AUTHOR QUERY FORM

ELSEVIER	Journal: MPB Article Number: 6238	Please e-mail or fax your responses and any corrections to: E-mail: corrections.eseo@elsevier.sps.co.in Fax: +31 2048 52799

Dear Author,

Please check your proof carefully and mark all corrections at the appropriate place in the proof (e.g., by using on-screen annotation in the PDF file) or compile them in a separate list. Note: if you opt to annotate the file with software other than Adobe Reader then please also highlight the appropriate place in the PDF file. To ensure fast publication of your paper please return your corrections within 48 hours.

For correction or revision of any artwork, please consult http://www.elsevier.com/artworkinstructions.

Any queries or remarks that have arisen during the processing of your manuscript are listed below and highlighted by flags in the proof. Click on the 'Q' link to go to the location in the proof.

Location in article	Query / Remark: click on the Q link to go Please insert your reply or correction at the corresponding line in the proof
Q1	Please confirm that given name(s) and surname(s) have been identified correctly.
Q2	The country name has been inserted for the affiliations 'a and e'. Please check, and correct if necessary.
Q3	Please provide more details for 'affiliations b and c'.
Q4	Please check whether the designated corresponding author is correct, and amend if necessary and also check the insertion of telephone number and email address.
Q5	The term 're' has been used throughout this article. Please check this meaning and correct if necessary.
Q6	Please check the hierarchy of the section headings.
Q7	Reference 'Wensveen (2012)' is cited in the text but not provided in the reference list. Please provide it in the reference list or delete this citation from the text.

Please check this box if you have no corrections to make to the PDF file ☐

Thank you for your assistance.

Highlights

• Pilot-whale sonar response thresholds higher than found for other cetaceans. • No effect of sonar frequency or previous exposures on probability of response. • US Navy dose–response underestimates probability of pilot-whales avoidance at long ranges.

Marine Pollution Bulletin xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Marine Pollution Bulletin

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

High thresholds for avoidance of sonar by free-ranging long-finned pilot whales (*Globicephala melas*)

R. Antunes[a,*], P.H. Kvadsheim[b], F.P.A. Lam[c], P.L. Tyack[a,d], L. Thomas[e], P.J. Wensveen[a], P.J.O. Miller[a]

[a] Sea Mammal Research Unit, Scottish Oceans Institute, University of St. Andrews, St. Andrews, Scotland KY16 8LB, UK
[b] Norwegian Defence Research Establishment (FFI), Norway
[c] Netherlands Organization for Applied Scientific Research (TNO), The Netherlands
[d] Biology Department, Woods Hole Oceanographic Institution (WHOI), Woods Hole, MA 94305, USA
[e] CREEM Centre for Research into Ecological and Environmental Modelling, University of St. Andrews, St. Andrews, Scotland KY16 9LZ, UK

ARTICLE INFO

Keywords:
Sonar
Pilot-whales
Marine mammals
Noise
Dose–response

ABSTRACT

The potential effects of exposing marine mammals to military sonar is a current concern. Dose–response relationships are useful for predicting potential environmental impacts of specific operations. To reveal behavioral response thresholds of exposure to sonar, we conducted 18 exposure/control approaches to 6 long-finned pilot-whales. Source level and proximity of sonar transmitting one of two frequency bands (1–2 kHz and 6–7 kHz) were increased during exposure sessions. The 2-dimensional movement tracks were analyzed using a changepoint method to identify the avoidance response thresholds which were used to estimate dose–response relationships. No support for an effect of sonar frequency or previous exposures on the probability of response was found. Median population response thresholds for avoidance (SPL_{max} = 179 dB re 1 μPa, SEL_{cum} = 183 dB re 1 μPa^2 s) were higher than previously found for other cetaceans. The US Navy currently uses a generic dose–response relationship to predict the responses of cetaceans to naval active sonar, which has been found to underestimate behavioural impacts on killer-whales and beaked-whales. The navy curve appears to match more closely our results with long-finned pilot-whales, though it might underestimate the probability of avoidance for pilot-whales at long distances from sonar sources.

1. Introduction

Sound propagates better in water than in air, and cetaceans have evolved sensitive hearing (e.g. Mooney et al., 2012). Sound is a primary sensory cue for cetaceans; they rely on sound for basic functions such as finding prey (e.g. Johnson et al., 2008), navigation (e.g. Verfuß et al., 2005), reproduction (e.g. Tyack, 1981), predator–prey interactions (e.g. Barrett-Lennard et al., 1996) and communication (e.g. King and Janik, 2013), making them particularly sensitive to disturbance caused by anthropogenic sounds. Military active sonar is amongst the most intense anthropogenic sound sources, with typical source sound pressure levels in excess of 220 dB re 1 μPa m (Ainslie, 2010) and has the potential to be detected over hundreds of kilometers of ocean. Several studies have reported avoidance (e.g. Buck and Tyack, 2000; Miller et al., 2012), injury and even mortality, caused by exposure to military sonar (Simmonds and Lopez-Jurado, 1991; NMFS, 2005; Claridge, 2001; Cox et al., 2006; Fernández et al., 2005; Nowacek et al., 2007; Parsons et al., 2008; Yang et al., 2008; D'Amico et al., 2009). Recognition of the potential of sound exposure to harm marine mammals has led legislators, international treaty bodies, environmental organizations and professional societies to express concern and to assess the potential adverse effects of anthropogenic sound in the ocean (e.g. ASCOBANS, 2006, 2009; ACCOBAMS, 2007; European Parliament and Council, 2008; IUCN, 2012; CMS, 2009; Dolman et al., 2011; Zirbel et al., 2011). The discovery of bubble-like lesions in the tissues of cetaceans that stranded following naval exercises suggested that auditory damage due to exposure to intense sounds was not the cause of death (Fernández et al., 2005). Investigation into the causes of these injuries suggested that changes in diving behavior could cause decompression sickness-like effects (Parsons et al., 2008; Kvadsheim et al, 2012; Fahlman et al., 2014). Understanding behavioral responses, which occur at lower sound levels than those that cause auditory damage, is critical for mitigation of the impacts of sonar on cetaceans (Parsons et al., 2008). Concerns about the effects of noise on cetaceans have shifted from an initial focus on direct mortality and physical injury to a broader range of sub-lethal

* Corresponding author. Tel.: +44 1334 208286.
E-mail address: rna@st-andrews.ac.uk (R. Antunes).

http://dx.doi.org/10.1016/j.marpolbul.2014.03.056
0025-326X/

and non-pathological effects such as reduction in feeding rates (Miller et al., 2009), reduction in fitness at the individual level and loss of habitat (Morton and Symonds, 2002). Dose–response relationships have been recognized as a useful management tool to evaluate the risk posed by the use of sonar by some of the world's navies (e.g. US Navy, 2008). Cetacean species have evolved diverse hearing capabilities and behavioral adaptations, and it is unrealistic to expect that a single dose–response relationship would fit all species. In particular, the species' hearing sensitivity at the frequency utilized by the sonar signal and behavioral responsiveness may affect the potential impact of the sound (Ellison et al., 2012).

Estimating dose–response relationships is common practice in toxicology, and is usually achieved by exposing groups of individuals to fixed doses and evaluating the proportion of individuals that are affected per dose.

Long-finned pilot whales have been reported to change diving (Sivle et al., 2012) and vocal behavior in response to sonar exposure (Rendell and Gordon, 1999; Alves et al., 2014). Here, we report a behavioral response study where we exposed long-finned pilot whales (*Globicephala melas*) to naval sonar signals within the 1–2 kHz (European LFAS: Low-Frequency Active Sonar) and 6–7 kHz (European MFAS: Mid-Frequency Active Sonar) frequency bands in order to investigate possible frequency effects on the response thresholds. A new method was used to quantify the dose threshold at which free-ranging long-finned pilot whales began to avoid an approaching vessel transmitting sonar. The method consists of two parts: statistical analysis of movement tracks to identify unusual change points indicating an avoidance response at a given threshold, and parameterizing a hierarchical Bayesian population-level dose–response model using the observed response thresholds. Although motivated by the study of anthropogenic disturbance caused by noise in the marine environment, this approach is generic and can be applied to other stimuli.

2. Methods

2.1. Experimental procedures

The experimental protocol is detailed in Miller et al. (2011, 2012) and summarized here. The experiments were conducted along the coast of Northern Norway between 66° and 70°N latitude in May/June of 2008 and 2009. Long-finned pilot whales were encountered in social groups of 3–35 individuals. These groups were approached in a small boat and one or more whales were instrumented with suction-cup attached archival tags (DTAGs; Johnson and Tyack, 2003). The DTAGs recorded pressure (20 Hz sampling rate, converted to depth using calibrated values) and stereo sound (192 kHz sampling rage).

The tagged whales' surfacings were tracked from an observation vessel (29 m MS Strønstad) aided by the VHF beacon on the tag. Observers on this vessel determined the tagged whale's position relative to the vessel approximately every 2 min. When multiple individuals in the same group were tagged, one was assigned as the focal animal, and sighting efforts were directed to it; non-focal tagged whales were also tracked whenever possible. Whale positions were determined from their azimuth relative to the bow of the vessel using a protractor with a sight, and measuring distance with a laser rangefinder or estimating distance by eye. The latitude and longitude of each sighting were calculated from the vessel's GPS position and heading measured by compass. The observation vessel stayed at least 400 m from the focal animal.

After a baseline (pre-exposure) period of 62–305 min, the whales were exposed to sonar signals transmitted from a naval sonar source (Socrates II; Kvadsheim et al., 2009) towed at a depth of 34–54 m astern of the source vessel (55 m FFI R/V H.U. Sverdrup II) moving at 3–4 ms^{-1}. Three types of sonar signals were played: LFAS-UP: 1–2 kHz hyperbolic upsweep, MFAS-UP: 6–7 kHz hyperbolic upsweep or LFAS-DO: 1–2 kHz hyperbolic downsweep. Maximum source levels were 214 dB re 1 μPa m (rms) for the 1–2 kHz band and 199 dB re 1 μPa m (rms) for the 6–7 kHz band. Sound transmissions were initiated when the source vessel was 6–8 km from the tagged whale and source levels were increased from 152 dB re 1 μPa m for LFAS and from 158 dB re 1 μPa m for MFAS to maximum level over a 10 min ramp-up period. Sound pulses (pings) were 1 s in duration (including two 50 ms cosine tapers at the start and end of each ping) and were transmitted every 20 s. The source vessel was steered toward the focal whale until a distance of 1 km, after which the course was fixed. The combination of source vessel approach and ramp-up of source level resulted in an escalation of sound pressure level (SPL) received by the focal whale. Transmissions stopped approximately 5 min after the source vessel passed the focal whale. During control approaches, the source vessel approached the whales in the same way, but no sonar was transmitted. Each tagged whale was exposed to 2–4 sonar and control exposure sessions, each separated by at least 55 min (Table 1).

These experiments were licensed under a permit provided by the Norwegian Animal Research Authority (Permit No. S-2007/61201), and were approved by the Univ. of St. Andrews Animal Welfare and Ethics Committee and the Woods Hole Oceanographic Institutional Animal Care and Use Committee. A mitigation protocol was in place during the experiments, calling for cessation of sound transmission if whales came within 100 m of the source, or if observed behavioral reactions posed a great risk to the exposed animals.

2.2. Measurements of sonar dose

Following the recommendations of Southall et al. (2007) for behavioral response studies on marine mammals, we quantified the sonar dose in terms of maximum sound pressure level (SPL_{max}; dB re 1 μPa, rms) and cumulative sound exposure level (SEL_{cum}; dB re 1 μPa^2 s) in the same way as Miller et al. (2014). The lack of hearing sensitivity values for this species at frequencies <4 kHz (Pacini et al., 2010) precluded the use of sensation levels for comparison between the 1–2 kHz and 6–7 kHz bands. However, the effect of sonar frequency band was included as a potential covariate in the dose–response model (see below).

We investigated changes in horizontal movement potentially caused by exposure to the source/vessel. We calculated the focal whale's horizontal speed from the sighting positions as:

$$v_j = \frac{d(s_{j-1}, s_j) + d(s_j, s_{j+1})}{t_{j+1} - t_{j-1}} \quad (1)$$

where $d(a, b)$ is the distance between the positions a and b, s_j is the position of the sighting at time t_j. Heading was calculated as the azimuth between one sighting and the next. Heading was decomposed into orthogonal components Easting and Northing that were linearly interpolated onto a 1 min grid.

2.3. Mahalanobis distance change-point analysis

We developed a generic multivariate change-point analysis for time-series of multivariate data to identify behavior changes. The magnitude of change in a time-series of multivariate data was calculated as the mean pairwise Mahalanobis distance (Mahalanobis, 1936) (D_{AB}) between adjacent windows (A and B, with n data points each):

$$D_{AB} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}\sqrt{(a_i - b_j)^T S^{-1}(a_i - b_j)}}{n^2} \quad (2)$$

Please cite this article in press as: Antunes, R., et al. High thresholds for avoidance of sonar by free-ranging long-finned pilot whales (*Globicephala melas*). Mar. Pollut. Bull. (2014), http://dx.doi.org/10.1016/j.marpolbul.2014.03.056

Table 1
Summary of sonar exposures to long-finned pilot whales with results from Mahalanobis changepoint analysis. The highest sonar dose measured just before the time of changepoint is given as SPL_{max}, SEL_{cum} and distance to source. Statistically significant changepoints are indicated in bold. Max dive depth indicates the maximum depth of the dive at the time of the changepoint.

