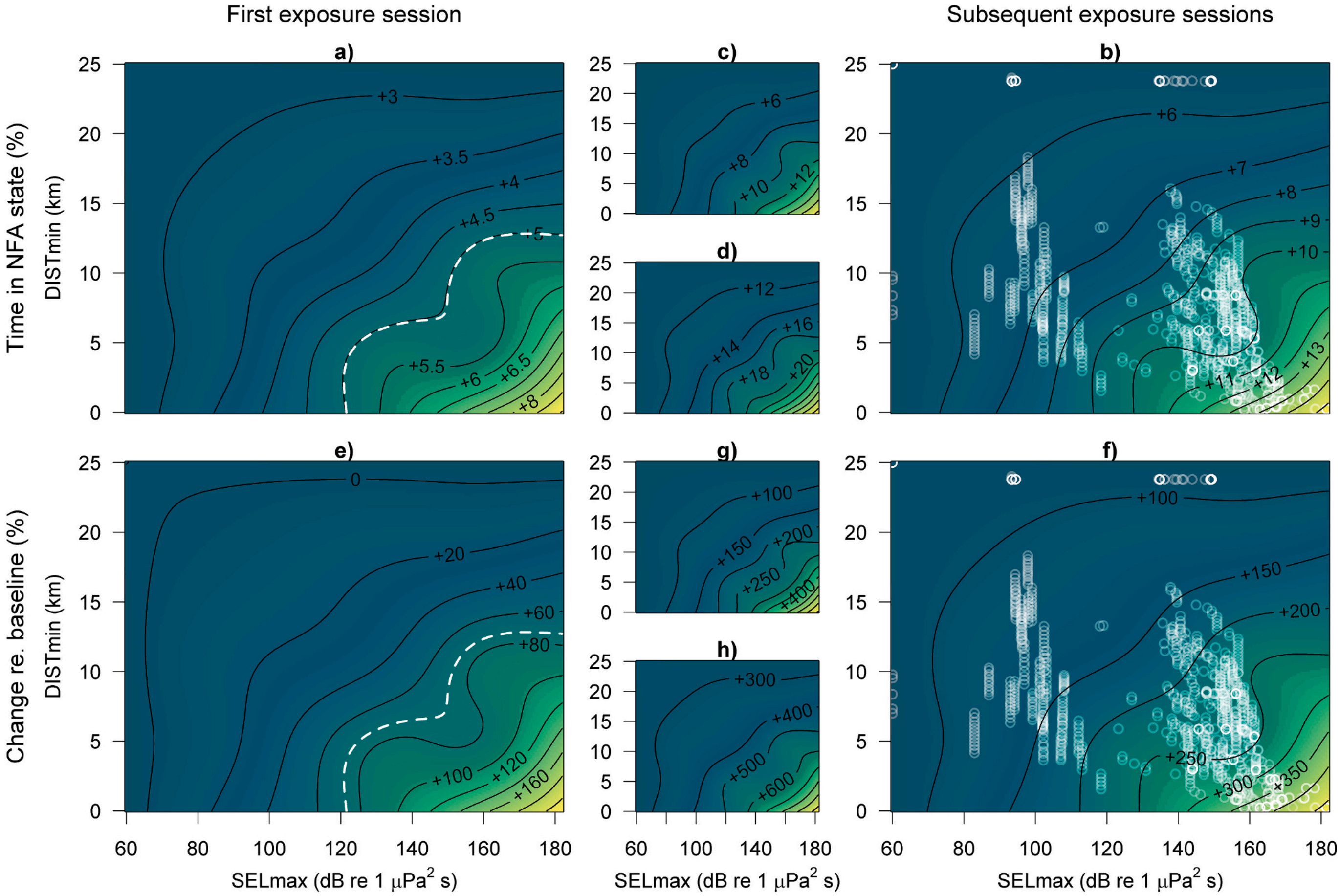

Fig. 3. Predicted percentage of time spent in non-foraging active (NFA) state. Mean predictions for the a) first exposure session (*DISTmin_prev* 25 km) and b) subsequent exposure sessions (*DISTmin_prev* 5 km; the average observed value). Panels c and d show the corresponding upper confidence bounds, respectively. e-h) Same as panels a-d but illustrated as percentage change in p(NFA) relative to the estimate for the baseline data (assigned the back-transformed values of *SELmax* 60 dB re 1 μPa^2 s and *DISTmin* 25 km for 'no response'). Predictions were generated using *solarnoon* 2 h, *BF_2h* true and *DISTmin_post* 25 km to represent the typical baseline condition (Fig. S2c). Dashed contour lines in panels a and e indicate the upper confidence bound of the baseline estimate. Circles indicate the data coverage (for clarity, only shown in panels b and f).

