PREDICTING ACOUSTIC DOSE ASSOCIATED WITH MARINE MAMMAL BEHAVIOURAL RESPONSES TO SOUND AS DETECTED WITH FIXED ACOUSTIC RECORDERS AND SATELLITE TAGS

Authors:

A.M. von Benda-Beckmann[1)], P.J. Wensveen[2,6)], M. Prior[1)], M. A. Ainslie [1,5)], R.R. Hansen[3,7)], S. Isojunno[2)], F.P.A. Lam[1)], P.H. Kvadsheim[4)], P.J.O. Miller[2)]

1) Netherlands Organisation for Applied Scientific Research (TNO), The Hague, The Netherlands.

2) Sea Mammal Research Unit, School of Biology, Scottish Oceans Institute, University of St Andrews, St Andrews, UK.

3) University of Oslo, Department of Biosciences, Oslo, Norway

4) Norwegian Defence Research Establishment (FFI), Maritime Systems, Horten, Norway.

5) currently at JASCO Applied Sciences (Deutschland) GmbH, Eschborn, Germany.

6) currently at Faculty of Life and Environmental Sciences, University of Iceland, Reykjavik, Iceland

7) currently at DNV GL AS, Environmental Risk Management, Høvik, Norway

ABSTRACT

To understand the consequences of underwater noise exposure for cetaceans, there is a need for assessments of behavioural responses over increased spatial and temporal scales. Bottom-moored acoustic recorders and satellite tags provide such long-term, and large spatial coverage of behaviour compared to short-duration acoustic-recording tags. However these tools result in a decreased resolution of data from which an animal response can be inferred, and no direct recording of the sound received at the animal. This study discusses the consequence of the decreased resolution of data from satellite tags and fixed acoustic recorders on the acoustic dose estimated by propagation modelling, and presents a method for estimating the range of sound levels that animals observed with these methods have received. This problem is illustrated using experimental results obtained during controlled exposures of northern bottlenose whales (*Hyperoodon ampullatus*) exposed to naval sonar, carried out near Jan Mayen, Norway. It is shown that variability and uncertainties in the sound field, resulting from limited sampling of the acoustic environment, as well as decreased resolution in satellite tag locations, can lead to quantifiable uncertainties in the estimated acoustic dose associated with the behavioural response (in this case avoidance and cessation of foraging).

I. INTRODUCTION

Behavioural response studies have carried out experiments to investigate the effects of navy sonar and other anthropogenic sounds on marine mammal behaviour (Miller e al., 2009; Southall et al., 2016; Harris et al., 2017). During controlled exposure experiments (CEE), sounds of interest are transmitted to subject whales at specified source levels and positions relative to the subject animal(s). Animal responses can be measured from on-animal tags, such as high-resolution sound and movement-recording tags (e.g. DTAG; Johnson & Tyack, 2003; Nowacek et al., 2003), position and depth-transmitting satellite tags (Schorr et al., 2014; Falcone et al., 2017), or using remote sensors, such as networks of bottom-moored acoustic sensors (Tyack et al., 2011; Moretti et al., 2014; Southall et al., 2016; Martin et al., 2015; Manzano-Roth et al., 2016). DTAGs have been commonly used as they provide detailed information on diving and acoustic behaviour, as well as a direct recording of the sound to which animals are exposed. To understand the consequences of behavioural responses for individual animals, there is a recognized need to measure the response over increased spatial and temporal scales, for which satellite tags and acoustic recorders are used as additional complementary tools (Southall et al., 2016). These tools provide more long-term, and larger spatial coverage of the response, at the cost of a decreased resolution in which the animal response can be measured. Since there is no direct acoustic sensor on the animals, these tools also do not enable direct measurement of the acoustic dosage to which the animal is exposed.

It is often unclear what specific characteristics of the sound field drive the behavioural responses of marine mammals, and therefore a range of metrics is usually reported in dose response studies. Common measures reported are rms sound pressure, peak sound pressure, sound exposure integrated over time, and signal rise time (Madsen et al., 2006; Southall et al., 2007; Götz and Janik, 2011). There are also indications that contextual factors, such as

distance to the source, behavioural state, or age/sex of the subject, mediate the responsiveness of the animals (Ellison et al., 2012; Goldbogen et al., 2013; Houser et al., 2013; Miller et al., 2014; DeRuiter et al., 2015; Southall et al., 2016; Falcone et al., 2017). Regulators and managers often estimate the extent of disturbance that sound producing activities might lead to using dose-response relationships (Department of the Navy, 2013; Dekeling et al., 2014) that relate the sound dosage, usually in terms of SPL, to probability of responses derived from these studies (e.g. Houser et al., 2013; Moretti et al., 2014; Miller et al., 2014; Antunes et al., 2014; Harris et al., 2015; Wensveen, 2016) regardless of other contextual variables that may have influenced the observed response thresholds.

Sound propagation in water can lead to strong gradients in the sound field amplitude which means that the sound levels measured with a stationary hydrophone may not be representative of the exposure levels received by animals which are detected on this hydrophone. Instead, estimates of the sound dose associated with behavioural responses need to be calculated using sound propagation models (e.g. Moretti et al., 2014). Reliable prediction of the sound field requires an accurate description of the oceanographic and geoacoustic parameters in the area. The availability and quality of such environmental data can vary greatly between areas. This will affect the reliability of the acoustic model predictions of the level of sound received by the subject whale, and therefore the accuracy to which dose-response relationships can be established using these methods.

To estimate the received level at the locations of animals detected on a moored or floating hydrophone, one needs to measure or estimate their location relative to the sound source.

If stationary hydrophone arrays are used to measure responses, the actual animal position may be established using acoustic localization (Ward et al., 2008; Wahlberg et al., 2011; Marques et al., 2009; Moretti et al., 2014; Gassmann et al., 2015), but in other cases where single

92 hydrophones are used animals must be assumed to be located within a volume of water
93 around the recorder (determined by the distance at which animal sounds can be detected). To
94 characterise the extent of this volume, efforts have been made in estimating detection
95 distances from hydrophones for beaked whales (e.g. Zimmer et al 2008; Marques et al., 2009;
96 von Benda-Beckmann et al., 2010, 2018; Ward et al., 2011; Hildebrand et al., 2015). If
97 assumptions about dive depths must be made, these are typically based on baseline
98 information obtained from other measurements, such as animal-borne tags.

99 While satellite tags provide information on animal location, measurements of sound dosage
100 are currently not attainable due to technical constraints of transmitting acoustic data. The
101 accuracy and dimensionality by which animal locations are recorded (sampled and
102 transferred) depend on the tag model used (e.g. with or without auxiliary Fastloc-GPS or
103 depth sensors), and the level of data compression and degradation exerted to enable efficient
104 transmission within a limited bandwidth and timeframe (Cooke et al., 2004; Tomkiewicz et
105 al., 2010; Carter et al., 2016). The limited bandwidth of the existing data transfer methods
106 often leads to a compromise where the resolution of the dive profile is substantially reduced to
107 maintain temporal coverage of diving activity. This leads to uncertainty in the actual depth of
108 animal at the time of each transmission, which adds to the uncertainty in the estimates of
109 sound dosage received by the tagged animal.

110 Multi-scale controlled sonar exposure experiments (Off-Range Beaked whale Studies, 3S-
111 ORBS) involving northern bottlenose whales (*Hyperoodon ampullatus*) were carried out near
112 the island of Jan Mayen (Norway, 71°N) in 2015 and 2016. The experiments involved DTAG
113 and satellite tags deployed on multiple northern bottlenose whales, and bottom-moored
114 acoustic recorders simultaneously monitoring the nearby animals' vocal behavioural
115 responses to 1-4 kHz sonar exposure at different temporal and spatial scales (Wensveen et al.,
116 2019). A previous experiment in this area indicated that northern bottlenose whales might be

very sensitive to sonar exposures (Miller et al., 2015) and showed avoidance responses consistent with that observed in other beaked whales species experimentally exposed (Tyack et al., 2011; DeRuiter et al., 2013). Observed responses of northern bottlenose whales to sonar during the controlled exposure experiments in 2015 and 2016 also followed the stereotypical response of beaked whales (Wensveen et al., 2019), with cessation of echolocation clicking, a change in dive behaviour, and strong avoidance of the sonar source location (Tyack et al., 2011; DeRuiter et al., 2013; Moretti et al., 2014; Miller et al., 2015). While the sound dosage of the individuals tagged with DTAGs could be calculated directly from the sound recording, the sound dosage associated with the observed responses of satellite tagged animals and responses detected by the acoustic recorders needed to be estimated.

This paper describes a method for estimating the sound dosage and the associated uncertainties around the fixed recorders and satellite tags, using the 2016 ORBS study as a case-study (Wensveen et al., 2019). The exposure area near Jan Mayen was situated in an oceanographic frontal zone with warmer, more saline, waters coming in from the south, and colder, less saline waters coming in from the Greenland Sea into the Norwegian Sea (Bourke et al., 1992; Rudels et al., 2005; Mork et al., 2014). Acoustic propagation in such frontal environments is notoriously difficult to model accurately (Heathershaw et al., 1991; Lynch et al., 2003; Finette, 2006; Katsnel'son et al., 2007; Pecknold and Osler, 2012; Shapiro et al., 2014), especially when detailed measurements of the oceanographic conditions are lacking. Due to logistical restrictions of the sailing vessel used during the experiment, only a limited number of measurements of oceanographic variables determining sound propagation could be obtained in the experimental area and time window. We discuss the consequences of the uncertainties in sound propagation, as well as the limited resolution of satellite tags and acoustic recorders on the estimated acoustic dose, and the ability to reliably establish dose-response relationships for marine mammals exposed to navy sonar.