		Start	End	Baseline duration (min)	Time of changepoint	SPL_{max} (dB re 1 μPa)	SEL_{cum} (dB re 1 μPa^2 s)	Distance (km)	Statistic	p	Max dive depth (m)
gm08_150c	MFAS-UP	16:12:00	16:50:01	62	**16:21**	**115**	**118**	**6.087**	**2.926**	**<0.0001**	**25**
	LFAS-UP	18:05:00	18:36:01		**18:22**	**159**	**162**	**2.226**	**3.853**	**<0.0001**	**26**
gm08_154d	LFAS-UP	01:15:00	02:35:01	129	[a]	[a]	[a]	[a]	[a]	[a]	
	MFAS-UP	03:35:00	04:00:01		03:35	70	67	6.000	2.403	0.4894	10
gm08_158b	SILENT	14:27:20	15:15:41	117	15:14	NA	NA	0.993	1.941	0.8137	NA
	LFAS-UP	16:15:00	16:51:01		16:51	171	179	1.139	2.327	0.6483	NA
	MFAS-UP	17:50:00	18:23:01		18:00	123	128	4.587	1.342	0.5072	NA
gm08_159a	SILENT	23:07:00	23:37:21	134	23:37	NA	NA	1.105	1.440	0.9499	405
	LFAS-UP	00:33:00	01:08:01		**01:00**	**160**	**168**	**1.239**	**2.391**	**<0.0001**	**14**
	MFAS-UP	02:10:00	02:46:01		**02:11**[b]	**80**	**79**	**7.837**	**2.617**	**<0.0001**	**429**
gm09_138a	LFAS-UP	14:42:00	15:14:01	193	15:02	156	166	1.610	2.279	0.4748	20
	MFAS-UP	16:40:00	17:15:01		16:59	123	130	2.269	3.018	0.4694	10
	SILENT	18:40:00	19:14:01		19:02	NA	NA	1.616	3.245	0.2602	441
	LFAS-DN	20:32:00	21:05:01		20:32	72	66	7.100	1.886	0.5377	15
gm09_156b	SILENT	23:30:00	00:02:01	305	00:02	NA	NA	1.495	2.364	0.2742	545
	LFAS-UP	01:36:00	02:09:01		**01:54**	**157**	**165**	**1.234**	**3.114**	**<0.0001**	**22**
	MFAS-UP	03:10:00	03:37:01		03:35	156	162	0.786	1.397	0.5156	548
	LFAS-DN	04:55:00	05:25:01		**05:10**	**159**	**168**	**1.507**	**3.374**	**<0.0001**	**546**

[a] Premature tag release during exposure session, with change of focal individual.
[b] Judged not to have been caused by sonar exposure (see text).

where a_i is the i^{th} datum in window A, b_j is the j^{th} datum in window B and S^{-1} is the inverse of the variance-covariance matrix of the whole data set. During the dose escalation period, the position of the maximum value of D_{AB} ($\max D_{AB}$) was taken as the time of the largest behavioral change (Fig. 1). To evaluate whether behavioral changes were likely to have been in response to the exposure stimulus, the probability of a change occurring during baseline was calculated. We compared the magnitude of $\max D_{AB}$ to maximum levels over identical time windows during the baseline period (between the tag boat leaving the whales and the first sonar exposure) using a randomization procedure. At each random iteration, a mock exposure period (with the same duration as the actual exposure period) was randomly placed within the baseline period. The magnitude of largest change in D_{AB} in each mock exposure ($\max D_{AB}^{mock}$) was identified in the same way as for the actual exposure period. The proportion of $\max D_{AB}^{mock}$ randomizations that exceeded the observed changepoint value $\max D_{AB}$ (p) is a measure of how unusual the changepoint observed during the exposure was, given the natural levels of variation during baseline behavior. Low p-values were likely to have been caused by the exposure. We carried out an assessment of this method by simulation, and present the results in supporting information, Appendix A.

We used speed, Easting and Northing position data from each exposure session and baseline periods for each whale as input into the multivariate changepoint analysis, using 5 min time windows. The changepoints that occurred during sonar exposure and had

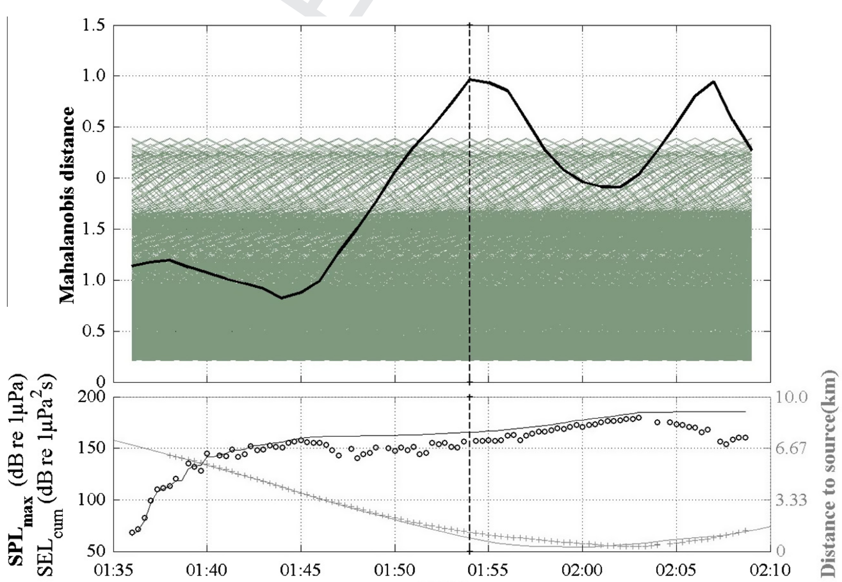


Fig. 1. Example of changepoint analysis for pilot-whale sonar exposure session (gm09_156b LFAS-UP). Top panel shows variation in Mahalanobis changepoint statistic during sonar exposure (black line) as well as variation in the same statistic for 10,000 mock exposures during baseline (gray lines). Bottom panel shows sound pressure (circles: SPL_{max}) and sound exposure levels (dashed line: SEL_{cum}), estimated distance between the sound source and the whale subject (gray line). Vertical dashed line indicates the time of the largest behavioural change as estimated by the changepoint analysis. Changepoint analysis for other exposures are shown in supporting information in Appendix B.

Please cite this article in press as: Antunes, R., et al. High thresholds for avoidance of sonar by free-ranging long-finned pilot whales (*Globicephala melas*). Mar. Pollut. Bull. (2014), http://dx.doi.org/10.1016/j.marpolbul.2014.03.056

$p < 0.003$ (0.05 with a Bonferroni correction for $N = 17$ exposure sessions) values from randomization tests were considered responses to the sonar and were used for dose–response relationship estimation. Exposure sessions with p-values above the significance threshold of 0.003 were considered as no-response exposure sessions. The highest SPL_{max} and SEL_{cum} before each of the changepoints were taken as the dose eliciting the response to sonar.

The number of exposure sessions with and without responses between sonar exposures and control approaches were also compared using a Barnard's test (Barnard, 1945).

2.4. Dose–response relationship

A hierarchical Bayesian model was used to estimate the dose–response relationships from response thresholds obtained from the Mahalanobis changepoint. Bayesian analysis provides a framework to combine prior information with data that is convenient for setting hierarchical models. Bayesian models are also robust to small sample sizes, such as those often encountered in dose escalation studies, and provide measures of parameter uncertainty that are directly interpretable in probabilistic terms. The model was fitted using Markov Chain Monte Carlo (MCMC) simulation with the software JAGS, version 3.2.0 (Plummer, 2003), via the rjags library in R, version 2.13 (R Development Core, 2013). The model assumes that each individual i has an expected response threshold μ_i and that the distribution of individual thresholds in the population is normal with a population average threshold μ and variance between individuals of ϕ^2:

$$\mu \sim N(\mu, \phi^2), \quad r_{min} \leq \mu_i \leq r_{max}. \tag{3}$$

The model further assumed that there is a minimum acoustic dose r_{min} below which no individual responds and a maximum acoustic dose r_{max} at which all individuals have responded; therefore the distribution of thresholds is truncated to the range r_{min} to r_{max}.

Each exposure session was coded by sonar stimulus type and whether or not it was the first exposure session for that animal. We accounted for any effects of these factors by assuming that the expected threshold for individual i during the exposure session j, μ_{ij} depends on the expected threshold of each individual μ_i, on the stimulus type and whether it has been previously exposed:

$$\mu_{ij} = \mu_i + \gamma_{order} \cdot \beta_{order} \cdot I(exposure) + \gamma_{stimulus} \cdot \beta_{stimulus} \cdot I(type). \tag{4}$$

Here β_{order} and $\beta_{stimulus}$ are parameters describing the effect of previous exposures and of stimulus type on the threshold respectively. $I(exposure)$ indicates if the individual has previously been exposed to sonar (0 for the first exposure and 1 for the subsequent exposures). $I(type)$ indicates stimulus type (0 for 1–2 kHz signal and 1 for 6–7 kHz signal). Gibbs Variable Selection (GVS, O'Hara and Sillanpää, 2009) was applied to assess the level of support for including β_{order} and $\beta_{stimulus}$ in the final dose–response model. In this procedure, two binary variables γ_{order} and $\gamma_{stimulus}$ were used to switch on/off the effects of the β terms in μ_{ij}. The proportion of posterior MCMC samples where γ_{order} and/or $\gamma_{stimulus}$ equal one is an estimate of the posterior model probability for models containing the parameters β_{order} and $\beta_{stimulus}$, respectively. In other words, it indicates the support in the data for the inclusion of parameters describing the effect of multiple exposure and stimulus type, respectively.

The actual response threshold for individual i, during exposure session j, r_{ij} is assumed to depend upon the individual whale's expected threshold and within-individual between-session variance σ^2, assumed to be constant for all individuals:

$$r_{ij} \sim N(\mu_{ij}, \sigma^2), \quad r_{min} \leq r_{ij} \leq r_{max} \tag{5}$$

An observation model is used because doses are usually presented in steps in the escalation procedure and the data are often collected using discrete sampling intervals, while the range of thresholds is assumed to be continuous. We assume that the observed responses o_{ij} have measurement error which is modeled as:

$$o_{ij} \sim N(r_{ij}, \varepsilon^2), \tag{6}$$

where ε is the standard deviation.

Escalation experiments in which no response was observed within the accomplished dose escalation range $[L_{ij}, U_{lj}]$ were assumed to be right censored (Plein and Moeschberger, 2003), where the response threshold is assumed *a priori* to be equally likely between U_{lj} and r_{max}.

Model parameters were estimated using 100,000 MCMC samples, after a burn-in of 10,000 (convergence of 3 MCMC chains was found to be rapid, so that such a burn-in is highly conservative).

The model was fitted assuming that an acoustic dose (in terms of both SPL_{max} and SEL_{cum}) below 60 dB is barely audible (Pacini et al., 2010; Schlundt et al., 2011) and will not cause a behavioral response and that all animals will avoid a sound source at an acoustic dose of 200 dB. We therefore chose a uniform prior distribution between $r_{min} = 60$ and $r_{max} = 200$ dB for the population mean (μ_{SPL} and μ_{SEL}). Uniform priors between 0 and 30 dB were used for both the between (ϕ_{SPL} and ϕ_{SEL}) and within-whale variation (σ_{SPL} and σ_{SEL}), considering that setting these values to 30 dB yields a probability density covering most of the range between 60 and 200 dB. We chose priors for β_{MFAS} and $\beta_{exposed}$ of $N(0,30)$ dB. Based on the calibration error of DTAG hydrophone sensitivity (s.d. = 2.5 dB re 1 μPa^{-1}), the measurement standard deviation ε was set to 2.5 dB for both SPL_{max} and SEL_{cum}. Priors for γ_{order} and $\gamma_{stimulus}$ were a Bernoulli distribution with $p = 0.5$.

3. Results

Six long-finned pilot whales instrumented with DTAGs were subjects in a total of 14 sonar exposure sessions (6 MFAS, 6 LFAS-UP and 2 LFAS-DO) and 4 control approaches (Table 1). Received SPL_{max} levels ranged from 68 to 180 dB re 1 μPa for the LFAS band and 70–161 dB re 1 μPa for the MFAS band. The maximum received levels measured in each exposure session ranged 163–180 dB re 1 μPa for the LFAS band and 150–161 dB re 1 μPa for MFAS. The closest approach distances ranged 0.14–0.47 km (mean 0.30 km) for the LFAS, 0.04–0.56 km (mean 0.23 km) for MFAS and 0.10–0.31 (mean 0.23) for control.