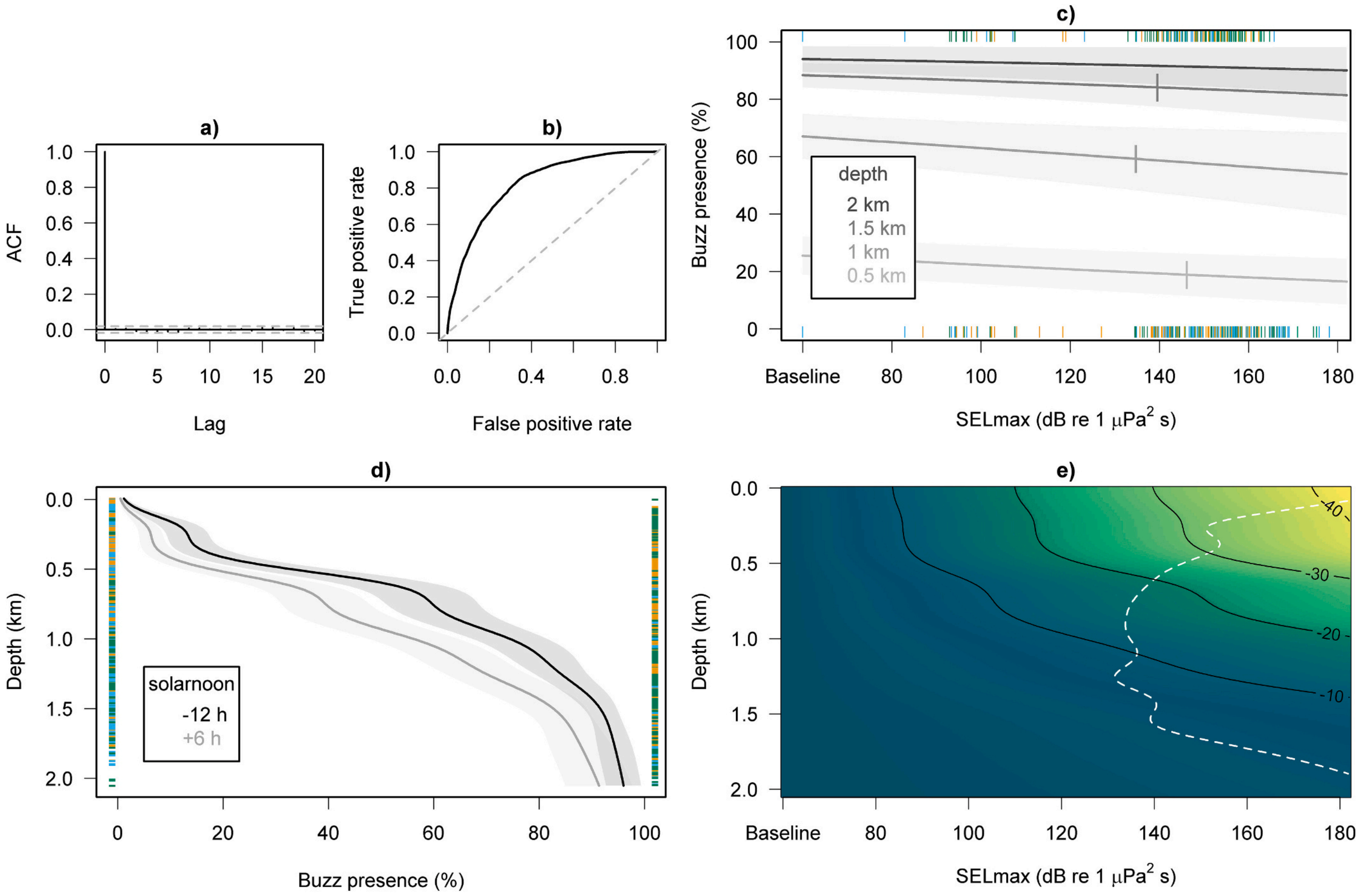

Fig. 4. Model evaluation and predictions for the presence of buzz starts during foraging. a) Autocorrelation function of the within-individual residuals; b) receiver operator characteristic (ROC) curve constructed from the model-predicted buzz presence vs. observed buzz presence; c) mean predictions of buzz presence, shown as a percentage of the 1-min time-bins, with 95 % confidence intervals, for c) *SELmax* at four whale *depths* and d) whale *depth* at the two extremes of *solarnoon*; and e) predicted percentage change in *buzz* relative to the estimate for the baseline data (assigned the back-transformed value of *SELmax* 60 dB re 1 μPa^2 s). Predictions in panels c and e were generated using *solarnoon* − 3 h to represent the typical baseline condition. Vertical line markers (panel c) or the dashed contour line (panel e) indicate the lower confidence bound of the baseline estimate.

initiated non-foraging active behaviour; the first at ~120 dB re 1 μPa^2 s for whales within 5 km from the source and the second at ~150 dB re 1 μPa^2 s for whales within 13 km (Fig. 3a, e). When the whale had been approached by the source to 5 km in a previous sonar exposure session (i.e., the average value of *DISTmin_prev*), time spent in the non-foraging active state was predicted to increase further to 15 % (0.1–31 %; 3.7/35 %) at *SELmax* 180 dB re 1 μPa^2 s and *DISTmin* 1 km, representing a 4.3× (−1.0–9.6; 3.6/6.0) increase from baseline levels (Figs. 3, S3).