142

143 ## II. METHODS & TECHNIQUES

144 ### A. 3S-ORBS experiments

145 #### *1. Experimental protocol and instrument deployment procedures*

146 A controlled sonar experiment was carried out 18 June 2016 east of the island of Jan Mayen.
147 Full details of the sonar exposure experiment are provided in Wensveen et al. (2019). Here we
148 provide a brief summary: In the days before the sonar exposure, animals were tagged, and one
149 bottom-moored hydrophone recorded data continuously over a period of 3 weeks before,
150 during, and after the sonar exposure. Six northern bottlenose whales were tagged with
151 position and depth-transmitting satellite tags (LIMPET configuration (SPLASH10, Wildlife
152 Computers), and one northern bottlenose whale was tagged with a DTAG (d3 core unit,
153 University of Michigan) that also housed a Fastloc-GPS logger and ARGOS transmitter, prior
154 to the sonar exposure. A sound source was then deployed at distances of approximately 15-30
155 km away from the animals and the bottom-mounted acoustic recorder. The source was located
156 at 70.76044N 6.0967W at start of sonar exposure, and started to transmit at 12:16:00Z. Due to
157 current (approximately 0.5 kn), the source vessel drifted by about 950 m during the
158 experiment in the direction of 70.7663N 6.1030W (location of final transmission). The
159 vertical source array consisted of 15 individual transducer elements with a 15.2-cm (6-inch)
160 centre-to-centre spacing and had an in-beam source level of 214 dB re 1 µPa·m, with strongly
161 decreasing output outside of the main-beam (beamwidth of ~20º (-10 dB full width) measured
162 at 3500 Hz) (Southall et al., 2012). The source array was used to play back a mid-frequency
163 sonar signal, similar to that of a typical US mid-frequency active sonar (MFA) (Southall et al.,
164 2012). Each transmitted pulse had a total duration of 1.6 s and consisted of three components,
165 one 3350 – 3450 Hz linear frequency modulated (LFM) upsweep, followed by two continuous

166 wave (CW) signals at 3600 Hz and 3900 Hz. Each component had 0.5 s duration (12.5 ms
167 Tukey window) with 0.05 s and 0.1 s pause between the components, respectively. The pulse
168 transmission was repeated every 25 s. The exposure consisted of a ramp-up from 154 to 214
169 dB re 1 µPa·m in 1 dB steps per pulse transmission for 20 min followed by 15 min of full
170 power transmissions at 214 dB re 1 µPa·m.

171 Tagging and experiments were conducted under permits from the Norwegian Animal
172 Research Authority (permit no 2011/38782 and 2015/23222) and Icelandic Ministry of
173 Fisheries in compliance with ethical use of animals in experimentation. Experimental
174 procedures were also approved by the Animal Welfare Ethics Committee at the University of
175 St Andrews.

176 ***2. Satellite tag data and DTAG data***

177 The DTAG recorded pressure, temperature, acceleration, magnetic field strength, and sound
178 in two channels (sensitivity in the sonar band: -188.5 dB re 1 V/µPa; sample rate: 240 kHz)
179 (Johnson & Tyack, 2003). The acoustic recording chain of the DTAG was calibrated in an
180 anechoic pool, just prior to the field work. Received levels of the sonar transmissions were
181 measured following Miller et al., (2012), and were computed over the entire sonar frequency
182 range (3350 Hz to 3900 Hz). The suction-cup attached DTAGs were programmed to release
183 from the animals, and recovered for data download. In addition to the whale carrying a
184 DTAG, six animals with SPLASH10 tags were tagged near the DTAGed whale. The satellite
185 tags have two different options of sampling dive records. 'Time series' mode provides an
186 estimate of the depth of the animal at regular intervals (every 2.5 min), whereas 'behaviour
187 log' mode only provides the start/end times, max depth, dive duration and stereotyped shape
188 (square, U or V) of the dive. The time series mode results in higher resolution dive records,
189 but since the amount of data transferred via satellite is limited, this is at the cost of time
190 coverage. Here, satellite tags were configured to sample both time series (1 day every 7 days)

and behavioural log (continuous). We reconstructed dive summary profiles from the behaviour log data. Surface periods were defined with a depth of 0 m and square-shaped, U-shaped, and V-shaped dives were symmetrical with bottom times determined as 75 %, 35 %, and 10 % of the total dive time, respectively. Dive depth was reported as the maximum of a given dive.

3. Acoustic recorder data

An autonomous deep sea acoustic recorder (Loggerhead DSG-ST, sensitivity −168 dB re 1 V/µPa, sampling at 144 kHz), with a flat frequency response between 100 Hz to 30 kHz, was deployed at (70.9254N 6.5607W) recording continuously from 2016-06-10T14:57 to 2016-06-22T12:11Z. The recorder was attached in the centre of a 200 m bottom-mounted mooring line, with three floats on top. The bottom depth was approximately 2300 m, and hence the estimated recorder depth was at approximately 2200 m, with an estimated location uncertainty of ± 50 m.

The presence of northern bottlenose whale clicks was detected using an automated energy detector in 2.5 min bins (Wensveen et al, 2019). During the sonar exposure, clicking around the recorder ceased, and clicking was not observed again until 14 h after the sonar exposure (Wensveen et al., 2019).

4. Defining sound dosage associated with responses

Acoustic quantities often reported in association with behavioural responses are sound pressure level (SPL) and sound exposure level (SEL) (Southall et al., 2007). For definitions of these and other acoustical quantities we follow ISO (18405). The terms SPL and SEL both require specification of a bandwidth and duration over which these are measured. In the case of cumulative SEL, often the sound exposure of each pulse is integrated over all pulse within a specified time window (e.g. duration of the entire exposure). Different studies also measure,

215 and report SPL associated with responses in different ways: the maximum SPL measured up
216 to a specific time of response (Miller et al., 2012, 2014; Sivle et al., 2016; Southall et al.,
217 2012; Tyack et al., 2011), or the maximum SPL (SPLmax) measured over the entire duration
218 of the exposure period (e.g. Moretti et al., 2014; Wensveen et al., 2017).

219 To be consistent with previous studies reporting SPL associated with sonar exposures (Miller
220 et al., 2012, 2014, 2015; Sivle et al., 2015), Table I summarizes the adopted methodology for
221 measuring SPL, SPLmax, and SEL.

222

223 TABLE I. Acoustic metrics and their definitions used to express the acoustic dose to which
224 animals were exposed.

Metric (Abbreviation)	Symbol	Units	Definition
Single pulse sound pressure level (SPL200ms)	$L_{p,200ms}$	dB re 1 μPa	The maximum value within each pulse of SPL for an averaging time of 200 ms, measured in the full sonar frequency band of 3350 – 3900 Hz band. This integration time was chosen because it is a typical integration time of the marine mammal hearing systems, and assumed to correlate with loudness of the signal (Kastelein et al., 2010; Miller et al., 2012).

Pulse duration	t_{-20dB}	s	Measured pulse duration, defined as time between first and last -20 dB point crossing of the SPL computed using short (10 ms) moving average. (Miller et al., 2012)
single pulse SEL (SELsp).	L_E	dB re 1 µPa^{2}s	The total sound exposure level of a single pulse, measured over the pulse duration t_{20dB}.
cumulative sound exposure level (SELcum)	$L_{E,cum}$	dB re 1 µPa^{2}s	Sound exposure measured as the sum over the sound exposure of all N transmissions, ie. $L_{E,cum} = 10 \log_{10} \sum_{i=1}^{N} 10^{L_{E,i}/10dB}$ dB
Sound pressure level averaged over the pulse duration (SPL20dB)	$L_{p,20dB}$	dB re 1 µPa	Sound pressure level averaged over the duration of the pulse. This quantity more closely resembles the SPL predicted using sound propagation models. It was determined here from the measured SELsp by $L_{p,20dB} = L_E - 10 \cdot \log_{10}(t_{20dB} /(1\ s))$ dB.
Maximum sound pressure (SPLmax)	$L_{p,max}$	dB re 1 µPa	The highest measured $L_{p,200ms}$ over a specified exposure sequence. Here, this sequence was either the entire transmission period, or the period between the start of transmission until the time at which an animal showed a specific response.