The changepoint statistic from the simulations (Appendix A) produced peaks that were associated with the simulated behavioral change points indicating that the Mahalanobis changepoint analysis can identify changes in autocorrelated time series of covarying variables.

Detailed results of the Mahalanobis changepoint analysis of the sonar exposures are described in Appendix B. Overall, the changepoint analysis highlighted six exposure sessions with responses likely to have been caused by sonar exposure. One of these was considered not to have been a response to sonar as it commenced before the start of sonar transmissions (gm08_159a MFAS-UP; Appendix B) and we found avoidance responses likely caused by exposure in one out of six (17%) MFAS-UP, three out of five (60%) LFAS-UP and one out of two (50%) LFAS-DN exposure sessions. In three (gm08_150 MFAS, gm08_150 LFAS-UP and gm08_150 LFAS-DN) out of the five sonar responses identified by the changepoint analysis, the behavioral changes consisted in heading changes away from the source vessel's position (160–180° relative to the source position) and to 0–30° relative to the heading of the source vessel (Fig. 2). However these responses did not last longer than

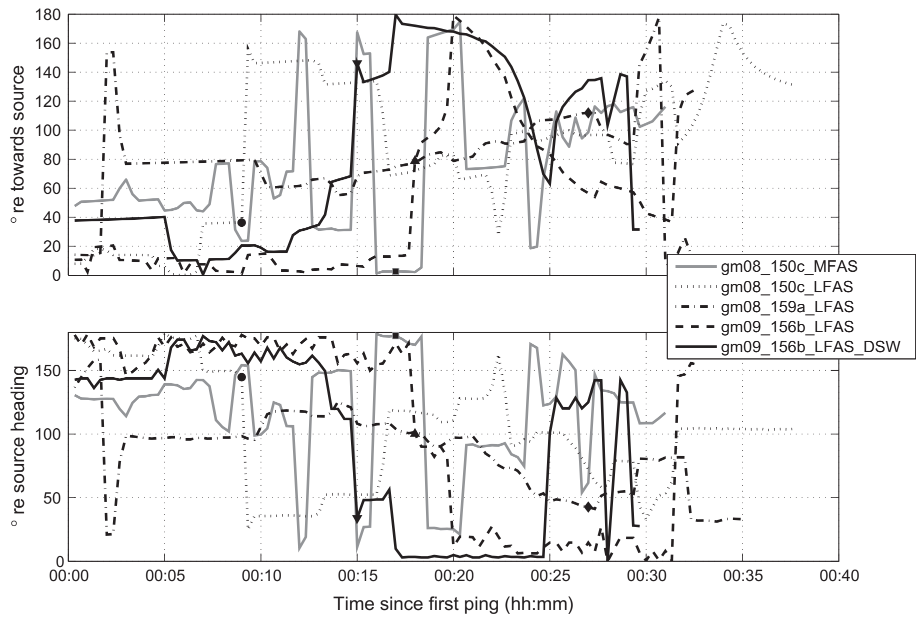


Fig. 2. Change in heading relative to the source for the five exposure sessions where responses to sonar were identified. Top panel shows whale heading relative to heading towards the source position and bottom panel shows whale heading relative to the source heading. Symbols mark the time of response as identified by the changepoint analysis. The values shown are the interpolated values as used in the Mahalanobis changepoint analysis.

the duration of the sonar exposure, and the whales appeared to have returned to previous movement patterns once the sonar ceased transmitting.

A change in movement behavior likely caused by the ship was not identified in any of the control approaches (with/without response = 0/4), a result that contrasts sharply with those for sonar exposure (with/without response = 5/8). Although the Barnard's test did not indicate independence between the sonar exposure and control approach responses at 0.05 significance level (Wald statistic = 1.4763; nuisance parameter = 0.9201; $p = 0.1096$) the observed difference suggests an effect of the sonar exposures.

The GVS procedure showed the strongest support for including β_{order} and excluding $\beta_{stimulus}$, but the level of support for this combination of variables was low ($p = 0.41$ for both SPL_{max} and SEL_{cum}). This indicates that there is little information in the data about whether these factors are important (Table 2). We therefore fitted a simpler model without the β terms of Eq. (4), resulting in estimated expected thresholds of 178.6 (95% CI: 155.2–198.5) dB re 1 μPa for SPL_{max} and 182.6 (160.2:199.1) dB re 1 μPa^2 s for SEL_{cum} (Table 3). Estimates for ϕ_{SPL}, σ_{SPL}, ϕ_{SEL} and σ_{SEL}, were close to 20 dB and had marginal posterior densities that appeared to have been constrained by the upper limit of the prior distribution (see Supporting information in Appendix C). Percentiles for values of probability of response for this model fit are given in supporting information of Appendix B.

Table 2
Proportion of posterior MCMC samples that supported the inclusion/exclusion of the effects of previous sonar exposure (β_{order}) and type of sonar stimulus ($\beta_{stimulus}$) in the dose–response model.

	SPL			SEL		
	Including $\beta_{stimulus}$	Excluding $\beta_{stimulus}$	Σ	Including $\beta_{stimulus}$	Excluding $\beta_{stimulus}$	Σ
Including β_{order}	0.23	0.41	0.64	0.25	0.41	0.66
Excluding β_{order}	0.12	0.24	0.36	0.11	0.23	0.34
Σ	0.35	0.65		0.36	0.64	

4. Discussion

The Mahalanobis changepoint analysis identified behavioral changes in movement during the sonar exposures. Within a dose

Table 3
Estimated dose–response model parameters for pilot-whale sonar exposure experiments, using both SPL_{max} (dB re 1 μPa) and SEL_{cum} (dB re 1 μPa^2 s) as dose. Mean, median, standard deviation and 95% credibility intervals for the marginal posterior densities are shown for each of the model parameters. Results are shown for full model and for model fit without effects of previous sonar exposure (β_{order}) and type of sonar stimulus ($\beta_{stimulus}$).

	Full dose–response model				Dose–response model without β effects			
	Mean	Median	St. dev.	95% CI	Mean	Median	St. dev.	95% CI
μ_{SPL}	173.2	173.4	14.4	144.4:198.0	178.6	178.8	11.8	155.2:198.5
ϕ_{SPL}	19.0	20.3	7.5	2.4:29.5	19.1	18.1	7.7	1.9:29.4
σ_{SPL}	20.1	20.4	5.8	8.9:29.5	20.8	20.9	5.2	10.9:29.5
$\beta_{stimulus}$	−5.4	−5.7	15.5	−35.2:25.9				
β_{order}	22.7	22.1	14.3	−4.4:52.7				
μ_{SEL}	179.0	177.9	13.4	149.9:198.7	182.6	183.4	10.8	160.2:199.1
ϕ_{SEL}	18.5	19.7	7.7	2.0:29.5	17.6	18.4	7.7	1.8:29.3
σ_{SEL}	20.0	20.2	5.8	8.1:29.4	20.5	20.5	5.2	10.9:29.4
$\beta_{stimulus}$	−6.1	−6.5	15.3	−35.6:24.8				
β_{order}	23.4	22.7	14.6	−4.2:54.0				

escalation context it might be possible that thresholds calculated using the time of the maximum value of the Mahalanobis distance statistic will be somewhat later in time than the onset of the response, which would result in thresholds being higher than that needed to initiate a response. In those exposure sessions considered to include a response to sonar and for which high-resolution tag data were available, we inspected the magnetic heading data to identify the onset of avoidance. The onset of avoidance identified from tag data differed by less than 30 s from those identified by the changepoint analysis, which would not influence the dose as specified for the onset of response. Although we cannot ascertain this in one case where sensor data was not available due to tag failure (gm08_158b), these observations indicate that any bias due to the discretization in changepoint analysis is likely to have been very small.

Changes in behavior are expected also in the absence of sonar exposure and we were interested in evaluating whether behavioral changes were more likely in response to the sonar exposure. Our randomization approach relied on the assumption that responses to sonar will be extreme in comparison to periods of similar duration without sonar. We have taken this approach for several reasons. Data on behavioral patterns in long-finned pilot whales with a high temporal resolution are scarce. *A priori* knowledge of behavioral patterns allows the identification of patterns of change potentially caused by disturbance and their biological implications (e.g. relative energetic costs of different behaviors and their disruption) and the formulation and testing of particular hypothesis in terms of responses to disturbance. Given the current lack of knowledge about the biological consequences of particular behavioral changes, we have chosen to apply the Mahalanobis changepoint analysis as an objective means of identifying changes in behavior.

Our decision about whether an observed behavioral change was caused by a stimulus was based on a comparison between the magnitude of the observed change and the magnitude of the behavioral changes observed during baseline. In this respect our analysis is not conservative and we may have missed changes in behavior that were caused by sonar exposure but fell within baseline variation. On the other hand, we expect that we would be able to detect any dramatic changes and the biological significance of undetected responses is probably limited. Nevertheless, our conclusions may be limited by the amount of baseline data available. Future studies would benefit from longer baseline periods for a better evaluation of behavioral patterns.

Our inability to identify any response to the control approaches suggests that the observed responses were caused by the sonar exposure and not by the approaching ship; however, the limited sample size provided only modest statistical support for this.

The Mahalanobis changepoint analysis objectively highlighted changes that were unusual given baseline variation; the method does not identify the form of behavioral change and required additional interpretation. In the case of horizontal movement analyzed here, the method could not distinguish between avoidance and attraction if both were characterized by unusual changes in the movement parameters, and our analysis relied upon some posterior interpretation of the detailed patterns of change. Observations of the change in heading relative to the source revealed that the identified changepoints were associated with characteristic patterns of avoidance where animals changed their heading so as to move between 0° and 45° relative to the source heading (i.e. parallel or acute angle to the source movement) and between 140° and 180° relative to the direction to the source (i.e. away from the source position). These avoidance patterns match those identified for other species in cases when the stimulus moves faster than or at similar speed as the source of disturbance (Domenici et al., 2011; Miller et al., 2014). This is consistent with the range of speed of the source and the animals during our experiments (source vessel approaching at 3–4 ms^{-1} and whales moving at approximately at 1–3 ms^{-1} except during occasional "sprints").

The observed avoidance responses of the pilot whales did not last beyond the sonar exposure session, suggesting a low impact of our experiments (Miller et al., 2012). The duration of the avoidance responses observed during sonar exposure in pilot whales was also shorter than for some other species. Tyack et al. (2011) reported Blainville's beaked whales (*Mesoplodon densirostris*) avoiding an area of several hundred square kilometers for several days during a sonar exercise. DeRuiter et al. (2013) reported Cuvier's beaked whales (*Ziphius cavirostris*) and Miller et al. (2012, 2014) reported killer whales (*Orcinus orca*) stopping feeding and maintaining high-speed avoidance for extended periods. The latter result was also consistent with a reduction in killer-whale sighting frequency during real naval sonar exercises (Kuningas et al., 2013). Actual sonar exercises often involve the use of multiple sources for longer periods and exposing wider areas than our short exposure sessions; they therefore have the potential to cause longer-lasting responses and higher impact than short exposure experiments. The relatively short responses observed might thus be a feature of our experiments which are designed to identify thresholds for onset of response, while minimizing the potential for adverse impact. On the other hand even in real naval exercise scenarios, the duration of exposure to levels above the high response threshold observed for pilot whales will likely be relatively short. Additional observations during actual sonar exercises are necessary to fully evaluate the impacts of operational sonar usage.

Houser et al. (2013) exposed bottlenose dolphins (*Tursiops truncatus*) to sonar signals in a captive setting and found habituation at received SPL $\leqslant$ 160 dB re 1 μPa but not at SPL $\geqslant$ 175 dB re 1 μPa. However, the captive dolphins used might not be an accurate model for naive wild animals as these captive, trained dolphins live in a noisy harbor, were trained using operant conditioning and have been used in multiple noise exposure experiments (Houser et al., 2013). Although we found some support for pilot whales having higher avoidance response thresholds during later sonar exposure sessions than during the first exposure session (i.e. habituation), this was not conclusive, possibly due to the small sample size. Subsequent studies and/or meta-analysis would benefit from additional data to elucidate the effects of sequential exposures to sonar. The most important question from a policy perspective is whether whales become less responsive or more responsive during longer exercises or repeated exposures.

The observed response thresholds occurred at higher levels than described for other cetacean species. Miller et al. (2014) fitted the same Bayesian dose–response model to the response thresholds of killer whales exposed to sonar signals and estimated an expected (± s.d.) response threshold of 142 ± 15 dB re 1 μPa for SPL_{max} and 149 ± 16 dB re 1 $\mu Pa^{2\circ}s$ for SEL_{cum}. Blainville's and Cuvier's beaked whales exposed to naval sonar showed SPL response thresholds below 142 dB re 1 μPa (Tyack et al., 2011) and 89–127 dB re 1 μPa (DeRuiter et al., 2013), respectively.

The observed avoidance responses in pilot whales were restricted to the duration of sonar exposure. This also contrasts with the responses reported for killer whales and beaked whales where some responses lasted longer than the sound exposure. These differences indicate that long-finned pilot whales are less sensitive to sonar exposure, compared with these species. Also, it indicates that a generic dose–response relationship for all odontocetes is not adequate for mitigation and impact assessment of sonar use, and that taxon specific dose–response relationships are necessary.