For the presence/absence of buzz starts (*buzz*) during foraging dives, modelling revealed strong evidence for effects of whale *depth* ($F = 128.5$, $p < 0.001$), time since *solarnoon* ($F = 10.8$, $p < 0.001$) and received *SELmax* ($F = 1.5$, $p = 0.005$) (Table 4). The sperm whales' production of these foraging buzzes, which are indicators of prey capture attempts, was predicted to decrease from over 90 % of the 1-min time-bins at 2 km depth to nearly 0 % at the surface during baseline (Fig. 4d). Buzz presence was also predicted to decrease with increasing *SELmax* (Fig. 4c, e). For example, the predicted decrease from baseline levels to *SELmax* 180 dB re 1 μPa^2 s corresponded to from 26 % (95 % CI 19–32 %) to 17 % (9–25 %) at 500 m depth and from 88 % (84–93 %) to 82 % (73–91 %) at 1.5 km depth (Fig. 4c). Based on the model predictions, we are 95 % confident that buzz production rose back to baseline levels for received *SELmax* below ~130–150 dB re 1 μPa^2 s and most dive depths (Fig. 4c, e).

Modelling also supported effects of behaviour *state* ($F = 254.0$, $p < 0.001$) and presence of blackfish (*BF_2h*; $F = 19.5$, $p < 0.001$) on the fluke stroke rate of the whales (Table 4; Fig. S4). A post-exposure covariate (*DISTmin_post*; $F = 3.0$, $p < 0.001$) in the model for fluke was also supported (Table 4). According to the model predictions, the extent to which the sonar exposure affected the whales' stroking locomotor activity during post-exposure as function of the DISTmin of the exposure session was small (Figs. S4d, S4e).

4. Discussion

This study was designed to separate the potential effects of received level of sonar and source distance on the behavioural time series of sperm whales. After controlling for baseline variations in behaviour, we observed that naval sonar exposure was associated with 1) increased time allocations of the 'active non-foraging' behaviour state and 2) decreased buzz presence when whales kept foraging, indicating that whales made fewer prey capture attempts (which might be indicative of reduced foraging success). Overall, our data supported the hypothesis that both sonar received level and source distance were driving the behavioural response. Source distance was found to modulate the effect of received level on the time spent in the non-foraging active behaviour state (Fig. 3), but not on buzz rate (Table 4).

Our study provides further evidence that male sperm whales in a high latitude habitat respond to naval sonar exposure by reducing time spent foraging. Short-term changes in foraging or in other fitness-enhancing activities have the potential to affect vital rates if such changes occur repeatedly, over long periods or during phenologically-important times, particularly for species with limited behavioural plasticity (Pirotta et al., 2018; Farmer et al., 2018). Naval sonar and its effects may also interact with other anthropogenic stressors and their effects (Tyack et al., 2022; Orr et al., 2020; Pirotta et al., 2022). Whether such long-term effects are likely for male sperm whales in northern Norway is beyond the scope of this study and requires additional information (e.g., on nutritional buffering capacity). Instead of a direct effect, acoustic interference of echolocation may also have caused the apparent change in buzz production during our experiments, especially during periods when the received single-pulse sound exposure level was above 140–160 dB re 1 μPa^2 s (von Benda-Beckmann et al., 2021; Isojunno et al., 2021).

The observed responses were very similar to those in previous behavioural response studies on sperm whales in the area, despite this study only reusing some of their baseline control data (Isojunno et al., 2016; Miller et al., 2012; Isojunno et al., 2020; Isojunno et al., 2021; Curé et al., 2021). The main similarities with these studies were the overall type and magnitude of the responses, increased responsiveness in subsequent sessions (indicating short-term sensitisation), support for context factors related to whale depth, behaviour state and time-of-day, lack of effects of incidental and unidentified sonar (likely from distant sources), and high levels of unexplained variation. The main differences were the decrease in the probability of being in the non-foraging active state during baseline associated with blackfish detections, and overall reduced sensitivity compared to 1–2 kHz PAS experiments conducted in the same area in 2008–2009 (Fig. S1). Differences in the effect of blackfish on the probability of non-foraging active behaviour may have resulted from an overall higher prevalence of blackfish in our dataset from 2019, with detections noted for each tag record (Table 1), compared with previous years (Isojunno et al., 2020; Curé et al., 2021). Our methods for detecting blackfish (often faint acoustic or distant visual detections) did not enable us to consistently identify the precise species of blackfish (or ecotype of killer whale) present. Therefore, it is possible that these differences may have been due to the (unknown) composition of the blackfish encountered, which can include potential predators in the form of mixed-diet killer whales as well as non-predator species (fish-specialist killer whales, long-finned pilot whales).

Overall reduced sensitivity compared to PAS experiments in 2008–2009 was previously observed by Isojunno et al. (2020), who suggested that the sonar signal was a more novel presentation in 2008–2009 when the Norwegian Navy had recently brought this low-frequency sonar system into service, though some detailed differences in the experimental sonar exposure protocols between 2008 or 2009 and 2019 could also have led to reduced responsiveness to the sonar presentations in the present study. However, the potential importance of the novelty of the sonar in the responsiveness of whales aligns with the relatively strong response of sperm whales during a naval exercise off eastern Canada in 2016 that was consistent with large-scale feeding cessation and area avoidance (Stanistreet et al., 2022). Those whales may have been less accustomed to sonar signals, as no other such sounds were observed at this and other recording sites in the region over a 2-year period (Stanistreet et al., 2022). Sperm whales are rapid social learners (e.g., (Whitehead et al., 2021)) so the whales in our study may already have learned to associate especially the distant sonar signals with a lower perceived risk, reducing their responsiveness to those signals. However, a lack of response could also simply reflect a lack of options (Gill et al., 2001), and the most responsive animals might not be those that are most vulnerable to disturbance (Beale and Monaghan, 2004; Burslem et al., 2022).