225

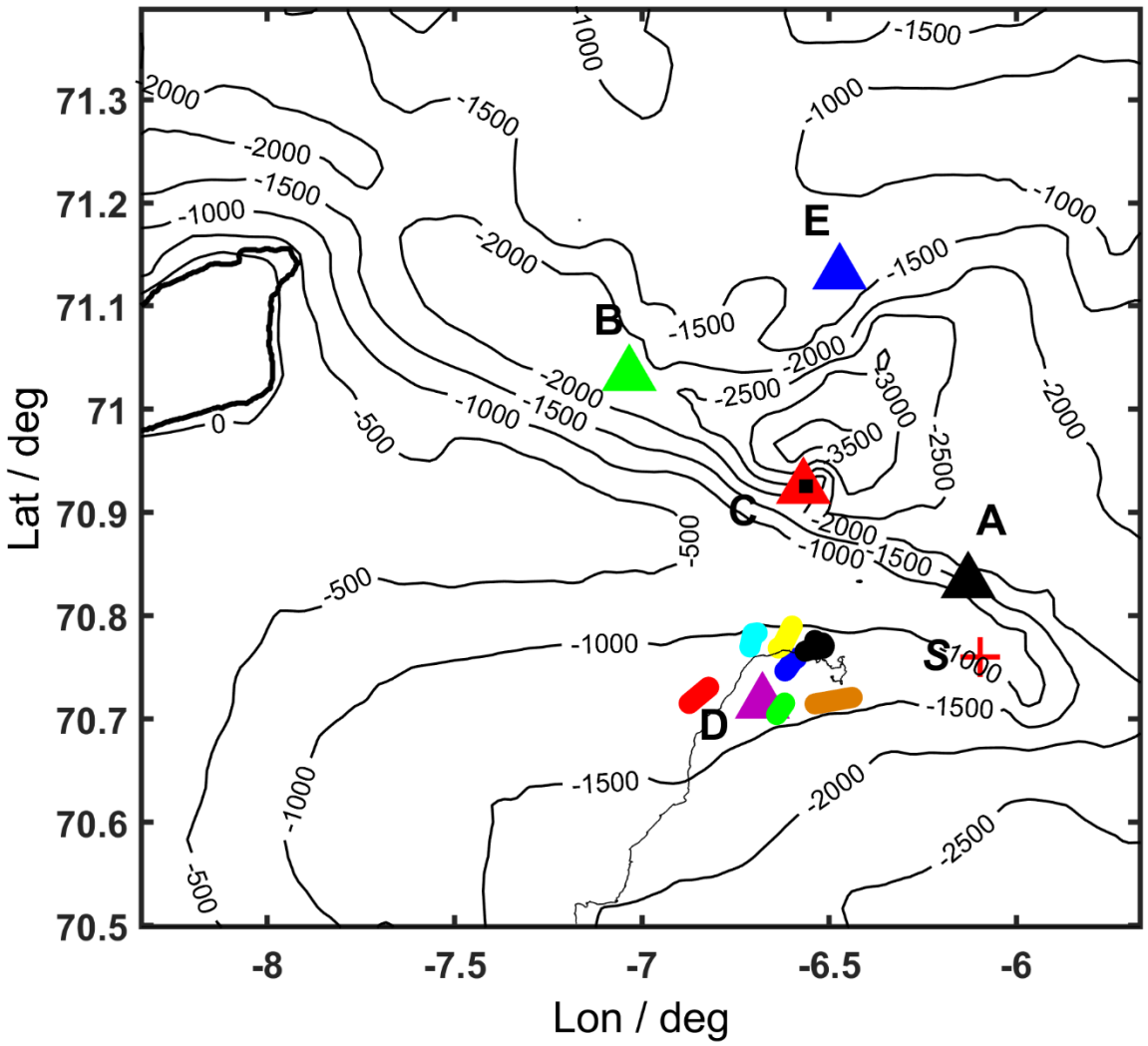
B. Sound propagation modelling

1. Characterizing the environmental conditions

To estimate the maximum SPL (SPLmax) to which the animals were exposed, SPL were computed using acoustic propagation modelling. Oceanographic predictions in the area of the recorders for that day, obtained with the Copernicus Marine Environment Monitoring Service (CMEMS) (von Schuckmann et al., 2017), suggested a gradient in the sound speed profile (SSP) from the transmission location towards the acoustic recorder location (see Appendix A; Fig. A1). The oceanographic predictions suggested that the satellite tagged animals were situated in the colder area, and that the source was located in an area with a strong temperature gradient.

2. Measured sound speed profiles

To characterize the acoustic environment, and support the exposure experiments, several CTD casts were carried out on different times at different locations: near the location of the source, in the area of the satellite tagged animals, and in the direction of the acoustic recorder (Fig. 1 and Fig. 2). CTD profiles were measured near the acoustic recorder locations and tag retrieval locations using a 1.8 kg Valeport Mini-CTD probe. The SSP was found to vary strongly between sites and over time (Fig. 2).



243

244 FIG. 1. Bathymetric map of Jan Mayen area showing depth contours in metres, with locations
245 of CTD measurements A – E (triangles) obtained in a two-week period around the sonar
246 exposure experiment. The sonar source location (*S*) is indicated by the red cross, the acoustic
247 recorder location by the black square, six satellite movement tracks of northern bottlenose
248 whales are indicated in different coloured dots, and one track of an individual tagged with
249 DTAG (black line).

250

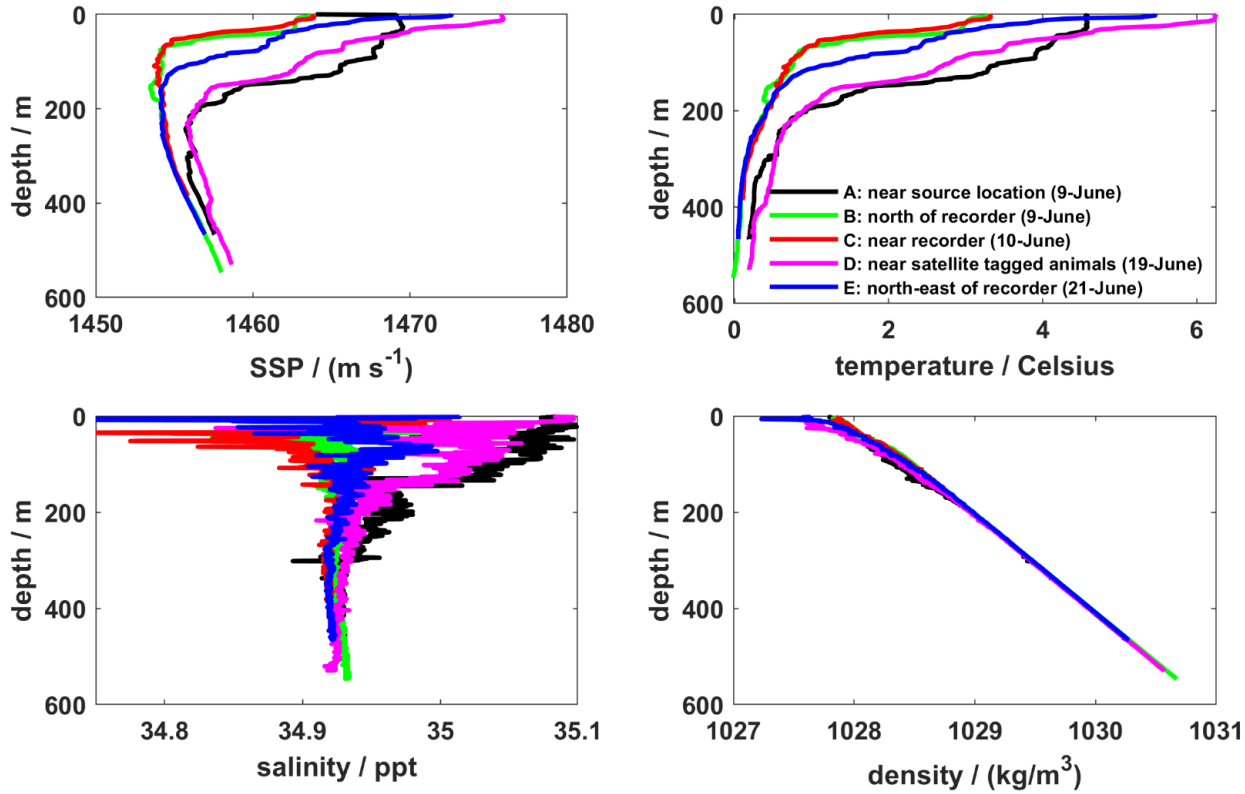


FIG. 2. Measured sound speed profiles (SSPs), salinity (conductivity), temperature and density (CTD) profiles obtained from casts A – E at the different locations from Fig. A1. CTD measurements were obtained at different times and locations during a 2-week period around the exposure session.

The measured SSP near the source 9 days prior to the sonar exposure (A), as well as the SSP measured close to the area of satellite tag deployment 1 day after the transmission (B), were both more consistent with the warmer/saline Atlantic waters in the area. The measured SSP near the recorder (C) and the northern-most SSP (B) were consistent with the colder and less saline waters as predicted with the models, although the measured sound speed at depth was somewhat higher (by 0.3 %, or 4 m/s) than predicted by the oceanographic model (see Appendix, Fig. A2).

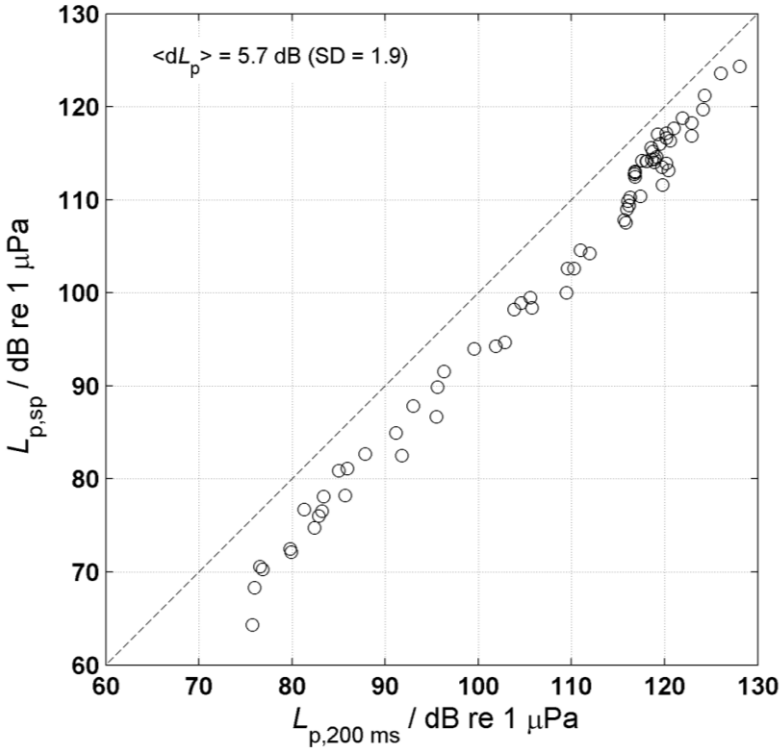
Because of the mismatch between the predicted and measured SSPs it was considered unfeasible to use the oceanographic model to improve the accuracy of the predicted levels on the tagged animals. Instead, measured SSPs were used to indicate uncertainty of the predicted SPLs at the time of the responses (here, cessation of echolocation detected from the recorder, and avoidance responses of the satellite tagged animals; Wensveen et al., 2019). Propagation loss was computed along the direction from source location to tag/recorder location using different SSP casts. For the satellite tagged animals, we used one cast representative location near the sonar source (A), and the one cast close to the satellite tag locations (D). The first cast was obtained 9 days before the transmission (9 June) and the latter one day after the transmissions (19 June). For the predictions around the recorder, the following CTD casts were used: one obtained around the sonar source location (A), a cast close to the recorder (C) location obtained 8 days before the transmission (10 June), and the cast north of the recorder (B) measured 9 days before the transmission (9 June) (Fig. 2).