Behavioral responses of marine mammals to sound stimuli often are strongly affected by the context of the exposure, which implies that species and the received sound level alone is not enough to predict type and strength of a response (Southall et al., 2007; Miller et al., 2012; DeRuiter et al., 2013; Goldbogen et al., 2013). High levels of unexplained within-individual variability in our model imply that observed response thresholds depended on contextual variables that are yet to be determined and/or have not been included in the model. Contextual variables are important and should be included in the assessment of the effects of noise on marine mammals (Ellison et al., 2012). The limited sample size of our dataset precludes a robust statistical analysis of several contextual variables with our response model, but the model is flexible and can be extended to include additional variables.

A more detailed analysis of contextual effects is required for extrapolating our assessment of the impacts of sonar exposure to long-finned pilot whales in other settings and seasons, to reduce the uncertainty associated with the current model, and possibly derive dose–response curves specific to other types of response. Also, basic understanding of the biology of long-finned pilot whales is lacking and further studies will provide insight into the biological significance of the responses observed during sonar exposure (e.g. energetic costs), which is of major importance for management and mitigation of sonar exposure.

The dose–response function currently used by the US Navy (US Navy, 2008) appears to be generally better suited for the prediction of the impacts of sonar exposure to pilot whales than to killer whales and beaked whales. Nonetheless there are still differences in shape between our dose–response relationship and the US Navy curve. Our curve predicts a higher probability of response for received levels <165 dB re 1 μPa and a lower probability of response for >165 dB re 1 μPa (Fig. 3). This indicates that mitigation using the US Navy curve is conservative for high received levels but underestimates the impact at lower received levels. Since the sound exposure area/volume increases at larger ranges, the number of animals impacted with lower received levels is potentially higher. With this effect in mind we further compared our dose–response relationship with the US Navy curve by calculating an impact index defined by the radial distance from the source, r:

$$I(r) = \int_0^r 2\pi \cdot r \cdot p_{RL}(SL - 20 \cdot log_{10}[r] - \alpha \cdot r)dr \qquad (7)$$

where p_{RL} is the probability of response as a function of received level, given by the dose–response relationship. If animals were evenly distributed, this index would indicate how many animals are affected cumulatively with distance r, by integrating both the effects of the dose–response relationship and the increase in area

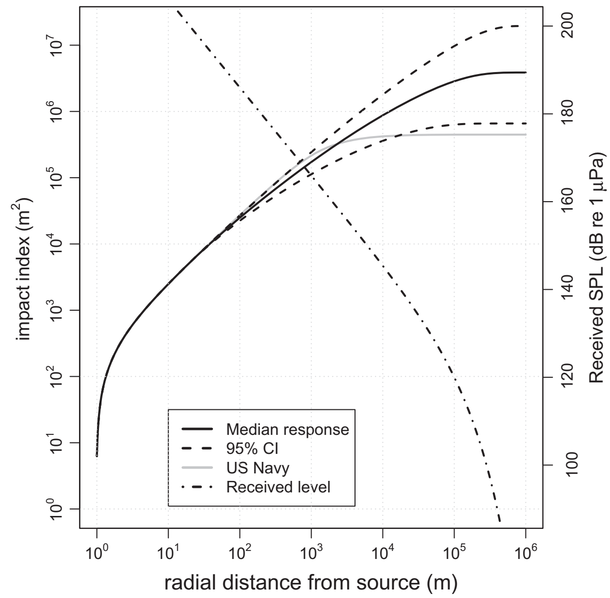


Fig. 4. Impact index (see text) of the estimated pilot-whale dose–response relationship and the curve currently in use by the US Navy (U.S. Navy EIS, 2008), for circular areas defined by radial distance from the source (r), assuming a source level of 226 dB re 1 μPa, transmission loss given by $20 \cdot \log_{10}(r)$ and 0.06 dB km^{-1} absorption loss. Also shown are the received SPLs at each distance.

with distance from the source. For comparison purposes we calculated this index for a realistic operational sonar SL of 226 dB re 1 μPa, using a transmission loss defined by a simple spherical spreading model ($20 \cdot \log_{10}[r]$) and attenuation loss of 0.06 dB km^{-1} (Fig. 4). Under these conditions both curves predict similar levels of impact up to 1 km but the US Navy dose–response relationship predicts little increase in impact at ranges >10 km while our curve predicts impacts at least one order of magnitude higher and increasing up to ranges beyond 100 km.

Impact assessment of more realistic sonar exposure scenarios can be carried out using a similar integrating approach but with added complexity, taking into account source characteristics, local sound propagation conditions, bathymetry/coastline and estimates of animal density and habitat preference, together with its uncertainties. This could potentially be implemented with computer programs running on board navy ships for real time assessment

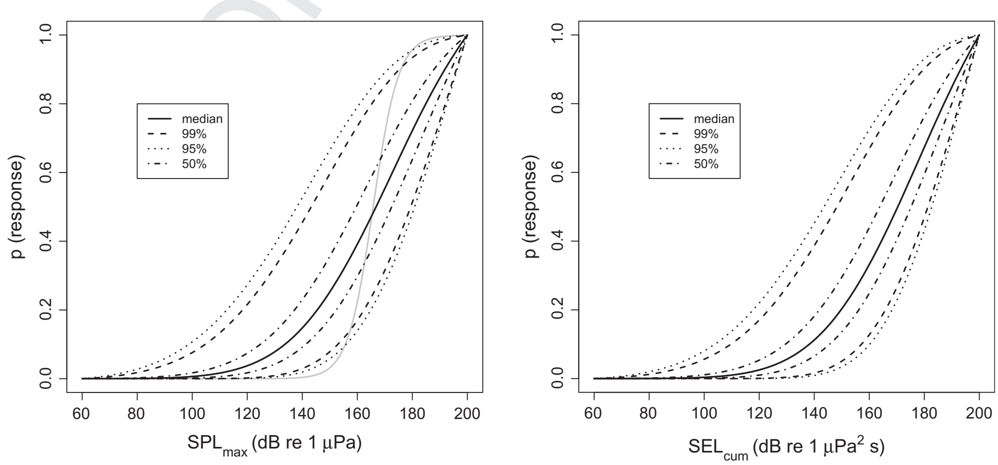


Fig. 3. Estimated dose–response relationship for avoidance responses of pilot whales as a function of SPL_{max} (left panel) and SEL_{cum} (right panel). Dashed lines show the 50%, 95% and 99% posterior credibility intervals. The dose–response curve currently in use by the US Navy for SPL (U.S. Navy EIS, 2008) is also shown overlaid (in grey) on the left panel.

of impacts. These types of analysis can also be used in management and mitigation of sonar exposure to establish the limits of impact area/volume, if thresholds for the maximum acceptable numbers of affected individuals are set.

Dose–response relationships provide the link between sonar exposure and the animals' responses that is necessary for the assessment of population level consequences. While early efforts produced a single generic dose–response relationship for marine mammal impact assessment and mitigation, current research supports the idea that species and context specific dose–response information is necessary. Our results provide the first dose–response relationship for exposure of long-finned pilot whales to sonar signals, providing an important basis for assessment and management of impacts for this species.

Acknowledgements

Thanks to the ship and science crews that supported our field efforts. The sonar exposure experiments were funded by the US Office of Naval Research (ONR) with grants to the 3S and MOCHA projects, the Royal Norwegian Navy and the Norwegian Ministry of Defence, the Netherlands Ministry of Defence and by WWF Norway. P.L.T. acknowledges the support of the MASTS pooling initiative (The Marine Alliance for Science and Technology for Scotland) in the completion of this study. MASTS is funded by the Scottish Funding Council (Grant Reference HR09011) and contributing institutions. P.J.W. was supported with a studentship of the VSB Foundation (Grant Number VSB.08/228-E).

Appendix A. Evaluation of the Mahalanobis changepoint analysis through simulation

A.1. Method

To assess the performance of the Mahalanobis changepoint analysis, we simulated avoidance responses in the context of sonar exposure to whales. Simulated individual whales were exposed one at a time to an approaching sound source. The movement of both the whale and source were simulated in a Cartesian space for 1500 iterations. Each simulated exposure session was initialized with the whale at position $(x_w, y_w) = (0, 0)$ and the source placed so that the distance between the whale and the sound source d was 8 km and the angle between the course of the source and the bearing from source to whale was set randomly between $-1/3\pi$ and $1/3\pi$. During the simulation, the source moved at a constant step length of 80 m. The direction of movement of source (v was towards the simulated whale's position while $d > 1$ km and fixed thereafter. Simulated transmissions started at $d = 7$ km and lasted for 120 iterations. This approaching behavior, including the angular range of approach was chosen to closely match the actual sonar experiments conducted (see below). The sonar source level (SL) was linearly increased from 140 to 210 dB re 1 μPa m in 30 pings and remained constant at 210 dB re 1 μPa m for 90 pings thereafter.

In each simulated exposure session, the whale moved according to a biased random walk model defined by the step length l and the angle θ between consecutive positions. At each iteration i, the values for l were randomly drawn from a Weibull distribution with scale and shape parameters λ and k, respectively. The values for θ

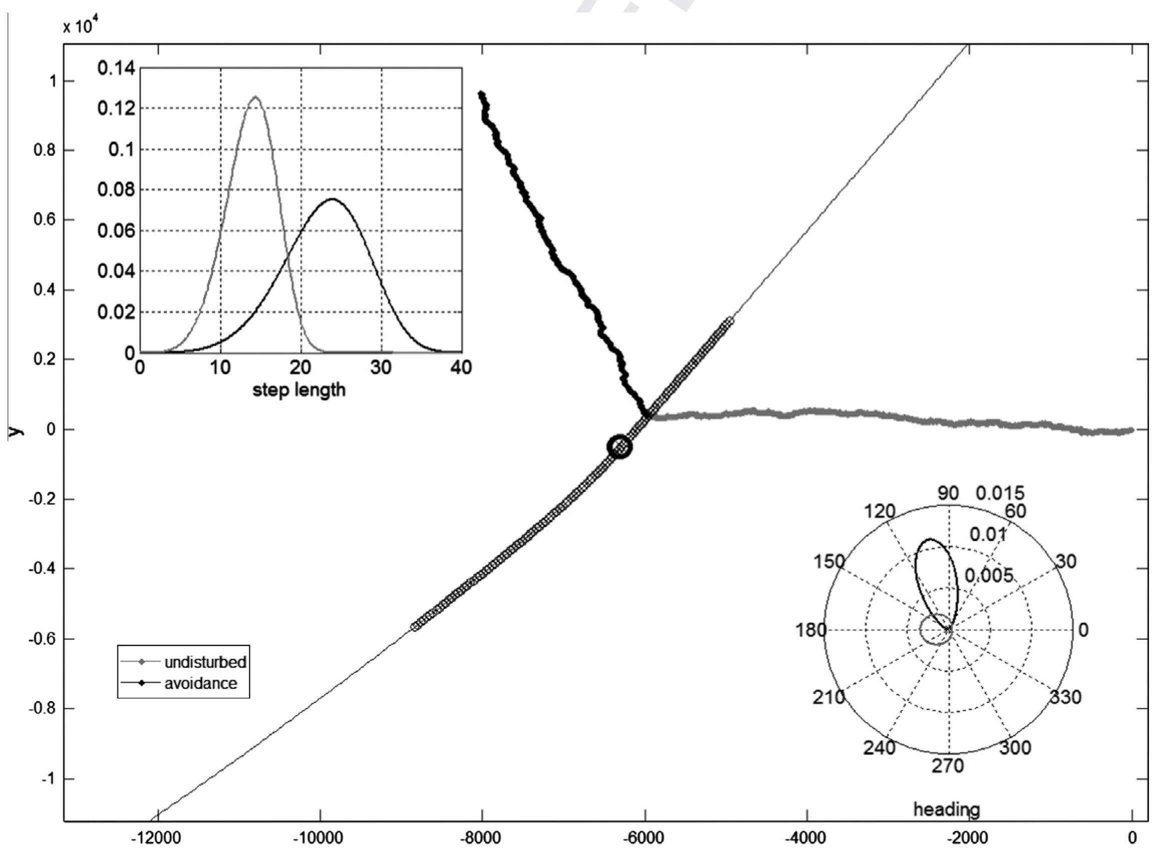


Fig. A1. Example of simulated sonar approach and distribution of movement parameters. Black line shows simulated movement during the undisturbed state and the gray line shows movement during the avoidance state. The simulated source vessel is approaching from the lower-left corner, with simulated sonar transmissions indicated by open circles on the ship track. Inset figures show distance travelled (upper left) and angle (lower right) distributions for undisturbed and avoidance states used in the simulation. Circle on the approaching source vessel's track indicates position at the time of change in state.