For whale responses to be affected by source distances of several kilometres, the whales likely use auditory cues to detect signals and estimate the distance to the source. The main auditory distance cues in humans and birds are based on frequency content, reverberation and received level of the sounds (Naguib and Wiley, 2001; Kolarik et al., 2016). Human listeners fail to accurately estimate source distance for unfamiliar sounds, and their estimates improve with repeated exposure (Kolarik et al., 2016; Coleman, 1962). Less research on auditory distance perception has been conducted in cetaceans compared to humans, but bottlenose dolphins (*Tursiops truncatus*) also use at least high-frequency harmonic attenuation and reverberation to estimate distance (Mulsow et al., 2018). Familiarity with the signal may also help cetaceans with source distance estimation.

Most of the experiments were conducted by a navy frigate towing an operational sonar source. The higher source level was crucial for addressing the study objective. Use of the navy frigate also increases the operational relevance of the results, as it alleviates the concern that disturbance effects of actual exercises could be very different than those inferred based on extrapolations from less powerful sources. Caveats related to the use and availability of the operational source included minor differences in exposure characteristics compared with earlier studies (mainly a more rapid increase in source level during ramp-up due to a limited number of attenuation levels; Fig. S1), a research period with some darkness hours (unsuitable for tagging and visual tracking), and challenges during pre-cruise planning and in scheduling of experiments. The experimental phase had to be shortened on some occasions and sonar sessions were given priority over no-sonar control sessions when frigate availability was limited.

The low number of no-sonar sessions ($n = 4$; Table 1) did not allow for strong statistical inference about the vessel itself being a response driver. During these four sessions the foraging behaviour of the focal whales did not appear to have been significantly affected, similar to 16 different sperm whales presented with such controls in previous studies using similar analyses (Isojunno et al., 2016; Isojunno et al., 2020). In other studies (Curé et al., 2021; Curé et al., 2016), no-sonar vessel approach sessions were associated with some changes in behaviour other than foraging but the probability and severity of these changes were lower than during sonar exposure sessions. The present work mainly focussed on foraging and thus did not consider all types of behavioural disturbance (e.g., social responses) that may occur due to sonar exposure.

High-resolution, multivariate timeseries data from animal-attached sensors are inherently challenging to analyse statistically. This study built upon the methodology developed by another study (Isojunno et al., 2020) which it extended by using the ARMA(1,1) correlation structure to reduce residual autocorrelation. While our statistical approach can be considered advanced, several future methodological improvements are possible, at least theoretically. Uncertainty in the behaviour state estimation (step 1) could be carried over to the GAMM analysis (step 2) by using random draws from the posterior distribution of the Bayesian model and refitting the GAMM multiple times, in a bootstrapping approach to uncertainty estimation. Further, the two-step modelling approach could also conceivably be combined into one by incorporating the use of covariates into the state estimation process (e.g. (Klappstein et al., 2023)). Such an approach could have the added benefit of more fully describing the behaviour of the animals as it reduces concerns about multiple testing of the same behavioural responses.

Quantitative assessments of the potential consequences of noise disturbance on areas or cetacean populations often use received level-based dose-response functions that are derived from different sound sources and exposure contexts (e.g., (Farmer et al., 2018)). A better understanding of when and in which contexts source distance affects behaviour independent of received level should improve the accuracy of their predictions and result in more effective conservation management. This study adds to the growing body of knowledge on the distance-mediated behavioural effects of anthropogenic noise in cetaceans (Southall et al., 2016; Wensveen et al., 2019; Southall et al., 2019b; DeRuiter et al., 2013; Falcone et al., 2017; Dunlop et al., 2017) and provides important input to noise criteria and risk thresholds for behavioural disturbance (Accomando et al., 2024). Its empirical data on sperm whale natural behaviour and responses can also be used to inform and provide input parameters to PCoD (population consequences of disturbance) modelling efforts (Pirotta et al., 2018; Farmer et al., 2018) which increasingly include effects of multiple stressors (Tyack et al., 2022; Pirotta et al., 2022).

CRediT authorship contribution statement

Paul J. Wensveen: Writing – review & editing, Writing – original draft, Visualization, Software, Resources, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Saana Isojunno:** Writing – review & editing, Visualization, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Petter H. Kvadsheim:** Writing – review & editing, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization. **Frans-Peter A. Lam:** Writing – review & editing, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization. **Charlotte Curé:** Writing – review & editing, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Alexander M. von Benda-Beckmann:** Writing – review & editing, Methodology, Formal analysis, Conceptualization. **Patrick J.O. Miller:** Writing – review & editing, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Funding

This work was supported by US Naval Facility Engineering Command / Living Marine Resources (LMR) research program (grant no. N39430–17-C-1935); The Netherlands Ministry of Defense; Defense Science and Technology Lab (UK Ministry of Defense; grant no. DSTLX-1000137649); and DGA (French Ministry of Defense; grant no. 1883003901). The Royal Norwegian Navy supported the work by supplying access to the frigate KNM Otto Sverdrup. PW was supported by the RANNÍS Icelandic Research Fund (grant no. 207081).