3. Sound propagation model

The received levels on satellite-tagged bottlenose whales from the Jan Mayen 2016 exposure experiment were modelled with a Gaussian beam ray-tracer (BELLHOP; Porter and Bucker, 1987). BELLHOP was run in a coherent ray mode, with a 30 m (horizontal) range-resolution, and 1 m depth-resolution, using 2000 rays to obtain convergence. A coherent model was used to account for the effect of pressure release near the surface for shallow diving animals. The BELLHOP propagation loss modelling included the vertical beam-pattern of source, with its acoustic centre located at a depth of 27 m, with a small tilt (5º downwards) to account for a slight drift with the current, and using the in-beam source level during the ramp-up and full-power transmissions.

The sound pressure level averaged over the entire pulse (L_{20dB}) was modelled using the power average over the coherent SPL for a number of frequencies (25 Hz steps between 3350 Hz and 3450 Hz for the LFM, and two CW at 3600 Hz and 3900 Hz, applying equal weighting for every frequency bin) for a 1.65 second pulse duration. A frequency-dependent absorption term was included (Urick, 1975), using the center frequency of each frequency band. Sound speed profiles were extrapolated assuming isothermal conditions from their deepest measurement points (~ 400 m to 550 m) to the bottom depth (between 1000 and 2500 m). Bottom parameters were estimated using United States Navy Bottom Sediment Type v2 database (which indicated 'Fine Sand' or 'Silty Sand', HFeva category 11). This corresponded to a bottom density ratio of 1.945, sound speed ratio in bottom of 1.1522, and bottom absorption of 0.89 dB/λ (with λ the acoustic wavelength in the sediment; values adopted for 'fine sand' from Ainslie, 2010, Table 4.18). Finally, the ETOPO-1 database was used for the bathymetry.

The measured levels for onset of behavioural responses on the DTAG were based on averaging times of 200 ms, which generally led to somewhat higher SPL (~ 5.7 dB) than when averaging over the entire pulses (Fig 3). The SPL predicted by the propagation model represents the SPL averaged over the entire pulse (i.e. SPL20dB). This difference in SPL value due to chosen averaging time was added to the modelled SPL obtained from the propagation model to predict the SPL200ms for the satellite tagged animals and animals near the recorder location.



308

309 FIG 3. Effect of averaging time for SPL of sonar transmissions measured on the DTAG.
310 Shown are maximum SPL averaged over 200 ms, $L_{p,200\ ms}$ (X-axis), for each pulse compared
311 to the SPL averaged over an entire single pulse, $L_{p,sp}$ (y-axis). The $L_{p,sp}$ was systematically
312 higher by an average 5.7 dB. This correction was used to correct the propagation models to
313 predict $L_{p,200\ ms}$ for the satellite-tagged animals and animals near the acoustic recorder.

314

315 Propagation loss was computed in a 2D slice towards the position of the animal carrying the
316 DTAG at the time halfway through the exposure. The propagation models indicated that the
317 sound was strongly refracted downwards (Fig. 4), and the resulting areas of high predicted
318 SPLs at the locations of the diving animals were strongly dependent on the adopted SSP, as
319 well as the bottom depth. Because the animals were diving in an area with a steeply sloped
320 sea bottom, initially a computation was performed where 2D slices were computed in steps of
321 0.5º horizontally, from which the SPLs were interpolated to the animal location at each

transmission time. This did not improve the match with the SPL measurements on the DTAG, which suggested that the mismatch with measurements were dominated by other uncertainties than the exact whale location (e.g. SSPs, bottom properties). For this reason, and to limit computation time, a 2D slice with propagation loss for the satellite tags and acoustic recorder location were estimated in the direction of the location at the time halfway through the exposure, which was then assumed to be representative for all transmissions.

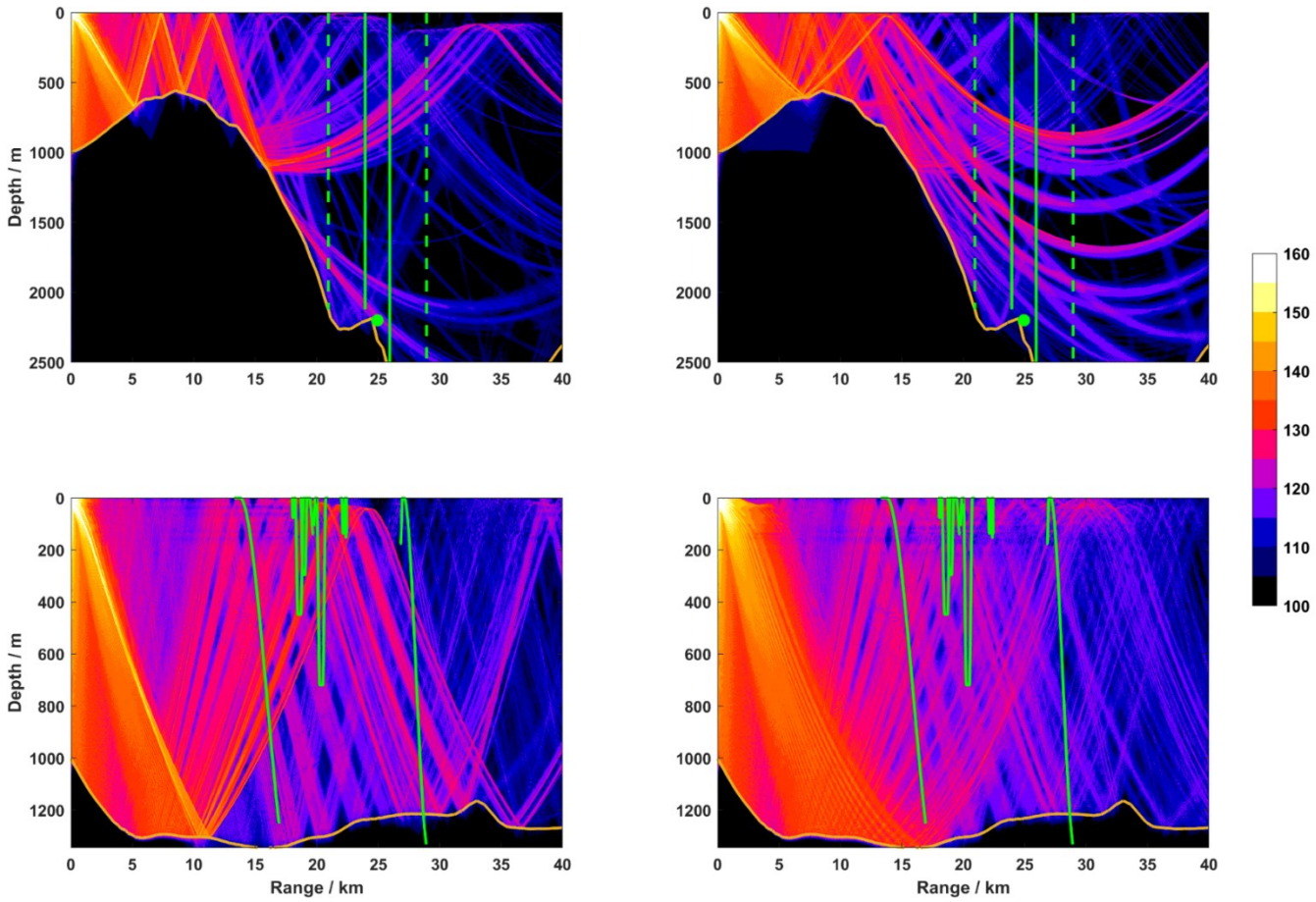


FIG. 4. Predicted sound pressure level, $L_{p,200ms}$, (colour-scale, in dB re 1 µPa) for the 3.3-3.9 kHz sonar pulse for a single beam in the direction towards the acoustic recorder location (top) and towards the satellite tagged animals (bottom) and, using two measured CTD SSPs (left vs. right panels), illustrating the uncertainty due to oceanographic conditions during the experiment. The vertical green lines in the top panels indicate the spread of assumed detection

ranges (1000 m solid lines; 4000 m dashed lines; see text for details) for northern bottlenose whales around the acoustic recorder (indicated by the green filled circle). The lower panel shows the predicted levels in the general direction of the tagged whales, with the dive profiles of the satellite tagged animals during the time of transmission superimposed in green. Note that one animal appears to have a dive depth that exceeded the bottom depth (indicated by the brown line). This only appears larger because all profiles are shown on a single slice (chosen in the direction of location of CTD cast D), whereas the bottom profiles in the direction of the different animals vary. These differences in bottom profiles between animals were incorporated into the predictions for each satellite tag. Note the difference in the scale of the y-axis between the top and bottom panels.

4. Estimating acoustic dose received by animals with satellite tags

Uncertainties in the sound dose received at the satellite tagged whale locations were expected to result from : 1) uncertainties in the propagation modelling 2) uncertainties in the animal xyz location determined by the coarse depth sampling of the satellite tags and the location uncertainty on the Argos locations. A Monte-Carlo approach was adopted to estimate the SPLmax on the satellite tags, using the measured location uncertainty of the satellite tags.

The depth uncertainty in the satellite tags in behaviour log mode could be quantified because the higher depth resolution of the time series log was available for a small part (113 h) of the total tag recording duration (1080 h, approximately 45 d). Normal distributions were fit to the depth error - the difference Δz between the low-resolution depth and higher resolution depth combined from three satellite tags. Separate distributions were fitted for animals at the surface ($z = 0$ in the behaviour log), and at depth ($z > 0$ in the behaviour log) (Fig. 5). For animals at

depth ($z > 0$), a depth correction was included to account for systematic discrepancy in depth estimates between the time series and behaviour log.