Please cite this article in press as: Antunes, R., et al. High thresholds for avoidance of sonar by free-ranging long-finned pilot whales (*Globicephala melas*). Mar. Pollut. Bull. (2014), http://dx.doi.org/10.1016/j.marpolbul.2014.03.056

were randomly drawn from a von Mises distribution with mean and dispersion parameters ω and γ respectively. The whale could be in one of two behavioral states each corresponding to a different set of movement parameters – undisturbed: $[\lambda_u, k_u, \omega_u, \gamma_u]$ and avoiding: $[\lambda_a, k_a, \omega_a, \gamma_a]$. The undisturbed state was characterized by slower movement (smaller step lengths) and wider turns centered at π: $[\lambda_u, k_u, \omega_u, \gamma_u] = [5, 15, \pi, 1]$. The avoiding state was characterized by faster movement and turns centered at an angle determined by heading of the source ν at the time of switching to the avoiding state: $[\lambda_a, k_a, \omega_a, \gamma_a] = [5, 25, \nu \pm \pi/2, 8]$ (Fig. 2). ω_u is chosen between $\nu - \pi/2$ and $\nu + \pi/2$ so that d is maximized in the iteration following the change in state. Avoidance responses

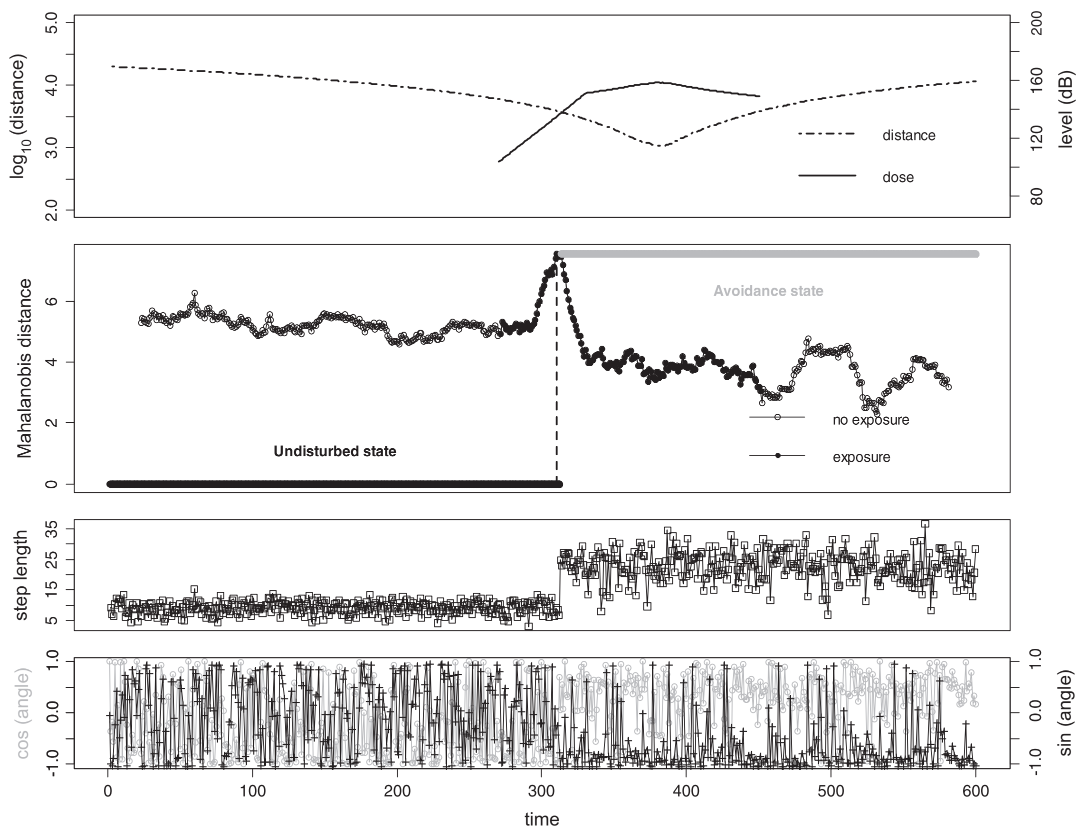


Fig. A2. Example of time-series data from a simulated sonar exposure session used to test the Mahalanobis distance changepoint analysis. Top panel shows variation of dose and distance. Second panel from the top shows variation of the Mahalanobis changepoint statistic and the simulated states, with identification of the estimated changepoint (vertical dashed line). Third panel from the top shows variation in step length and bottom panel shows variation in sine and cosine of step angle.

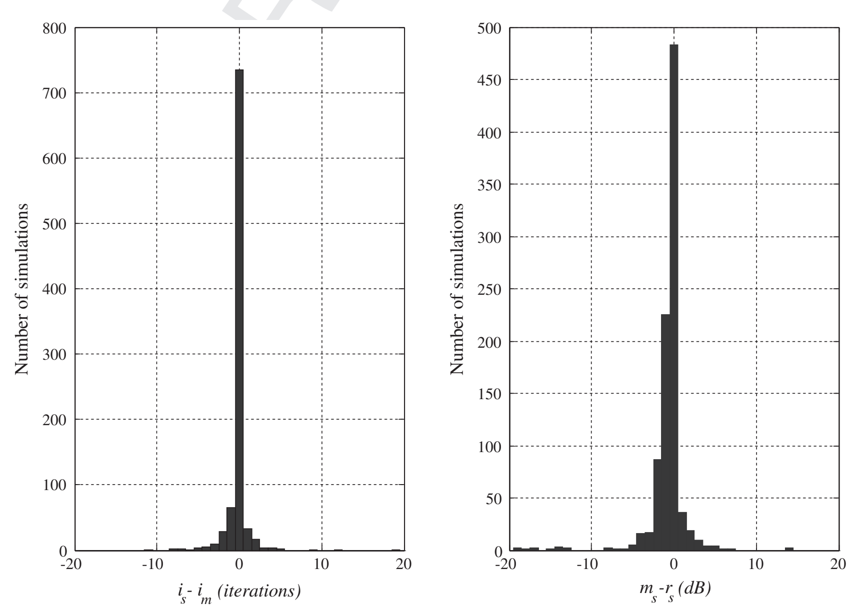


Fig. A3. Histograms of the difference between the simulation iteration at response i_s and the estimated response iteration i_m (left) and of the dose estimated using Mahalanobis changepoint analysis m_s and the simulated response threshold r_s (right), for 1000 simulations.

Please cite this article in press as: Antunes, R., et al. High thresholds for avoidance of sonar by free-ranging long-finned pilot whales (*Globicephala melas*). Mar. Pollut. Bull. (2014), http://dx.doi.org/10.1016/j.marpolbul.2014.03.056

that consisted in movement approximately perpendicular to the heading of the incoming sound source were commonly observed in real sound exposures (Miller et al., 2011, 2012) and it was found by simulation to be the best strategy to avoid an approaching source (Wensveen, 2012).

In each simulation s the whale switched from the undisturbed state to the avoidance state at iteration i_s, when the received sound pressure level (RL) exceeded a response threshold r_s and did not switch back until the end of the simulated exposure session. For each simulation, r_s was randomly chosen from a normal distribution with mean = 160 dB re 1 re 1 µPa and standard deviation = 10 dB, truncated to the range 80–200 dB. The RL at the whale was calculated at every iteration as $RL = SL - 17 \cdot \log_{10}(d)$; with d in meters. A sound transmission loss of $17 \cdot \log_{10}(d)$ was chosen as it was found to approximate the propagation conditions encountered in some real sound exposures sessions previously carried out in this area at this time of the year (Miller et al., 2011). The orthogonal components of the direction of movement of the whale $\sin(\theta)$ and $\cos(\theta)$ and the step length values (l) were used as variables in the aforementioned Mahalanobis changepoint analysis to estimate the iteration i_m and doses m_s corresponding to the response thresholds. The Mahalanobis sliding window width was 5 min. The ability of the Mahalanobis procedure to identify the simulated thresholds was evaluated by running 1000 simulations. For each simulation we calculated the difference between the Mahalanobis changepoint iteration and simulation iteration at which the change from undisturbed to avoidance took place $i_m - i_s$. This gives a measure of the offset between the actual and estimated changepoints that is independent from the dose escalation. We also calculated the difference between the dose at the Mahalanobis changepoint analysis and the simulated response threshold ($m_s - r_s$), that reflects the combined effect of the Mahalanobis changepoint estimation and the dose escalation.

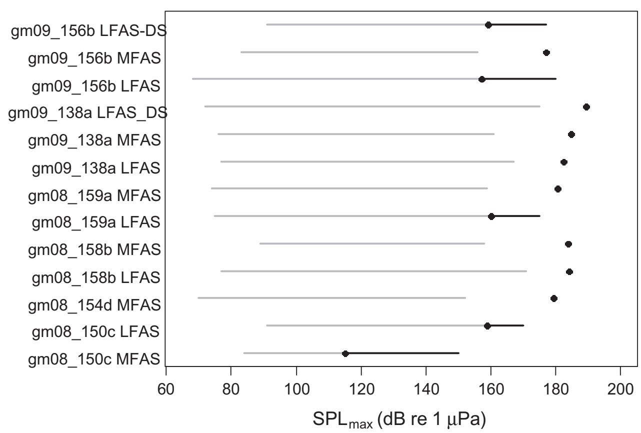


Fig. C1. Exposure ranges for sonar exposures used in dose–response estimation. Gray lines show SPL_{max} range where no response was identified and black lines show SPL ranges beyond the observed response. Black dots show the median posterior of r_{ij}, i.e. the expected behavioural response threshold for whale i during exposure session j estimated by the model.

A.2. Results

The sonar exposure simulations generated tracks where changes in movement patterns between the undisturbed and avoidance states could be observed with the simulated animal moving away from the source after response (Fig. A1). The variation of the Mahalanobis changepoint statistic in each simulation generally showed a single peak that was associated with the simulated change from undisturbed to avoidance behaviour (Fig. A2).

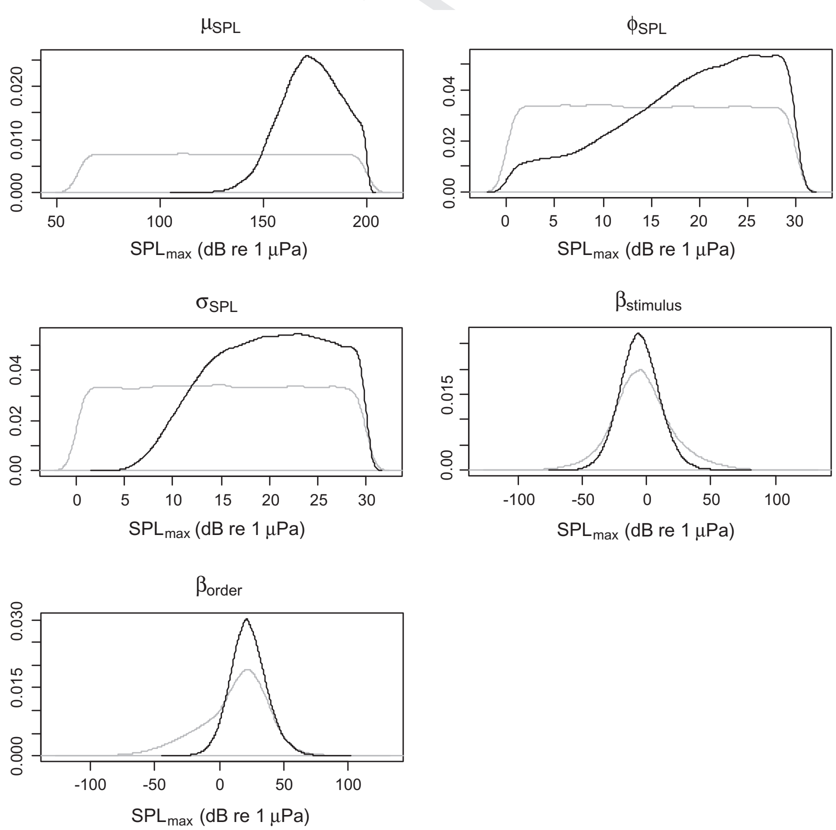


Fig. C2. Prior (grey) and posterior (black) marginal densities for parameters of dose–response function using SPL_{max} as dose.

The median difference between the simulated and the estimated changepoints ($i_s - i_m$) was 0 iterations (pings) (95% quantiles [−35]) and the median difference between the doses associated with these changepoints ($m_s - r_s$) was 0 dB (95% quantiles [−92]). Variability (i.e. error) around the median difference values decreased when the standard deviation of population level response threshold was lowered to 1 ($i_s - i_m$: median 0 iterations, 95% quantiles [−32]; $m_s - r_s$: median −0 dB, 95% quantiles [−21]). Reducing the overlap between the movement parameter distributions in avoidance vs undisturbed modes, while keeping other simulation parameters constant also resulted in a reduction in error ($i_s - i_m$: median 0 iterations, 95% quantiles [−60]; $m_s - r_s$: median 0 dB, 95% quantiles [−43]).

Fig. A3.