Declaration of competing interest

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

Acknowledgements

We sincerely thank everyone involved in the preparation, execution, and data processing of the 3S-2019-OPS field experiments including Peter Tyack, René Dekeling, Ron Filipowicz, Odile Gerard, Rune Roland, Lars Kleivane, Benjamin Benti, Alec Burslem, Lucia Martin López, Kalliopi Gkikopoulou, Heather Vance, Catarina Alves, Marije Siemensma, Liz Henderson McGhee, Jacqueline Bort, Mark van Spellen, Sander van IJsselmuide, Martijn van Riet, Jene Nissen, Julie Rivers, CDR Tom Erik Lindhjem, and the crews of R/V H.U. Sverdrup II and KNM Otto Sverdrup.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.marpolbul.2025.117742.

Data availability

Relevant data and code are within the paper or is available from https://github.com/PaulWens/sperm-whale-RL-v-Distance

References

Accomando, A., Finneran, J., Henderson, E., Jenkins, K., Kotecki, S., Martin, C., Mulsow, J., Zapetis, M., 2024. Criteria and thresholds for U.S. In: Navy Acoustic and Explosive Effects Analysis (Phase 4). Technical Report of the Naval Information Warfare. Center Pacific (San Diego, CA, pp. 92152–95001.

Åtland, K., Nilsen, T., Pedersen, T., 2022. Military muscle-flexing as interstate communication: Russian NOTAM warnings off the coast of Norway, 2015–2021. Scand. J. Mil. Stud. 5, 63–78. https://doi.org/10.31374/sjms.133.

Azzara, A.J., von Zharen, W.M., Newcomb, J.J., 2013. Mixed-methods analytic approach for determining potential impacts of vessel noise on sperm whale click behavior. J. Acoust. Soc. Am. 134, 4566–4574. https://doi.org/10.1121/1.4828819.

Beale, C.M., Monaghan, P., 2004. Behavioural responses to human disturbance: a matter of choice? Anim. Behav. 68, 1065–1096.

Bernaldo de Quirós, Y., et al., 2019. Advances in research on the impacts of anti-submarine sonar on beaked whales. P. Roy. Soc. B 286, 20182533. https://doi.org/10.1098/rspb.2018.2533.

Booth, C.G., Brannan, N., Dunlop, R., Friedlander, A., Isojunno, S., Miller, P., Quick, N., Southall, B., Pirotta, E., 2022. A sampling, exposure and receptor framework for identifying factors that modulate behavioural responses to disturbance in cetaceans. J. Anim. Ecol. 91, 1948–1960. https://doi.org/10.1111/1365-2656.13787.

Bowles, A.E., Smultea, M., Wursig, B., DeMaster, D.P., Palka, D., 1994. Relative abundance and behavior of marine mammals exposed to transmissions from the heard island feasibility test. J. Acoust. Soc. Am. 96, 2469–2484.

Burslem, A., Isojunno, S., Pirotta, E., Miller, P.J.O., 2022. Modelling the impact of condition-dependent responses and lipid-store availability on the consequences of disturbance in a cetacean. Conserv. Physiol. 10, 1–42. https://doi.org/10.1093/conphys/coac069.

Coleman, P.D., 1962. Failure to localise the source distance of an unfamiliar sound. J. Acoust. Soc. Am. 34, 345–346. https://doi.org/10.1121/1.1928121.

Cox, T.M., et al., 2006. Understanding the impacts of anthropogenic sound on beaked whales. J. Cetacean Res. Manag. 7, 177–187.

Crane, A.L., Ferrari, M.C.O., 2017. Patterns of predator neophobia: a meta-analytic review. P. Roy. Soc. B 284, 20170583. https://doi.org/10.1098/rspb.2017.0583.

Curé, C., Isojunno, S., Visser, F., Wensveen, P.J., Sivle, L.D., Kvadsheim, P.H., Lam, F.A., Miller, P.J.O., 2016. Biological significance of sperm whale responses to sonar : comparison with anti-predator responses. Endanger. Species Res. 31, 89–102. https://doi.org/10.3354/esr00748.

Curé, C., et al., 2021. Severity scoring of behavioural responses of sperm whales (*Physeter macrocephalus*) to novel continuous versus conventional pulsed active sonar. J. Mar. Sci. Eng. 9, 444. https://doi.org/10.3390/jmse9040444.

Depledge, D., 2020. Train where you expect to fight: why military exercises have increased in the high north. Scand. J. Mil. Stud. 3, 288–301. https://doi.org/10.31374/sjms.64.

DeRuiter, S.L., et al., 2013. First direct measurements of behavioural responses by Cuvier's beaked whales to mid-frequency active sonar. Biol. Lett. 9, 20130223. https://doi.org/10.1098/rsbl.2013.0223.

Dominoni, D.M., et al., 2020. Why conservation biology can benefit from sensory ecology. Nat. Ecol. Evol. 4, 502–511. https://doi.org/10.1038/s41559-020-1135-4.

Dunlop, R.A., Noad, M.J., McCauley, R.D., Kniest, E., Slade, R., Paton, D., Cato, D.H., 2017. The behavioural response of migrating humpback whales to a full seismic airgun array. P. Roy. Soc. B 284, 20171901. https://doi.org/10.1098/rspb.2017.1901.