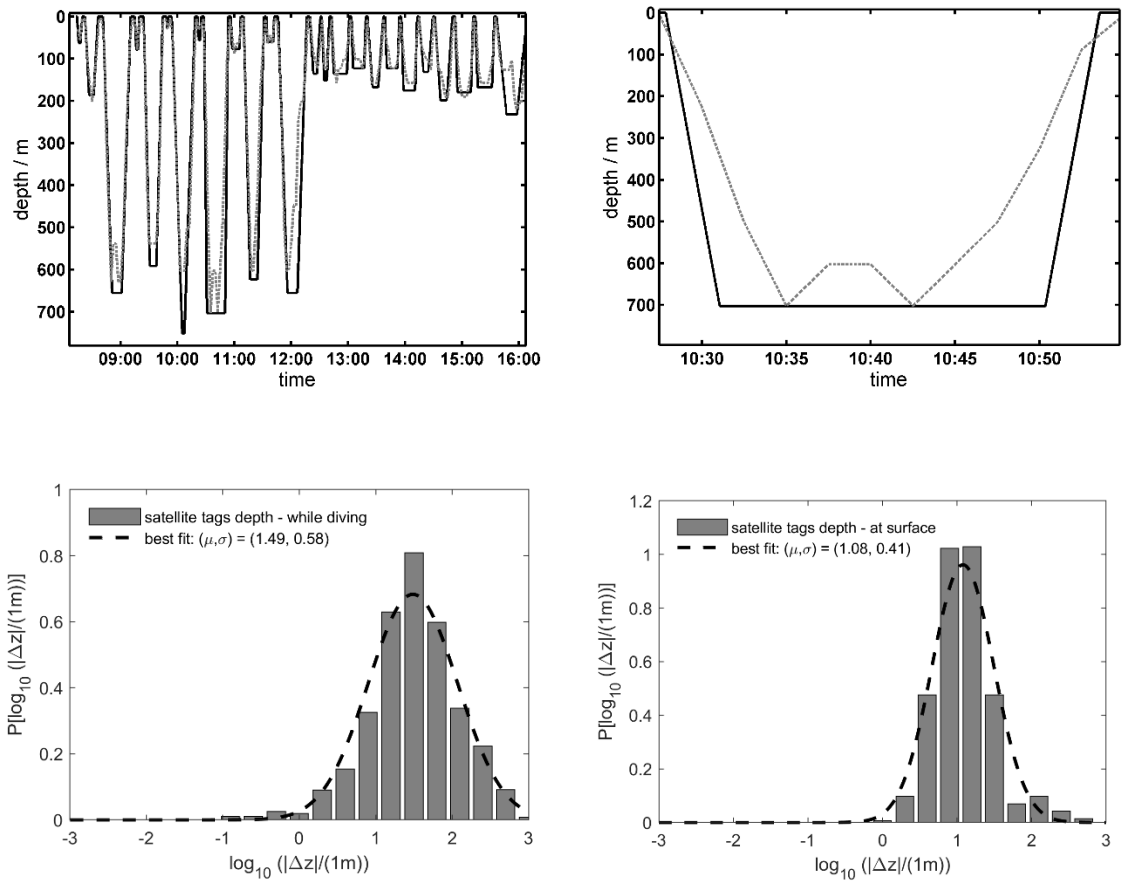


FIG. 5. *Top:* Comparison of a section of a simultaneously collected high-resolution time series dive record (grey dashed) and a lower-resolution behaviour log dive record (black solid) from a satellite-tagged northern bottlenose whale. Top right panel shows a zoom-in on the deep dive starting around 10:30Z. The lower resolution mode can result in substantial uncertainties in depth, in particular if the dive shape is asymmetrical. *Bottom:* Probability density distributions of the depth error, $\Delta z = z_{low} - z_{high}$, for periods when the animal is diving ($z_{low} > 0$) (bottom left panel) and at the surface ($z_{low} = 0$) (bottom right panel). The dotted lines indicate best fit of a normal distribution to the $\log_{10}$ ($|\Delta z|/(1m)$) used to draw realizations for animal depth in the Monte-Carlo approach. For the periods during diving ($z_{low} > 0$) a

constant depth offset of 10 m (mean difference between behaviour log and time series over all dives) was applied before fitting the normal distribution.

For each whale carrying a DTAG, the animal's horizontal track was estimated using track reconstruction (Wensveen et al., 2015). For each whale carrying a satellite tag, raw Argos locations were filtered using a random walk model fitted in a state-space framework with modifications to incorporate error ellipse data, resulting in standard deviation estimates for each transmission location (see Wensveen et al., 2019, for details).

The fitted uncertainty distributions in depth and horizontal range were used to generate a random realization of xyz location for each transmission, which was then used to predict the SPL200ms for each realization. The process was repeated 100 times, resulting in a probability distribution for SPLmax, and SELsp for each transmission, from which the mean, median, and percentiles were derived. For each realization, also the SPL_{max}, and SELcum were computed. This process was performed twice, using two different SSPs (A and D).

5. Estimating acoustic dose near the bottom-moored recorder

To estimate the levels of the sonar transmissions near the diving animals detected on the recorder, representative thresholds for onset of disturbance were computed in two different ways: the first using the distribution of SPL200ms within a chosen detection distance of the recorder, assuming a uniform distribution up to a maximum dive depth of the bottlenose whales. Here, the dive depth was assumed to extend to a depth (either 1000 m or to the seafloor), and was randomly sampled from a uniform distribution. Second, a separate estimate was made by creating mock exposures using baseline satellite dive data. These mock

393 exposures were obtained by randomly sampling 35 min of dive data from the baseline period
394 (30 min after tag on, until the time of transmission) of the satellite tag data, with the method
395 described above to estimate the animal depth, which were distributed randomly within a
396 chosen detection range.

397 The animal locations were placed at different random horizontal distances within a maximum
398 detection range of 1000 m or 4000 m. These distances were based on the similarities between
399 echolocation clicks produced by diving northern bottlenose whales and other beaked whales
400 species, for which it is expected that these clicks would be detectable at distances of up to
401 several kilometres (Hooker et al., 2002; Zimmer et al, 2008; Marques et al., 2009; von Benda-
402 Beckmann et al., 2010; Wahlberg et al., 2011; von Benda-Beckmann et al., 2018).

403 To compare the SPL200ms measured on the acoustic recorder to the modelled SPL200ms, an
404 SPL-distribution within a smaller volume (100x100x100 m^3) of water centered around the
405 estimated deployment location of the acoustic recorder was computed for the last transmission
406 before the cessation of clicking.

407 **III. RESULTS**

408 ***A. Sound propagation in exposed area during CEEs***

409 Model predictions with the SSP measured near the source location (9 days prior to the
410 exposure; cast A) gave a better match between the predicted and DTAG measured SPL200ms
411 than those with the SSP measured around the deployment locations of the DTAG and satellite
412 tags (1 day after the exposure; cast D), but systematically exceeded the measured levels on
413 average by 6 dB (Fig. 6). The mean modelled SPLmax over all transmissions (128 dB re 1
414 µPa) was in closer agreement with the measured SPLmax (128 dB re 1 µPa).

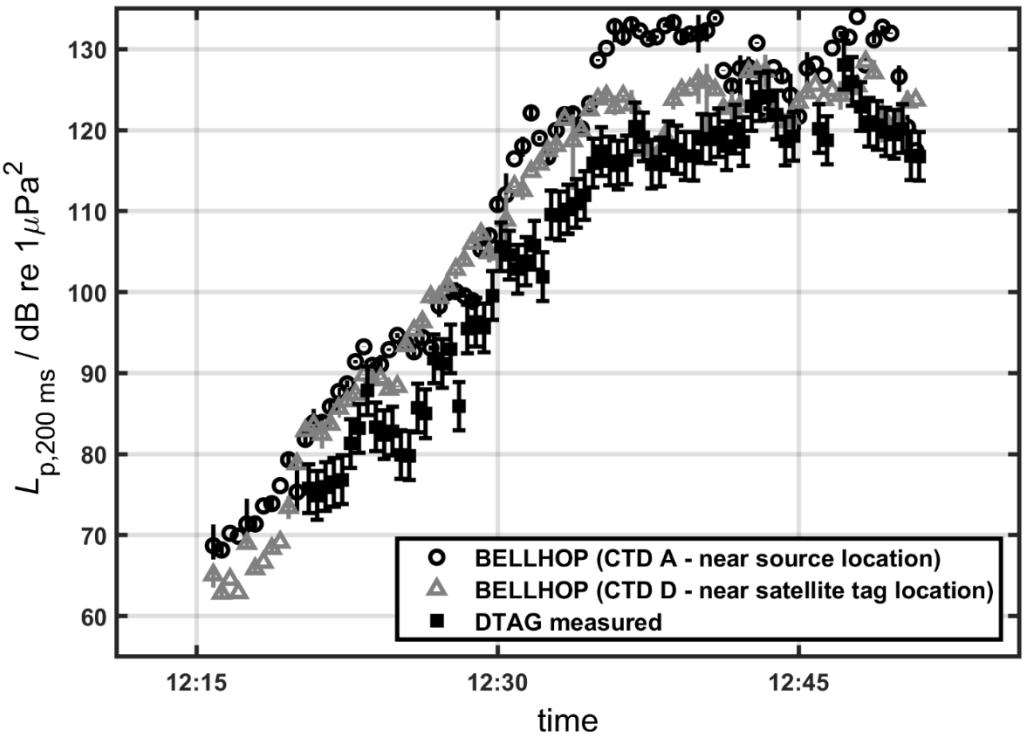


FIG. 6. Modelled and measured sound pressure level, $L_{p,200\ ms}$, of sonar pulses received on the DTAG. Black squares indicate the SPLmax for each transmission as measured on the DTAG. Circles and triangles indicate modelled levels using the CTD measured close to the source location (location A) (9 days before transmission) and using the CTD measured (1 day after transmission) close to the tagged animals (location D), respectively. Black error bars on the DTAG measured levels are indicative of the calibration error.