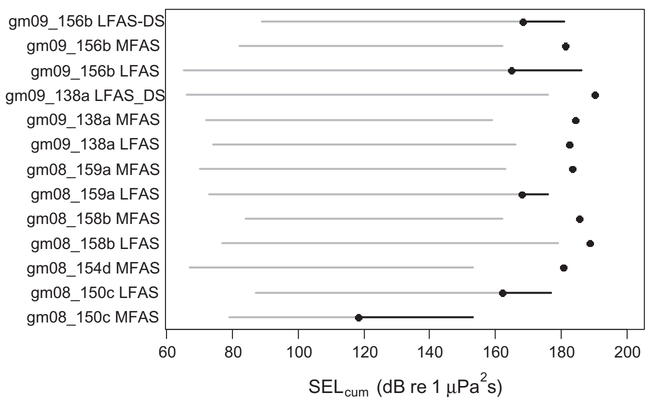


Fig. C3. Exposure ranges for sonar sonar exposures. Gray lines show *SEL$_{cum}$* range where no response was identified and black lines show *SEL$_{cum}$* ranges beyond the observed response. Black dots show the median posterior of r_{ij}, i.e., the expected behavioural response threshold for whale *i* during exposure session *j* estimated by the model.

Appendix B. Summary of results from individual experiments including variation in Mahalanobis changepoint statistic during sonar exposures

The first half of the movement track for the gm08_154d LFAS-UP exposure session is missing because the tag on the initial focal animal released prematurely, and therefore the changepoint analysis could not be carried out for this subset of the data. In 11 of the remaining 17 exposure/control approach sessions, the Mahalanobis changepoint statistic did not show any peaks outside the natural level of variation in the baseline period, and therefore we considered that there were no responses in the measured parameters were scored during these sessions.

Three exposure sessions (gm08_159a LFAS-UP; gm08_159a MFAS-UP; gm08_156b LFAS-DN) showed one peak in the changepoint statistic that was outside baseline variation. The gm08_159a MFAS-UP peak corresponded to a slow (~15 min) heading change. This response was initiated before the transmission of the first sonar ping and was therefore not considered to be a response to sonar. The changepoint peak for gm08_159a LFAS-UP was associated with a reduction in speed (from 2.1–2.6 ms^{-1} to <1.3 ms^{-1}) and no apparent change in heading. The changepoint peak for gm08_156b LFAS-DN was associated with a 135° change in heading away from the source.

Three exposure sessions (gm08_150a LFAS-UP; gm08_150a MFAS-UP; gm08_156b LFAS-UP) showed two or three peaks in the changepoint statistic that were outside baseline variation. The first peak was taken as the response threshold for exposure

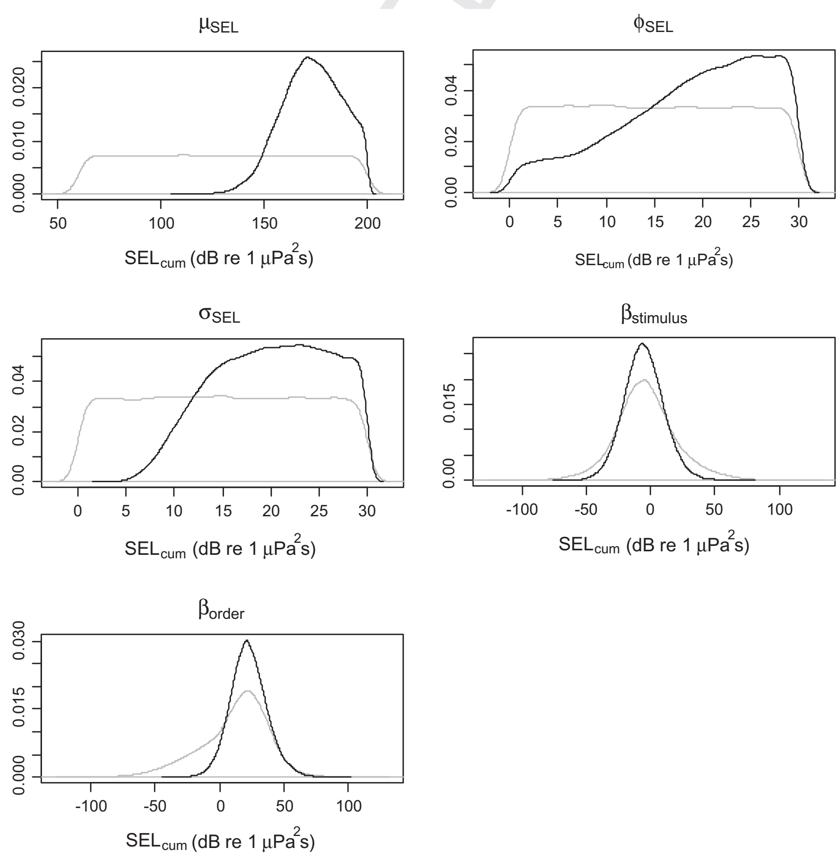


Fig. C4. Prior (gray) and posterior (black) marginal densities for parameters of dose–response function using *SEL$_{cum}$* as dose. Prior distribution was obtained by running the model without data.

sessions gm08_150a MFAS-UP and gm08_156b LFAS-UP. The first changepoint peak for gm08_156b LFAS-UP was associated with a change of heading of 144°, away from the source vessel. This response also was associated with an increase in the production of social sounds (Miller et al., 2011) and an increase in speed to >2 ms^{-1}. The first changepoint peak for gm08_150a MFAS-UP was associated with a sharp heading change (>140°) turning away from the source vessel. Several social and echolocation sounds

Table C1
Probability of response at different levels of SPL_{max} as estimated by the Bayesian dose–response model, in 5 dB re 1 μPa steps. Shown are the mean, median, and quantiles.

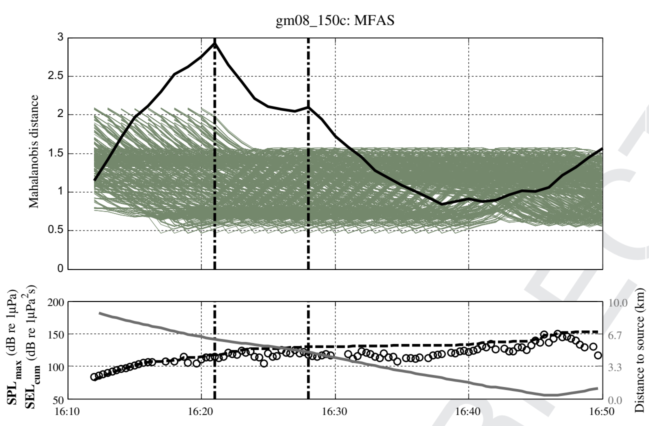
SPL_{max} (dB re 1 μPa)	Mean	Median	Quantiles					
			0.5%	2.5%	25.0%	75.0%	97.5%	99.5%
60	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
65	0.0001	0.0001	<0.0001	<0.0001	<0.0001	0.0003	0.0025	0.0057
70	0.0001	0.0001	<0.0001	<0.0001	<0.0001	0.0007	0.0060	0.0133
75	0.0003	0.0003	<0.0001	<0.0001	<0.0001	0.0014	0.0109	0.0235
80	0.0006	0.0006	<0.0001	<0.0001	0.0001	0.0025	0.0178	0.0366
85	0.0012	0.0012	<0.0001	<0.0001	0.0002	0.0043	0.0271	0.0534
90	0.0022	0.0022	<0.0001	<0.0001	0.0005	0.0070	0.0394	0.0741
95	0.0037	0.0037	<0.0001	<0.0001	0.0010	0.0110	0.0554	0.1000
100	0.0063	0.0063	<0.0001	0.0001	0.0019	0.0169	0.0759	0.1312
105	0.0102	0.0102	<0.0001	0.0001	0.0034	0.0253	0.1016	0.1675
110	0.0162	0.0162	0.0001	0.0004	0.0061	0.0370	0.1334	0.2105
115	0.0251	0.0251	0.0002	0.0009	0.0104	0.0529	0.1715	0.2600
120	0.0377	0.0377	0.0005	0.0021	0.0171	0.0741	0.2160	0.3136
125	0.0550	0.0550	0.0014	0.0045	0.0273	0.1015	0.2674	0.3728
130	0.0783	0.0783	0.0032	0.0090	0.0422	0.1361	0.3250	0.4366
135	0.1088	0.1088	0.0072	0.0167	0.0630	0.1788	0.3874	0.5034
140	0.1476	0.1476	0.0146	0.0295	0.0914	0.2298	0.4547	0.5689
145	0.1955	0.1955	0.0266	0.0494	0.1282	0.2891	0.5248	0.6358
150	0.2524	0.2524	0.0465	0.0783	0.1746	0.3560	0.5962	0.7036
155	0.3180	0.3180	0.0777	0.1177	0.2311	0.4292	0.6670	0.7656
160	0.3911	0.3911	0.1212	0.1698	0.2975	0.5061	0.7343	0.8216
165	0.4704	0.4704	0.1795	0.2359	0.3731	0.5855	0.7969	0.8710
170	0.5538	0.5538	0.2549	0.3153	0.4561	0.6639	0.8517	0.9133
175	0.6382	0.6382	0.3480	0.4084	0.5450	0.7386	0.8981	0.9457
180	0.7216	0.7216	0.4569	0.5133	0.6377	0.8075	0.9340	0.9687
185	0.8013	0.8013	0.5815	0.6280	0.7321	0.8687	0.9612	0.9839
190	0.8752	0.8752	0.7163	0.7497	0.8255	0.9214	0.9800	0.9928
195	0.9418	0.9418	0.8573	0.8748	0.9156	0.9651	0.9924	0.9977
200	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

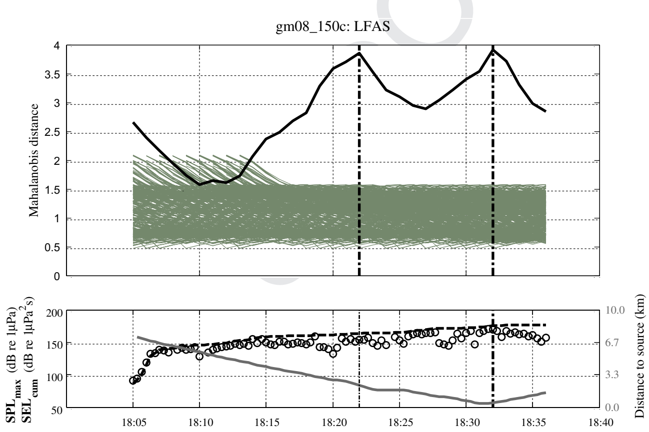
Table C2
Probability of response at different levels of SEL_{cum} as estimated by the Bayesian dose–response mode in 5 dB re 1 μPa² s steps. Shown are the mean, median, and quantiles.

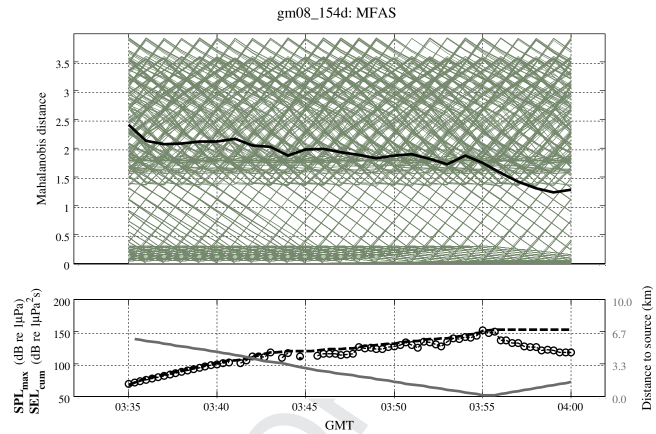
SEL_{cum} (dB re 1 μPa² s)	Mean	Median	Quantiles					
			0.5%	2.5%	25.0%	75.0%	97.5%	99.5%
60	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
65	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.0001	0.0016	0.0038
70	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.0004	0.0040	0.0091
75	0.0002	0.0002	<0.0001	<0.0001	<0.0001	0.0008	0.0074	0.0162
80	0.0003	0.0003	<0.0001	<0.0001	<0.0001	0.0015	0.0123	0.0258
85	0.0006	0.0006	<0.0001	<0.0001	0.0001	0.0026	0.0190	0.0383
90	0.0012	0.0012	<0.0001	<0.0001	0.0002	0.0043	0.0282	0.0544
95	0.0021	0.0021	<0.0001	<0.0001	0.0005	0.0070	0.0403	0.0745
100	0.0037	0.0037	<0.0001	<0.0001	0.0009	0.0111	0.0562	0.1001
105	0.0062	0.0062	<0.0001	<0.0001	0.0018	0.0170	0.0764	0.1311
110	0.0101	0.0101	<0.0001	0.0001	0.0033	0.0255	0.1017	0.1669
115	0.0162	0.0162	<0.0001	0.0003	0.0059	0.0374	0.1331	0.2088
120	0.0252	0.0252	0.0001	0.0007	0.0102	0.0537	0.1703	0.2573
125	0.0381	0.0381	0.0004	0.0018	0.0171	0.0751	0.2147	0.3106
130	0.0561	0.0561	0.0010	0.0040	0.0277	0.1031	0.2660	0.3691
135	0.0804	0.0804	0.0025	0.0082	0.0433	0.1384	0.3239	0.4313
140	0.1122	0.1122	0.0060	0.0160	0.0656	0.1818	0.3867	0.4973
145	0.1527	0.1527	0.0130	0.0293	0.0959	0.2341	0.4544	0.5667
150	0.2029	0.2029	0.0257	0.0506	0.1358	0.2950	0.5249	0.6343
155	0.2628	0.2628	0.0470	0.0823	0.1862	0.3635	0.5976	0.7010
160	0.3316	0.3316	0.0815	0.1275	0.2481	0.4392	0.6688	0.7629
165	0.4095	0.4095	0.1314	0.1879	0.3211	0.5194	0.7373	0.8196
170	0.4937	0.4937	0.2020	0.2651	0.4041	0.6016	0.8001	0.8698
175	0.5823	0.5823	0.2928	0.3592	0.4957	0.6838	0.8554	0.9125
180	0.6728	0.6728	0.4060	0.4690	0.5940	0.7622	0.9019	0.9455
185	0.7623	0.7623	0.5380	0.5922	0.6964	0.8343	0.9388	0.9698
190	0.8482	0.8482	0.6858	0.7245	0.8001	0.8986	0.9669	0.9853
195	0.9279	0.9279	0.8417	0.8620	0.9021	0.9539	0.9868	0.9949
200	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