Fais, A., Aguilar Soto, N., Johnson, M., Pérez-González, C., Miller, P.J.O., Madsen, P.T., 2015. Sperm whale echolocation behaviour reveals a directed, prior-based search strategy informed by prey distribution. Behav. Ecol. Sociobiol. 69, 663–674. https://doi.org/10.1007/s00265-015-1877-1.

Falcone, E.A., Schorr, G.S., Watwood, S.L., DeRuiter, S.L., Zerbini, A.N., Andrews, R.D., Morrissey, R.P., Moretti, D.J., 2017. Diving behaviour of Cuvier's beaked whales exposed to two types of military sonar. R. Soc. Open Sci. 4, 170629. https://doi.org/10.1098/rsos.170629.

Farmer, N.A., Baker, K., Zeddies, D.G., Denes, S.L., Noren, D.P., Garrison, L.P., Machernis, A., Fougères, E.M., Zykov, M., 2018. Population consequences of disturbance by offshore oil and gas activity for endangered sperm whales (*Physeter macrocephalus*). Biol. Conserv. 227, 189–204. https://doi.org/10.1016/j.biocon.2018.09.006.

Gill, J.A., Norris, K., Sutherland, W.J., 2001. Why behavioural responses may not reflect the population consequences of human disturbance. Biol. Conserv. 97, 265–268.

Harris, C.M., Sadykova, D., DeRuiter, S.L., Tyack, P.L., Miller, P.J.O., Kvadsheim, P.H., Lam, F.P.A., Thomas, L., 2015. Dose response severity functions for acoustic disturbance in cetaceans using recurrent event survival analysis. Ecosphere 6, art236. https://doi.org/10.1890/ES15-00242.1.

Harris, C.M., et al., 2018. Marine mammals and sonar: dose-response studies, the risk-disturbance hypothesis and the role of exposure context. J. Appl. Ecol. 55, 396–404. https://doi.org/10.1111/1365-2664.12955.

Harrison, X.A., 2021. A brief introduction to the analysis of time-series data from biologging studies. Philos. Trans. R. Soc. B Biol. Sci. 376. https://doi.org/10.1098/rstb.2020.0227.

Heide-jørgensen MP *et al.* 2021 Behavioral response study on seismic airgun and vessel exposures in narwhals. Front. Mar. Sci. 8, 658173. (doi:https://doi.org/10.3389/fmars.2021.658173).

Hersh, T.A., et al., 2022. Evidence from sperm whale clans of symbolic marking in non-human cultures. P. Natl. Acad. Sci. 119. https://doi.org/10.1073/pnas.2201692119.

Hooker, S.K., et al., 2012. Deadly diving? Physiological and behavioural management of decompression stress in diving mammals. P. Roy. Soc. B 279, 1041–1050. https://doi.org/10.1098/rspb.2011.2088.

Isojunno, S., Curé, C., Kvadsheim, P., Lam, F., Tyack, P., Wensveen, P., Miller, P., 2016. Sperm whales reduce foraging effort during exposure to 1-2 kHz sonar and killer whale sounds. Ecol. Appl. 26, 77–93.

Isojunno, S., Miller, P.J.O., 2015. Sperm whale response to tag boat presence: biologically informed hidden state models quantify lost feeding opportunities. Ecosphere 6, 6. https://doi.org/10.1890/ES14-00130.1.

Isojunno, S., Wensveen, P.J., Lam, F.-P.A., Kvadsheim, P.H., von Benda-Beckmann, A.M., Martín López, L.M., Kleivane, L., Siegal, E.M., Miller, P.J.O., 2020. When the noise goes on: received sound energy predicts sperm whale responses to both intermittent and continuous navy sonar. J. Exp. Biol. 223, jeb219741. https://doi.org/10.1242/jeb.219741.

Isojunno, S., et al., 2021. Sperm whales exhibit variation in echolocation tactics with depth and sea state but not naval sonar exposures. Mar. Mamm. Sci. 38, 682–704. https://doi.org/10.1111/mms.12890.

Johnson, M.P., Tyack, P.L., 2003. A digital acoustic recording tag for measuring the response of wild marine mammals to sound. IEEE J. Ocean. Eng. 28, 3–12. https://doi.org/10.1109/Joe.2002.808212.

Joyce, T.W., Durban, J.W., Claridge, D.E., Dunn, C.A., Hickmott, L.S., Fearnbach, H., Dolan, K., Moretti, D., 2019. Behavioral responses of satellite tracked Blainville's beaked whales (*Mesoplodon densirostris*) to mid-frequency active sonar. Mar. Mamm. Sci. 36, 29–46. https://doi.org/10.1111/mms.12624.

Klappstein, N.J., Thomas, L., Michelot, T., 2023. Flexible hidden Markov models for behaviour-dependent habitat selection. Mov. Ecol. 11, 1–13. https://doi.org/10.1186/s40462-023-00392-3.

Kleivane, L., Kvadsheim, P.H., Bocconcelli, A., Øien, N., Miller, P.J.O., 2022. Equipment to tag, track and collect biopsies from whales and dolphins: the ARTS. DFHorten and LKDart systems. Anim. Biotelemetry 10, 32. https://doi.org/10.1186/s40317-022-00303-0.

Kolarik, A.J., Moore, B.C.J., Zahorik, P., Cirstea, S., 2016. Auditory distance perception in humans: a review of cues, development, neuronal bases, and effects of sensory loss. Atten. Percept. Psychophysiol. 78, 373–395. https://doi.org/10.3758/s13414-015-1015-1.