B. Predicted levels for satellite tagged animals

The spread in estimated SPL200ms on the satellite tagged animals was substantial, and was affected both by location uncertainty, as well as the assumed SSP (Fig. 7). The difference between model predictions of SPLmax using two measured SSPs was between 0 dB and 6 dB, depending on tag (Table II), but variation in SPL200ms between transmissions could exceed 10 dB (Fig. 7). The mean predicted SPLmax over the entire exposure for each satellite tagged

animal ranged between 122 – 132 dB re 1 µPa, and 118 – 130 dB re 1 µPa, depending on the SSP considered (Table II).

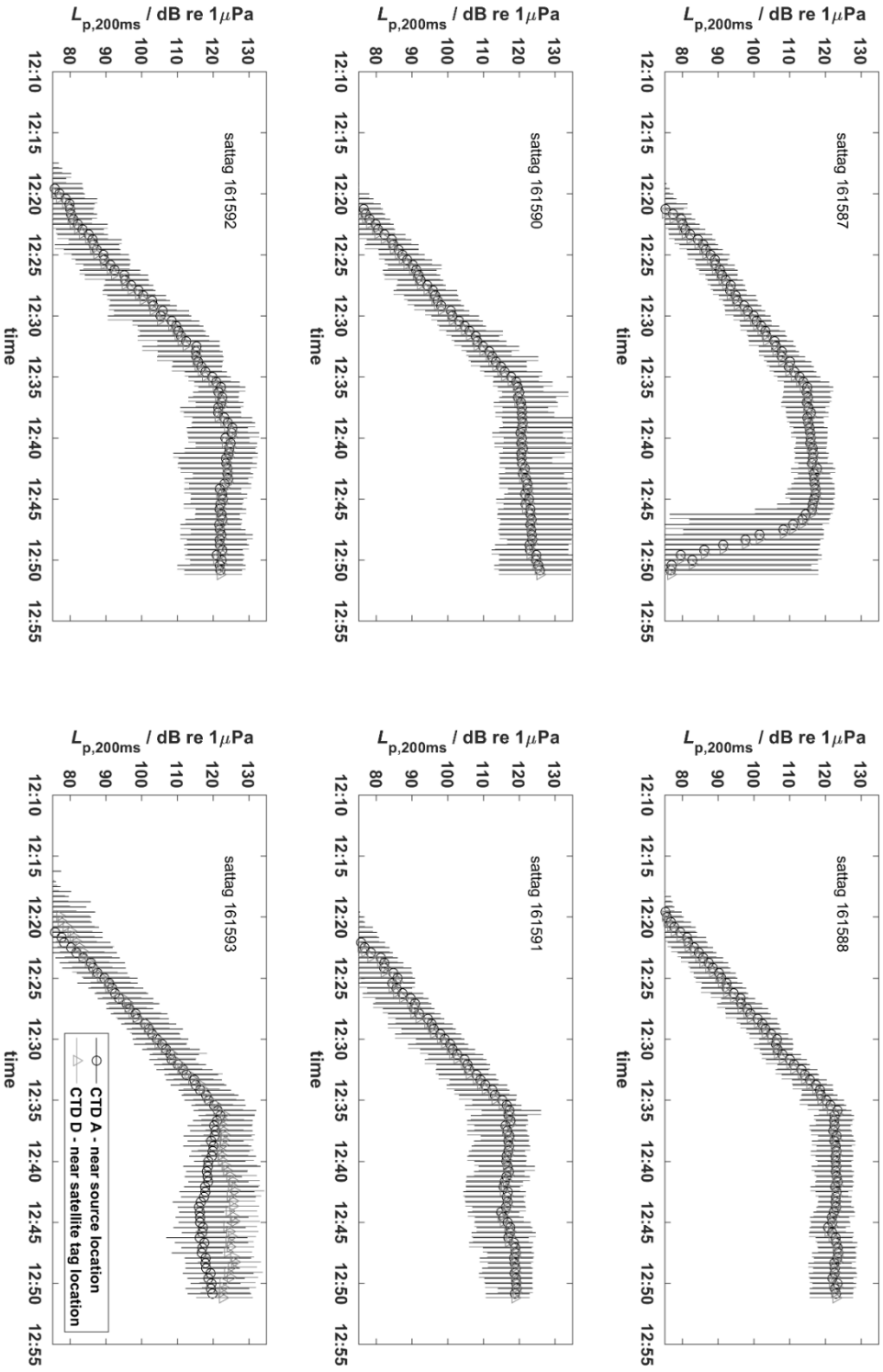


FIG. 7. Modeled sound pressure level $L_{p,200ms}$ of sonar pulses received on each satellite tag. Symbols indicate median levels, crosses the mean (of p^2) and error bars 5^{th} – 95^{th} percentile ranges. The uncertainty was estimated using a Monte-Carlo simulation incorporating the uncertainties in depth and range resulting from the coarse depth information, and estimated location of the animal from the filtered ARGOS track (Wensveen et al., 2019). Grey triangles indicates the modelled $L_{p,200ms}$ values using the CTD measured close to the source location (A). Black circles indicates the modelled $L_{p,200ms}$ values using the CTD measured close to the satellite tagged animals (D).

TABLE II. Summary statistics of distributions of the modelled maximum received sound pressure level $L_{p,max}$ and cumulative received sound exposure level $L_{E,cum}$ on satellite tagged animals over all sonar transmissions using CTD cast near source location (A) (unshaded columns), and using CTD cast near satellite tag location (D) (shaded columns). Mean values reported represent the arithmetic mean of the computed quantities.

TAG ID	$L_{p,max}$ dB re 1µPa								$L_{E,cum}$ dB re 1µPa²s							
	5th perc		50th per (Median)		Mean		95th perc		5th perc		50th perc (Median)		Mean		95th perc	
161587	115	113	121	117	122	118	127	122	132	127	133	128	133	129	134	129
161588	116	118	124	124	129	125	137	130	137	135	138	136	139	136	141	137
161590	116	117	126	127	132	130	138	135	137	138	139	139	140	140	144	142
161591	108	114	121	120	122	120	129	125	131	130	133	131	133	131	135	132
161592	114	116	120	125	127	127	134	133	133	136	136	137	136	137	139	139
161593	116	119	122	126	127	128	137	134	133	137	136	139	136	139	139	140

C. Response thresholds for animals detected around the bottom-moored acoustic recorder

The measured SPL200ms on the bottom-mounted recorder at the last transmission before the cessation of clicking was 80 ± 2 dB re 1 µPa. The limited accuracy of the hydrophone position, as well as uncertainties in SSPs, resulted in different median predicted SPL200ms at the recorder between 88 dB to 99 dB re 1 µPa, with lowest 5th and highest 95th percentiles of 83 dB and 104 dB re 1 µPa (Fig. 8). The levels measured on the acoustic recorder were low compared to the model predictions using the SSP measured close to the source and recorder locations, and more consistent with the lower end of the predicted SPL200ms distribution using the CTD measurements obtained 9 days before the transmission in the area north of the recorder location (Fig. 8).

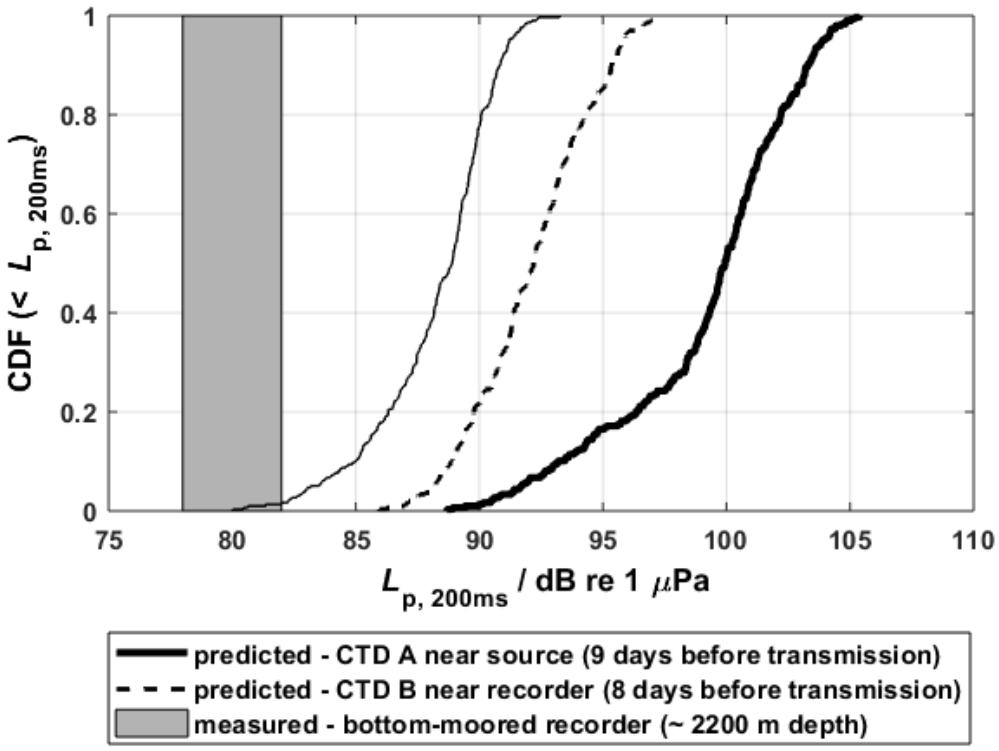


FIG. 8. Cumulative distribution functions (CDFs) of modelled SPL_{200ms} around the acoustic recorder location at the time of cessation of clicking observed in the recording. The range of