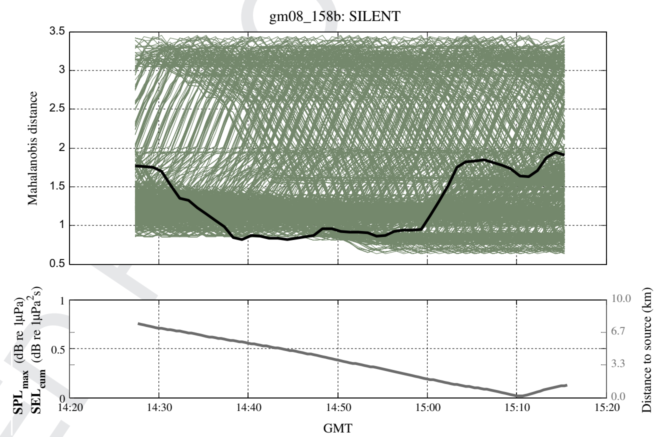
vocalizations could be detected in the audio record of the DTAG up to this point, but vocal output was reduced hereafter until the end of the MFAS-UP exposure session (Miller et al., 2011). Prior to the start of ramp-up of the gm08_150a LFAS-UP exposure session, the focal group of whales was approached by a whale watching vessel. This approach caused a change in the focal whale's heading that was detected by the Mahalanobis change point analysis as the first peak outside baseline range at the start of the exposure session. The second largest value of the Mahalanobis change point statistic occurred later during the exposure session and this second peak was taken as the earliest response elicited by the sonar exposure. This change point was associated with the onset of a series of heading changes and an increase in speed from a mean of 1.6 ms^{-1} (SD 0.4 ms^{-1}) to a mean of 3.2 ms^{-1} (SD 0.8 ms^{-1}).

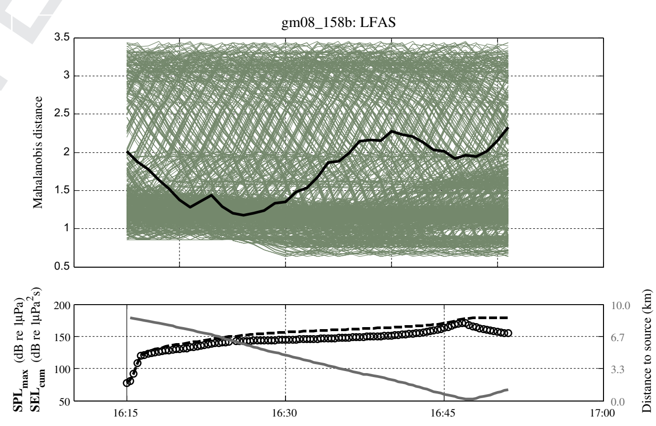
Plots of results from changepoint analysis for pilot-whale sonar exposure sessions. Top panel shows variation in Mahalanobis changepoint statistic during sonar exposure (black line) as well as variation in the same statistic for 10,000 mock exposures during baseline (grey lines). Bottom panel shows sound pressure (circles and triangles: SPL_{max}), sound exposure levels (dashed line: SEL_{cum}), and estimated distance between the sound source and to the source vessel (grey line). Vertical dashed line indicates the times of peaks of the Mahalanobis statistic with values outside the randomization range discussed in the text above. Whale code and sonar signal types are indicated above the top plot (e.g., gm08_150c: MFAS). Plots for exposure session of whale Gm09_138b also show received levels for whale Gm09_138a which was exposed simultaneously:

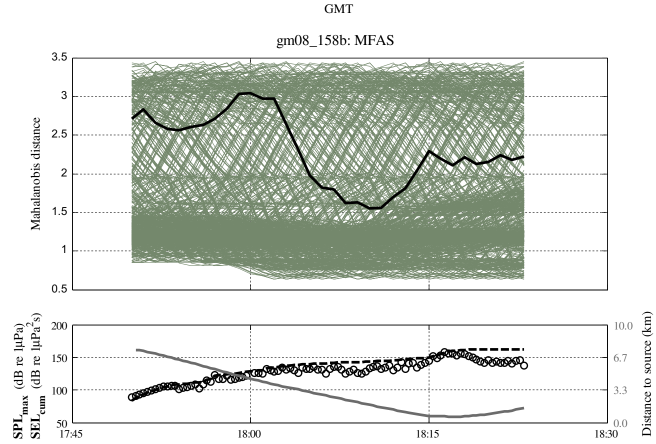


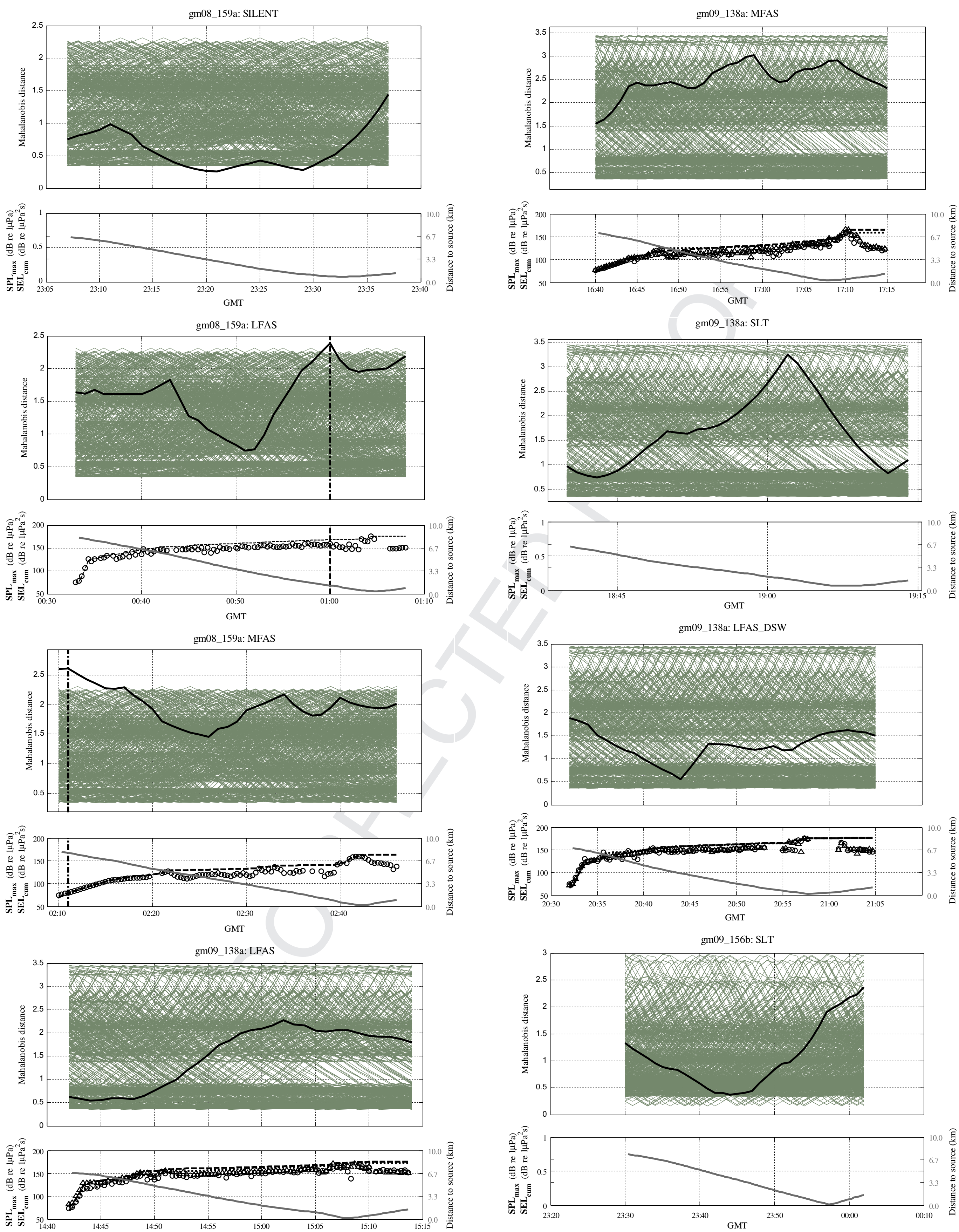


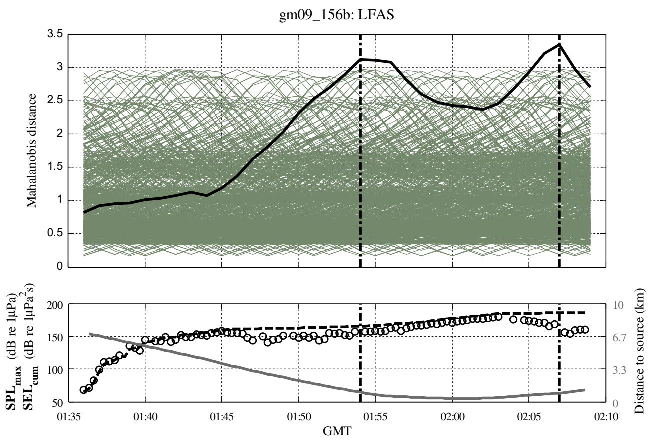


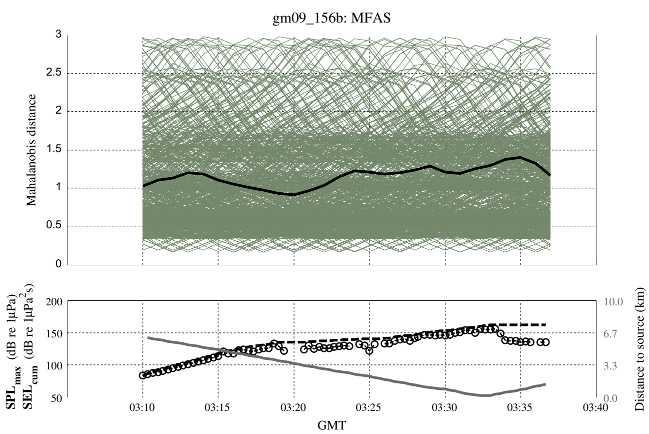


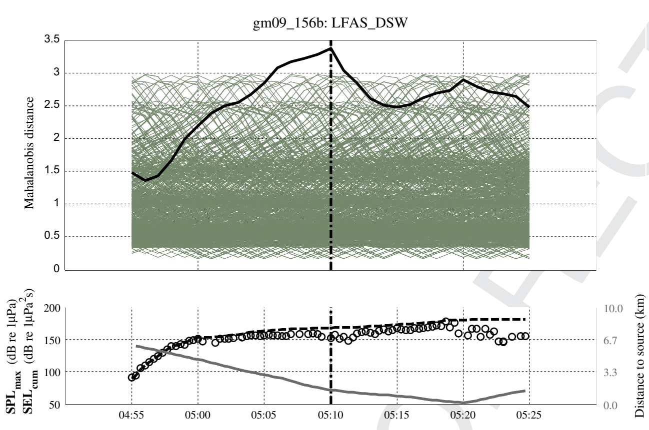





gm08_159a: SILENT
gm09_138a: MFAS
gm08_159a: LFAS
gm09_138a: SLT
gm08_159a: MFAS
gm09_138a: LFAS_DSW
gm09_138a: LFAS
gm09_156b: SLT
Mahalanobis distance
SPLmax (dB re 1μPa)
SELcum (dB re 1μPa²s)
Distance to source (km)
GMT







Appendix C. Dose–response estimation for sonar exposures: additional figures

Figs. C1–C4.

Tables C1 and C2.

References

ACCOBAMS, Agreement on the Conservation of Cetaceans in the Black Sea Mediterranean Sea and Contigous Atlantic Area, 2007. Guidelines to Address the Impact of Anthropogenic Noise on Marine Mammals in the ACCOBAMS Area. Resolution 3.10 Adopted at the Third Meeting of Parties.