Kunc, H.P., McLaughlin, K.E., Schmidt, R., 2016. Aquatic noise pollution: implications for individuals, populations, and ecosystems. P. Roy. Soc. B 283, 20160839. https://doi.org/10.1098/rspb.2016.0839.

Kvadsheim P *et al.* 2021 The 3S3 experiment data report - using operational naval sonars to study the effects of continuous active sonar, and source proximity, on sperm whales. FFI Report 21/00688.

Kvadsheim PH *et al.* 2020 Studying the effect of source proximity in sperm whales and continous sonar in pilot whales using operational sonars - the 3S-2019-OPS cruise report. FFI Report 20/01749.

Lavery, T.J., Roudnew, B., Gill, P., Seymour, J., Seuront, L., Johnson, G., Mitchell, J.G., Smetacek, V., 2010. Iron defecation by sperm whales stimulates carbon export in the Southern Ocean. P. Roy. Soc. B 277, 3527–3531. https://doi.org/10.1098/rspb.2010.0863.

Madsen, P.T., Møhl, B., Nielsen, B.K., Wahlberg, M., 2002. Male sperm whale behaviour during exposures to distant seismic survey pulses. Aquat. Mamm. 28, 231–240.

Martin, M.J., et al., 2022. Exposure and behavioral responses of tagged beluga whales (*Delphinapterus leucas*) to ships in the Pacific Arctic. Mar. Mamm. Sci. 39, 387–421. https://doi.org/10.1111/mms.12978.

Merchant, N.D., Putland, R.L., André, M., Baudin, E., Felli, M., Slabbekoorn, H., Dekeling, D., 2022. A decade of underwater noise research in support of the European marine strategy framework directive. Ocean Coast. Manag. 228, 106299. https://doi.org/10.1016/j.ocecoaman.2022.106299.

Miller, P., et al., 2011. The 3S experiments: Studying the behavioural effects of naval sonar on killer whales (*Orcinus orca*). In: Sperm Whales (*Physeter macrocephalus*), and Long-Finned Pilot Whales (*Globicephala melas*) in Norwegian Waters. Scottish Oceans Institute, St Andrews, UK. Technical Report SOI-2011-001.

Miller, P.J., Johnson, M.P., Madsen, P.T., Biassoni, N., Quero, M., Tyack, P.L., 2009. Using at-sea experiments to study the effects of airguns on the foraging behavior of sperm whales in the Gulf of Mexico. Deep Sea Res. I 56, 1168–1181. https://doi.org/10.1016/j.dsr.2009.02.008.

Miller, P.J., Johnson, M.P., Tyack, P.L., 2004. Sperm whale behaviour indicates the use of echolocation click buzzes 'creaks' in prey capture. P. Roy. Soc. B 271, 2239–2247. https://doi.org/10.1098/rspb.2004.2863.

Miller, P.J.O., Isojunno, S., Siegal, E., Lam, F.P.A., Kvadsheim, P.H., Cure, C., 2022. Behavioral responses to predatory sounds predict sensitivity of cetaceans to anthropogenic noise within a soundscape of fear. P. Natl. Acad. Sci. 119, 1–8. https://doi.org/10.1073/pnas.2114932119.

Miller, P.J.O., et al., 2012. The severity of behavioral changes observed during experimental exposures of killer (*Orcinus orca*), long-finned pilot (*Globicephala melas*), and sperm (*Physeter macrocephalus*) whales to naval sonar. Aquat. Mamm. 38, 362–401. https://doi.org/10.1578/AM.38.4.2012.362.

Mulsow, J., Finneran, J.J., Schlundt, C.E., Jones, R., 2018. Bottlenose dolphin (*Tursiops truncatus*) discrimination of harmonic stimuli with range- dependent signal degradation. J. Acoust. Soc. Am. 3434, 3434. https://doi.org/10.1121/1.5040498.

Naguib, M., Wiley, R.H., 2001. Estimating the distance to a source of sound: mechanisms and adaptations for long-range communication. Anim. Behav. 62, 825–837. https://doi.org/10.1006/anbe.2001.1860.

Orr, J.A., et al., 2020. Towards a unified study of multiple stressors: divisions and common goals across research disciplines. P. Roy. Soc. B 287, 20200421. https://doi.org/10.1098/rspb.2020.0421.

Pirotta, E., et al., 2018. Understanding the population consequences of disturbance. Ecol. Evol. 8, 9934–9946. https://doi.org/10.1002/ece3.4458.

Pirotta, E., et al., 2022. Understanding the combined effects of multiple stressors: a new perspective on a longstanding challenge. Sci. Total Environ. 821, 153322. https://doi.org/10.1016/j.scitotenv.2022.153322.

Ramirez-Martinez, N.C., Víkingsson, G.A., Øien, N.I., Mikkelsen, B., Gunnlaugsson, T., Hammond, P.S., 2024. Distribution and habitat use of deep-diving cetaceans in the central and North-Eastern North Atlantic. NAMMCO Sci. Publ. 13, 1–27. https://doi.org/10.7557/3.7416.

Roman, J., et al., 2014. Whales as marine ecosystem engineers. Front. Ecol. Environ. 12, 377–385. https://doi.org/10.1890/130220.