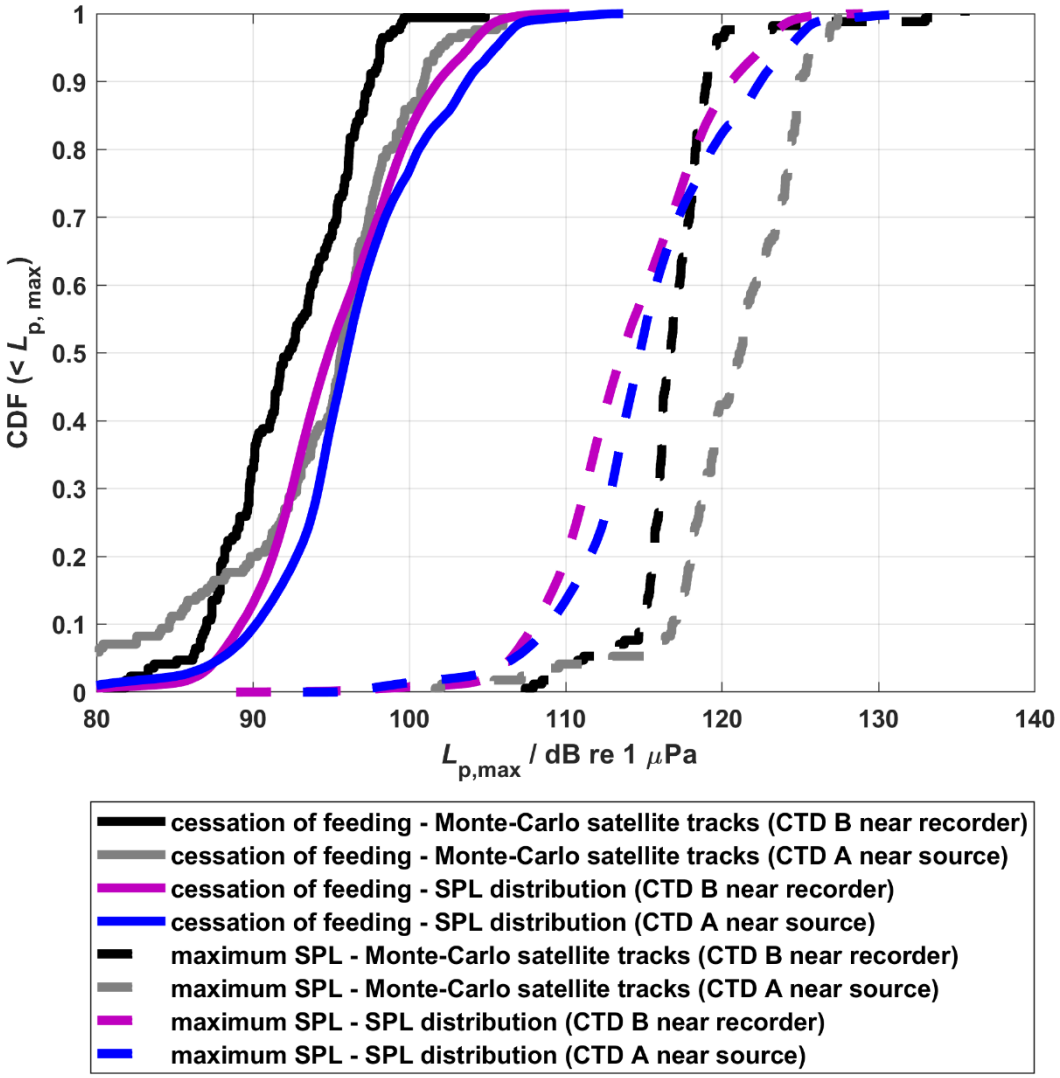
460 predicted values for each predicted distribution reflect the uncertainty in estimated receiver
461 (i.e. animal) location. Sound propagation modelling was performed using three different
462 sound speed profiles (A, B, and C). The grey region indicates the level (mean +/- SD) of the
463 sonar measured at the time of cessation of clicking.

464

465 Predicted SPL200ms for animals detected around the acoustic recorder within a detection
466 distance of 4 km at the time of cessation of clicking had mean values of 95 dB re 1 µPa with
467 5^{th} and 95^{th} percentiles of 84 dB re 1 µPa and 106 dB re 1 µPa, respectively (Fig. 9).
468 Assumptions on detection distances (1 km or 4 km) and assumed maximum dive depth (up to
469 1 km, or up to bottom) were varied. The CDFs in Fig. 9 show the estimates assuming a 4 km
470 detection distance, and dive depth up to 1 km. The two methods used here to estimate CDF
471 for the diving animals (the direct CDF within a volume, or sampling from baseline dive data)
472 had some effect on the resulting CDF (Fig. 9), and were on the same order as differences in
473 SSPs used. The Monte-Carlo method (i.e., generating mock-exposures from the tag data) led
474 to systematically lower SPLs than when sampling the CDF in the entire water column.

475

476



477

478 FIG. 9. Cumulative distribution functions (CDFs) of SPLmax, $L_{p,max}$ for animals around the
479 moored acoustic recorder, estimated at different times windows (solid lines: up to cessation of
480 echolocation clicks; dashed lines: maximum up to the end of the exposure) using different
481 methods (CDF over assuming uniform depth distribution within detection range; Monte-Carlo
482 using baseline dive data from satellite tags within detection range) and using different sound
483 speed profiles (location A near the source, and C near the recorder). The estimated CDFs
484 assumed a detection range of 4000 m and dive depth up to 1 km are shown here for the SPL
485 distribution model.

The predicted SPLmax for the baseline dives detected within the detection ranged was on average 113 dB re 1 µPa to 115 dB re 1 µPa (depending on assumed SSP), with 5th and 95th percentiles of 103 dB re 1 µPa and 124 dB re 1 µPa, respectively. The resulting cumulative distribution functions (CDFs) of the SPL200ms were not strongly affected by the exact choice (median differences were < 4 dB). The assumption of a large detection distance (4 km) combined with animals assumed to dive to the seafloor resulted in a slightly lower 5th percentile (93 dB re 1 µPa to 100 dB re 1 µPa, depending on the SSP).

IV. DISCUSSION

The results from this study demonstrate the challenges in obtaining accurate and precise estimates of the acoustic dose associated with behavioural responses of marine mammals observed with satellite tags and bottom-moored acoustic recorders. These challenges are exacerbated in acoustically complex environments such as the Jan-Mayen oceanographic frontal system.

A. Comparison of measured acoustic dose and model predictions

Recordings of the sonar signals on the DTAG and bottom-moored acoustic recorder allowed for a quantitative model-measurement comparison to assess the validity of the modelled acoustic dose.

Model predictions of the SPL200ms for the DTAG were systematically higher (> 6 dB) than those measured on the DTAG, and the model-measurement mismatch strongly depended on the assumed SSP. The predicted maximum SPL over the entire exposure session (SPLmax) was closer to the observed value on the DTAG than the SPL at any given individual transmission (SPL200ms).

509 The better match between measured and predicted SPL using the CTD cast near the source
510 location (A) than using a CTD cast made near the animal location (D), may be attributed to
511 the fact that the sound propagation condition there indicated a strongly downward refracted
512 path of the transmitted sound. The strength of the downward refraction strongly depended on
513 the which SSP was used (Fig. 4). A representative measurement of the local SSP conditions
514 may therefore explain the better match of the model, because the conditions at greater depth
515 were more similar between the two locations, and less likely to lead to big differences
516 between the two model predictions. However, since the SSPs were obtained at different times,
517 and neither of them during the actual transmissions, it cannot be conclusively stated that the
518 model predictions using the CTD cast near the source location represented the best estimate of
519 the SPLs on the acoustic recorder and satellite tags.

520 Overall, the range of levels predicted close (within 50 m) to the bottom-moored acoustic
521 recorder location was similar to those predicted within a larger area (4000 m) around the
522 acoustic recorder in which animals could be detected. This was mainly due to the exact
523 vertical location of the recorder being somewhat uncertain, and the large variability in the
524 vertical distribution of the modelled sonar sound field. However, the resulting CDFs
525 computed using the CTD cast that provided predicted levels closest to the measured
526 SPL200ms on the recorder, suggested that the highest values predicted around the recorder
527 (91 dB re 1 μPa; 95^{th} percentile) were much lower than the upper range predicted for diving
528 animals (104 dB re 1 μPa; 95^{th} percentile). This illustrates that the levels recorded on the
529 acoustic recorder are not necessary representative of the SPLs received by the animals
530 associated with the observed cessation of foraging around the recorder.

531

532

533 **B. Characterizing uncertainty on the modelled acoustic dose**

534 Characterizing uncertainties associated with sound propagation modelling is challenging and
535 is a continuing point of study on which a large body of literature exists (Colosi et al., 1999;
536 Lynch et al., 2003; Finette, 2006; Lermusiaux et al., 2010; Pecknold and Osler, 2012).
537 Comparisons between a measured quantity and predictions of that quantity made by computer
538 models should consider the effects of imperfections in both predictions (environmental
539 uncertainties and variability, model accuracy) as well as measurements (e.g. acoustic tags
540 might be shielded by the body; Madsen et al., 2006; Wensveen, 2012).

541 To predict the sound field incident on animals, the propagation model BELLHOP required a
542 description of the ocean environment, the acoustic source and the animal location. The
543 acoustic scenario in our study was one of propagation from a shallow source to a receiver that
544 was at a depth of up to ~1000 to 2500 m. The model predicted high spatial variability of SPL
545 in deep (Fig. 4). Consequently, both the depth and distance of the animals were critical factors
546 in determining the SPL in the animal's vicinity. The SSP changed the paths followed by
547 beams of high SPL, and therefore uncertainty in seawater sound-speed led to high uncertainty
548 in SPL at any given location. Uncertainty in SPL would therefore generally be lower for
549 sound sources with a broader vertical beam than the vertical line array used here.

550 BELLHOP required a description of the acoustic source in terms of its level and directivity.
551 Because of the slight drift of the sailing vessel deploying the source, the source array suffered
552 a tilt from the vertical that might have been significant with respect to its directivity. A small
553 tilt correction was applied in the model, but the sensitivity of the model results for tilt
554 variations was not further investigated.