Ainslie, M.A., 2010. Principles of Sonar Performance Modeling. Springer-Praxis, Chichester, UK, pp. 521–522.

Alves, A., Antunes, R., Bird, A., Tyack, P.L., Miller, P.J.O., Lam, F.P.A., Kvadsheim, P.H., 2014. Vocal matching of naval sonar signals by long-finned pilot whales (*Globicephala melas*). Mar. Mamm. Sci. http://dx.doi.org/10.1111/mms.12099.

ASCOBANS, Agreement on the Conservation of Small Cetaceans of the Baltic and North Seas, 2006. Adverse Effects of Sound, Vessels and Other forms of Disturbance on Small Cetaceans. Resolution 4 Adopted at the Fifth Meeting of Parties.

ASCOBANS, 2009. Report of the Intersessional Working Group on the Assessment of Acoustic Disturbance. Noise Working Group, ASCOBANS, Bonn.

Barnard, G.A., 1945. A new test for 2x2 tables. Nature 156, 177.

Barrett-Lennard, L.G., Ford, J.K.B., Heise, K.A., 1996. The mixed blessing of echolocation: differences in sonar use by fish-eating and mammal-eating killer whales. Anim. Behav. 51, 553–565.

Buck, J.R., Tyack, P.L., 2000. Response of gray whales to low frequency sounds. J. Acoust. Soc. Am. 107, 2774.

Claridge, D.E., 2001. A mass stranding of cetaceans caused by naval sonar in the Bahamas. Baham. J. Sci. 8, 1–12.

Convention on Migratory Species (CMS), 2009. Adverse anthropogenic marine/ocean noise impacts on cetaceans and other biota UNEP/CMS/resolution 9.19. In: Proceedings of the Ninth Meeting of the Conference of the Parties. Part I: Resolutions Adopted by the Conference of the Parties at Its Ninth Meeting. Annex VIII, pp. 155–250.

Cox, T.M., Ragen, T.J., Read, A.J., Vos, E., Baird, R.W., Balbomb, K., Barlow, J., Caldwell, J., Cranford, T., Crum, L., D'Amico, A., D'Spain, G., Fernández, A., Finneran, J., Gentry, R., Gerth, W., Gulland, F., Hildebrand, J., Houser, D., Hullar, T., Jepson, P.S., Ketten, D., MacLeod, C.D., Miller, P.M., Moore, S., Mountain, D.C., Palka, D., Ponganis, P., Rommel, S., Rowles, T., Taylor, B., Tyack, P., Wartzok, D., Gisiner, B., Mead, J., Mead, Benner, L., 2006. Understanding the impacts of anthropogenic sound on beaked whales. J. Cetacean Res. 7, 177–187.

D'Amico, A., Gisiner, R.C., Ketten, D.R., Hammock, J.A., Johnson, C., Tyack, P.L., Mead, J., 2009. Beaked whale strandings and naval exercises. Aquat. Mammal. 35, 452–472.

DeRuiter, S.L., Southall, B.L., Calambokidis, J., Zimmer, W.M.X., Sadykova, D., Falcone, E.A., Friedlaender, A.S., Joseph, J.E., Moretti, D., Schorr, G.S., Thomas, L., Tyack, P.L., 2013. First direct measurements of behavioural responses by Cuvier's beaked whales to mid-frequency active sonar. Biol. Lett. 9, 20130223.

Dolman, S.J., Parsons, E.C., Wright, J.A., 2011. Cetaceans and military sonar: a need for better management. Mar. Pollut. Bull. 63, 1–4.

Domenici, P., Blagburn, J.M., Bacon, J.P., 2011. Animal escapology I: theoretical issues and emerging trends in escape trajectories. J. Exp. Biol. 214, 2463–2473.

Ellison, W.T., Southall, B.L., Clark, C.W., Frankel, A.S., 2012. New context-based approach to assess marine mammal behavioral responses to anthropogenic sounds. Conserv. Biol. 26, 21–28.

European Parliament and Council, EU Marine Strategy Framework Directive 2008/56/EC of 17 June 2008 establishing a framework for community action in the field of marine environmental policy (Marine Strategy Framework Directive). Off. J. Eur. Union L164/19.

Fahlman, A., Tyack, P.L., Miller, P.J., Kvadsheim, P.H., 2014. How man-made interference might cause gas bubble emboli in deep diving whales?. Front. Physiol. 5, article 13.

Fernández, A., Edwards, J.F., Rodriguez, F., Espinosa de los Morteros, A., Herraez, P., Casstro, P., Jaber, J.R., Martin, V., Arbelo, M., 2005. "Gas and fat embolic syndrome" involving a mass stranding of beaked whales (Family Ziphiidae) exposed to anthropogenic sonar signals. Veterin. Pathol. 42, 446–457.

Goldbogen, J.A., Southall, B.L., DeRuiter, S.L., Calambokidis, J., Friedlaender, A.S., Hazen, E.L., Falcone, E.A., Schorr, G.S., Douglas, A., Moretti, D.J., Kyburg, C., McKenna, M.F., Tyack, P.L., 2013. Blue whales respond to simulated mid-frequency military sonar. Proc. Roy. Soc. B 280, 1–8.

Houser, D.S., Martin, S.W., Finneran, F., 2013. Exposure amplitude and repetition affect bottlenose dolphin behavioral responses to simulated mid-frequency sonar signals. J. Exp. Mar. Biol. Ecol. 443, 123–133.

IUCN, International Union for Conservation of Nature, 2012. IUCN Resolution WCC-212-Res-081-EN, Resolutions and Recommendations from the World Conservation Congress, Jeju, Republic of Korea. IUCN, Gland, Switzerland.

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.

Johnson, M., Hickmott, L.S., Aguilar-Soto, N., Madsen, P., 2008. Echolocation behaviour adapted to prey in foraging Blainville's beaked whale (*Mesoplodon densirostris*). Proc. Roy. Soc. B 1631, 133–139.

King, S.L., Janik, V., 2013. Bottlenose dolphins can use learned vocal labels to address each other. Proc. Natl. Acad. Sci. U.S.A. 110 (32), 13216–13221.

Kuningas S., Kvadsheim, P.H., Lam, F.P.A., Miller, P.J.O., 2013. Killer whale presence in relation to naval sonar activity and prey abundance in northern Norway. ICES J. Mar. Sci. http://dx.doi.org/10.1093/icesjms/fst127.

Kvadsheim, P., Lam, F.A., Miller, P., Alves, A.C., Antunes, R., Brocconcelli, A., van Ijsselmuide, S., Kleivane, L., Olivierse, M., Visser, F., 2009. Cetaceans and Naval Sonar – The 3S-2009 Cruise Report. FFI Report 2009/01140.

Kvadsheim, P.H., Miller, P.J.O., Tyack, P., Sivle, L.D., Lam, F.P.A., Fahlman, A., 2012. Estimated tissue and blood N2 levels and risk of in vivo bubble formation in deep, intermediate and shallow diving toothed whales during exposure to naval sonar. Front. Aquat. Phyisol. 3, article 125.

Mahalanobis, P.C., 1936. On the generalised distance in statistics. Proc. Natl. Inst. Sci. India 2, 49–55.

Miller, P.J.O., 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.

Miller, P., Antunes, R., Alves, A.C., Wensveen, P., Kvadsheim, P., Kleivane, L., Nordlund, N., Lam, F., van IJsselmuide, S., Visser, F., Tyack, P., 2011. The 3S Experiments: Studying the Behavioural Effects of Naval Sonar on Killer Whales (*Orcinus orca*), Sperm Whales (*Physeter macrocephalus*), and Long-finned Pilot Whales (*Globicephala melas*) in Norwegian Waters. Scottish Oceans Institute Technical Report SOI-2011-001.

Miller, P.J.O., Kvadsheim, P.H., Lam, F.A., Wensveen, P.J., Antunes, R., Alves, A.C., Visser, F., Kleivane, L., Tyack, P.L., Sivle, L.D., 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. Mammal. 38, 362–401.

Miller, P.J.O., Antunes, R.N., Wensveen, P.J., Samarra, F.I.P., Alves, A.C., Tyack, P.L., Kvadsheim, P.H., Kleivane, L., Lam, F.A., Ainslie, M.A., Thomas, L., 2014. Dose–response relationships for the onset of avoidance of sonar by free-ranging killer whales. J. Acoust. Soc. Am. 135 (2), 975–993.

Mooney, T.A., Yamato, M., Branstetter, B.K., 2012. Hearing in cetaceans: from natural history to experimental biology. Adv. Mar. Biol. 63, 197–246.

Morton, A.B., Symonds, H.K., 2002. Displacement of *Orcinus orca* by high amplitude sound in British Columbia, Canada. ICES J. Mar. Sci. 59, 71–80.

NMFS, 2005. Assessment of Acoustic Exposure on Marine Mammals in Conjunction with USS SHOUP Active Sonar Transmissions in the Eastern Strait of Juan de Fuca and Haro Strait, 5 May, 2003. NOAA Office of Protected Resources, Washington, DC. Report Dated January 21, 2005.

Nowacek, D.P., Thorne, L.H., Johnston, D.W., Tyack, P.L., 2007. Responses of cetaceans to anthropogenic noise. Mammal. Rev. 37, 81–115.

O'Hara, R.B., Sillanpää, M.J., 2009. A review of Bayesian variable selection methods: what, how and which. Bayesian Anal. 4, 85–118.

Pacini, A.F., Nachtigall, P.E., Kloepper, L.N., Linnenschmidt, M., Sogorb, A., Matias, S., 2010. Audiogram of a formerly stranded long-finned pilot whale (*Globicephala melas*) measured using auditory evoked potentials. J. Exp. Biol. 213, 3138–3143.

Parsons, E.C.M., Dolman, S.J., Wright, A.J., Rose, N.A., Burns, W.C.G., 2008. Navy sonar and cetaceans: just how much does the gun need to smoke before we act? Mar. Pollut. Bul. 56, 1248–1257.

Plein, J.P., Moeschberger, M.L., 2003. Survival Analysis: Techniques for Censored and Truncated Data. Springer, New York.

Plummer, M., 2003. JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling, Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003), March 20–22, Vienna, Austria.

R Development Core Team, 2013. R: A Language and Environment for Statistical Computing, Reference Index Version 2.2.1. R Foundation for Statistical Computing, Vienna, Austria. ISBN:3-900051-07-0.

Rendell, L., Gordon, J.C.D., 1999. Vocal response of long-finned pilot whales (*Globicephala melas*) to military sonar in the Ligurian Sea. Mar. Mammal Sci. 15, 198–204.

Schlundt, C.E., Dear, R.L., Houser, D.S., Bowles, A.E., Reidarson, T., Finneran, J.J., 2011. Auditory evoked potentials in two short-finned pilot whales (*Globicephala macrorhynchus*). J. Acoust. Soc. Am. 129, 1111–1116.

Simmonds, M., Lopez-Jurado, L.F., 1991. Whales and the military. Nature 337, 448.

Sivle, L.D., Kvadsheim, P.H., Fahlman, A., Lam, F.P.A., Tyack, P.L., Miller, P.J.O., 2012. Changes in dive behavior during naval sonar exposure in killer whales, long-finned pilot whales, and sperm whales. Front. Physiol. 3, 1–11.

Southall, B.L., Bowles, A.E., Ellison, W.T., Finneran, J.J., Gentry, R.L., Greene Jr., C.R., Kastak, D., Ketten, D.R., Miller, J.H., Nachtigall, P.E., Richardson, W.J., Thomas, J.A., Tyack, P.L., 2007. Criteria for injury: TTS and PTS. Aquat. Mammal. 33, 437–445.

Tyack, P., 1981. Interactions between singing Hawaiian humpback whales and conspecifics nearby. Behav. Ecol. Sociobiol. 8, 105–116.

Tyack, P.L., Zimmer, W.M.X., Moretti, M., Southall, B.L., Claridge, D.E., Durban, J.W., Clark, C.W., D'Amico, A., DiMarzio, N., Jarvis, S., McCarthy, E., Morrissey, R., Ward, J., Boyd, I.L., 2011. Beaked whales respond to simulated and actual navy sonar. Plos One 6, e17009.

US Navy, 2008. Southern California Range Complex, Final Environmental Impact Statement/Overseas Environmental Impact Statement. December 8, 2008.

Verfuß, U.K., Miller, L.A., Schnitzler, H., 2005. Spatial orientation in echolocating harbour porpoises (*Phocoena phocoena*). J. Exp. Biol. 208, 3385–3394.

Yang, W.C., Chou, L.S., Jepson, P.D., Brownell, R.L., Cowan, D., Chang, P.H., Chiou, H.I., Yao, C.J., Yamada, T.K., Chiu, J.T., Wang, P.J., Fernández, A., 2008. Unusual cetacean mortality event in Taiwan, possibly linked to naval activities. Veterin. Rec. 162, 184–186.

Zirbel, K., Balint, P., Parsons, E.C.M., 2011. Navy sonar, cetaceans and the US Supreme Court: a review of cetacean mitigation and litigation in the US. Mar. Pollut. Bull. 63, 40–48.