Santos, M.B., et al., 1999. Stomach contents of sperm whales *Physeter macrocephalus* stranded in the North Sea 1990-1996. Mar. Ecol. Prog. Ser. 183, 281–294. https://doi.org/10.3354/meps183281.

Similä, T., Haug, T., Lindblom, L., Lockyer, C., O'Callaghan, S.A., 2022. Stomach contents of three sperm whales (*Physeter macrocephalus*) stranded on Andøya, northern Norway. Aquat. Mamm. 48, 449–455. https://doi.org/10.1578/AM.48.5.2022.449.

Southall, B., Nowacek, D., Miller, P., Tyack, P., 2016. Experimental field studies to measure behavioral responses of cetaceans to sonar. Endanger. Species Res. 31, 293–315. https://doi.org/10.3354/esr00764.

Southall, B.L., Finneran, J.J., Reichmuth, C., Nachtigall, P.E., Ketten, D.R., Bowles, A.E., Ellison, W.T., Nowacek, D.P., Tyack, P.L., 2019a. Marine mammal noise exposure criteria: updated scientific recommendations for residual hearing effects. Aquat. Mamm. 45, 125–232. https://doi.org/10.1578/am.45.2.2019.125.

Southall, B.L., et al., 2019b. Behavioral responses of individual blue whales (*Balaenoptera musculus*) to mid-frequency military sonar. J. Exp. Biol. 222, jeb190637. https://doi.org/10.1242/jeb.190637.

Stanistreet, J.E., Beslin, W.A.M., Kowarski, K., Martin, S.B., Westell, A., Moors-Murphy, H.B., 2022. Changes in the acoustic activity of beaked whales and sperm whales recorded during a naval training exercise off eastern Canada. Sci. Rep. 12, 1973. https://doi.org/10.1038/s41598-022-05930-4.

Tablado, Z., Jenni, L., 2017. Determinants of uncertainty in wildlife responses to human disturbance. Biol. Rev. 92, 216–233. https://doi.org/10.1111/brv.12224.

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

Tyack, P.L., Johnson, M., Soto, N.A., Sturlese, A., Madsen, P.T., 2006. Extreme diving of beaked whales. J. Exp. Biol. 209, 4238–4253. https://doi.org/10.1242/jeb.02505.

Tyack, P.L., et al., 2011. Beaked whales respond to simulated and actual navy sonar. PLoS One 6, e17009. https://doi.org/10.1371/journal.pone.0017009.

Tyack, P.L., et al., 2022. Managing the effects of multiple stressors on wildlife populations in their ecosystems: developing a cumulative risk approach. P. Roy. Soc. B 289. https://doi.org/10.1098/rspb.2022.2058.

von Benda-Beckmann, A.M., Isojunno, S., Zandvliet, M., Ainslie, M.A., Wensveen, P.J., Tyack, P.L., Kvadsheim, P.H., Lam, F.P.A., Miller, P.J.O., 2021. Modeling potential masking of echolocating sperm whales exposed to continuous 1–2 kHz naval sonar. J. Acoust. Soc. Am. 149, 2908–2925. https://doi.org/10.1121/10.0004769.

Wensveen, P.J., Kvadsheim, P.H., Lam, F.-P.A., von Benda-Beckmann, A.M., Sivle, L.D., Visser, F., Curé, C., Tyack, P.L., Miller, P.J.O., 2017. Lack of behavioural responses of humpback whales (*Megaptera novaeangliae*) indicate limited effectiveness of sonar mitigation. J. Exp. Biol. 220, 4150–4161. https://doi.org/10.1242/jeb.161232.

Wensveen, P.J., Thomas, L., Miller, P.J.O., 2015. A path reconstruction method integrating dead-reckoning and position fixes applied to humpback whales. Mov. Ecol. 3, 31. https://doi.org/10.1186/s40462-015-0061-6.

Wensveen, P.J., et al., 2019. Northern bottlenose whales in a pristine environment respond strongly to close and distant navy sonar signals. P. Roy. Soc. B 286, 20182592. https://doi.org/10.1098/rspb.2018.2592.

Whitehead, H., 2003. Sperm Whales: Social Evolution in the Ocean. University of Chicago Press.

Whitehead, H., Shin, M., 2022. Current global population size, post-whaling trend and historical trajectory of sperm whales. Sci. Rep. 12, 1–12. https://doi.org/10.1038/s41598-022-24107-7.

Whitehead, H., Smith, T.D., Rendell, L., 2021. Adaptation of sperm whales to open-boat whalers: rapid social learning on a large scale? Biol. Lett. 17. https://doi.org/10.1098/rsbl.2021.0030 rsbl.2021.0030.

Williams, T.M., Blackwell, S.B., Tervo, O., Garde, E., Sinding, M.S., Richter, B., Heide-Jørgensen, M.P., 2022. Physiological responses of narwhals to anthropogenic noise: a case study with seismic airguns and vessel traffic in the Arctic. Funct. Ecol. 36, 2251–2266. https://doi.org/10.1111/1365-2435.14119.

Wood, S.N., 2008. Fast stable direct fitting and smoothness selection for generalized additive models. J. R. Statist. Soc. B 70, 495–518. https://doi.org/10.1111/j.1467-9868.2007.00646.x.