555 The sound propagation model indicated that the dominant sound-paths interacted with the
556 seabed and reflectivity of the seabed, driven by the sediment's composition, were an

557 important factor in determining the received SPL. Seabed acoustic properties are generally not
558 well known, and this is particularly the case for deep environments where seabed sampling is
559 difficult. The values of "fine sand" used represent a reasonable estimate but the actual
560 properties will vary with position, particularly in areas of significant seabed slope.
561 Uncertainty in seabed reflection properties is likely to be high in all deep-water environments.
562 An attempt to capture this uncertainty could involve the use of different values of the
563 sediment properties applied at different locations. A more sophisticated, location-dependent
564 model of seabed properties would require ground-truth information that is unlikely to be
565 available for the foreseeable future in our study area. However, given the small number of
566 bottom interactions for the important rays, and the low grazing angle relative to the seafloor
567 (Fig. 4), it can be expected that in the conditions of this experiment this uncertainty has a
568 relatively small effect on the predicted levels around the animals and recorders. The depth of
569 the ocean was also subject to some uncertainty, but few studies have considered it to date as it
570 is usually one of the better-known input parameters (e.g. Lermusiaux et al., 2010).

571 Seawater sound-speed profiles were measured, but due to logistical restrictions of the vessel
572 used during the experiment, only a limited number of CTD casts could be obtained in the
573 experimental area and time window. Alternatively, SSPs can be predicted using numerical
574 ocean models. However, a significant mismatch was observed between the ocean-model
575 prediction and the measured data from the CTD casts (Fig. A2). This was likely a result of the
576 high oceanographic variability in the region around Jan Mayen.

577 Ocean-model data describing the sound speed, density and salinity profiles represents a high-
578 quality estimate of environmental conditions but even this level of data is shown to have
579 mismatches with data gathered in-situ (see Appendix A). An estimate of the uncertainty
580 associated with the seawater sound speed could in principle be obtained using model outputs
581 produced at different stages of the "forecast-nowcast-hindcast" process. This would require

repeated runs of BELLHOP, each using multiple sound-speed profiles. Although the details of the propagation paths predicted in each model-output dataset would differ, the general conditions of high spatial variability in SPL would not change. This means that the uncertainty in the animal's location, especially when the animals were deep diving, was one of the dominant sources of uncertainties in predicted SPL in this experiment.

The presence of range dependent sound speed profiles was not accounted for in the model predictions presented here, but could affect the sound propagation. Because to the number of CTD casts obtained in the experimental area and time window was very limited, the actual change in SSPs along the direction towards the tagged animals and recorder could not be calculated. Although this would affect the predicted levels for individual transmissions, the maximum received level during the exposure would be less affected by these uncertainties, as these would be less sensitive to the exact time at which an animal reaches a certain depth with high sonar intensities. Future studies should assess the effect of neglecting the range dependence SSP on the predicted SPLs.

A full assessment of the causes of data spread in the model output- would require extensive sensitivity analysis, and ideally also include complementary sound propagation models (e.g. parabolic equation-based models). This process was complicated by the fact that the impact of uncertainty in one parameter was affected by the values chosen for other parameters. That is, the importance of lack of knowledge regarding the seabed was affected by the seawater SSP and the water depth. In the anticipation of such a study, it can be conjectured that the location of the animals was likely to be a very strong driver in determining the uncertainty of SPL predictions. This is illustrated in Figure 7 where the error bars arising from depth uncertainty overlap between calculations for the two SSPs. It should be noted that the use of two SSPs that are "extreme" in terms of their measurement locations does not guarantee that acoustic predictions provide brackets within which the actual values lie. Nonetheless, the acoustic

propagation paths shown in Figure 4 illustrate how any uncertainty in the animal location equates to a very high uncertainty in predicted SPL.

The model-measurement comparison (Appendix A) provided useful insight for interpreting the differences between the different CTD measurement locations and suggested that the oceanographic hind-cast was limited in predicting the sound propagation at the time of the exposure. We could not determine whether this was specific for this location and time, and more systematic studies are required to assess the optimal method to incorporate oceanographic models and measurements in predicting sound propagation in acoustically challenging environments.

C. Uncertainties in animal location

The results from this study demonstrated how uncertainty in animal location translates into a wide range in sound dosage associated with the response to the sonar, which depended on the sensor used for detecting the response.

The uncertainties of the estimated acoustic dose around the acoustic recorder were determined by the uncertainty in location of the animal detected on the recorder. The assumptions made here can be improved on, for instance by modelling detection range explicitly (although we did not find sensitivity to assumed detection distance), or by using more sophisticated agent-based modelling methods based on empirical dive data (e.g. Langrock et al., 2013). In our current approach, low-resolution dive data transmitted by satellite tags were used for simulating the variation in dive behaviour (because of good temporal coverage), but this could also be achieved by using shorter-duration, high-resolution dive profiles from a greater number of DTAGs. Given that the two methods applied (CDF of the SPL in the water column, or Monte-Carlo approach using satellite tag baseline information) did not result in a different range of SPLs, the use of such methods is not likely to reduce much the spread in the

predicted acoustic dose. Methods that estimate location of the animals, i.e. through passive localization, could be used to estimate location of animals at the time of exposure, which requires more advanced system design (i.e. Moretti et al., 2014; Gassmann et al., 2015) and sufficient SNR, which may not work for fainter clicks detected on the recorder, due to the highly directional nature of the echolocation clicks of beaked whales (e.g. Wahlberg et al., 2011; Shaffer et al., 2013). However, such an approach would also likely increase the statistical power to detect responses by the individual clicking animals compared to single hydrophone recordings.

The depth and range uncertainty of the low-resolution satellite tag data were a major cause of large uncertainties in the predicted SPLs. The method described here to quantify the depth uncertainty assumed that the depth uncertainties between transmissions were uncorrelated, although these may be correlated in practice. Dive uncertainties may be estimated by downsampling and summarizing dive profiles with higher time and spatial resolutions or using double tagging experiments. So far the limited amount of data which can be transferred via the Argos satellites forces researchers to make tradeoffs between time resolution and time coverage. This trade off will be different if base stations are used to bypass the ARGOS satellites (e.g. Mote from Wildlife Computers), but such antenna-based systems are mostly useful in areas with high vantage points and for animals with high site fidelity.

D. Relating acoustic dose to measured behavioural responses

The acoustic doses reported here were associated with different types of responses that could be measured with data from the satellite tags and acoustic recorders. High-resolution multi-sensor tags like DTAGs have the temporal resolution to obtain precise estimates for the onset of a response. However, for the satellite tags used here the precise onset times of the responses to the sonar, and therefore the associated received levels, were not possible to

establish due to the tags' courser temporal resolution. Sound pressure levels at onset of response may therefore have been lower, as potential cessation of feeding and small-scale changes in dive behaviour during the ramp-up could not be detected using the satellite tags (Wensveen et al. 2019).

Model predictions based upon satellite tag positions contain large uncertainties for individual transmissions (Fig. 7). Estimate of maximum received level over the entire exposure period were more robust against dive uncertainties, and assumed SSP, as they were less sensitive to timing when an animal is at depth with higher sound intensities.

For single acoustic recorders as used in the Jan Mayen experiments, responses in vocal behaviour can only be reliably measured in areas with high animal presence, and strong long-lasting responses (such that it could be distinguished from natural variability in click presence detected on the recorder; Wensveen et al., 2019). The observed cessation of echolocation clicks on the acoustic recorder was associated with lower modelled SPLs than the maximum SPLs predicted for the satellite tags due to the ramp-up in source level and the difference in temporal resolution between the two types of sensors. Since animals cannot be tracked acoustically while silent, the SPLmax experienced by animals near the acoustic recorder could not be established, as the direction in which they might have moved is unknown. The modelled SPLmax within the detection distance of the recorder may therefore have overestimated (or underestimated) the SPLmax associated with the response if the animals moved away from (or towards) the source. Although it is possible that all animals near the recorder became silent but did not avoid the sound source during the sonar exposure, this seems unlikely, since avoidance of the exposure area is typical for this species at levels similar to the SPLmax that was predicted at recorder location (Miller et al., 2015; Wensveen et al., 2019).

679 The acoustic dose estimated in this study is a best estimate, but given the environmental
680 uncertainties and variability associated with this experiment, these estimates should be taken
681 with some caution. Deep diving species such as beaked whales could be attracted to dynamic
682 regions with upwelling of nutrient-rich water because of increased biological productivity,
683 and therefore are often naturally found in acoustically complex environments. Studies aimed
684 at these species could therefore suffer from larger uncertainties in sound propagation and
685 resulting estimated sound dose. It is recommended to carefully characterize the environment
686 and uncertainties associated with propagation conditions when using satellite tags and
687 acoustic recorders in challenging environments, such as oceanographic frontal zones.

688 The extent to which satellite tags and acoustic recorders add value to quantify dose-response
689 relationships of effects of sound on marine mammals depends on a balance of the quality
690 versus quantity of the data collected. Acoustic recorders and satellite tags offer practical
691 benefits compared to high-resolution DTAGs, such as the practicality of deploying multiple
692 tags without compromising data recovery (i.e. through satellite link), which allows for
693 collecting data from multiple individuals over large spatial scales. In addition, they provide a
694 benefit of monitoring over much longer timescales. Disadvantages include sampling of a
695 limited aspect of the behavioural response and, for satellite tags, lower-resolution
696 observations, periods of missing data and less developed analysis methods and, for acoustic
697 recorders, less power to detect responses. The current study demonstrated the large range of
698 values in the acoustic dose associated with the observed response to sound. The methods to
699 incorporate positional uncertainty of animal locations presented in this study can be used to
700 make a quantitative power analyses to assess the added benefit of these devices for future
701 controlled exposure studies, which are likely to be species and site specific. This is especially
702 relevant in conditions where the environmental conditions are highly uncertain and variable,
703 which will limit the accuracy to which received levels can be predicted.

V. CONCLUSION

This study quantified the uncertainty in the estimated acoustic dose associated with responses of northern bottlenose whales monitored with satellite tags and a bottom-mounted acoustic recorder to controlled sonar exposure in the Jan Mayen area in 2016. The effects of uncertain animal location, particularly in the vertical plane (depth), and uncertainties in environmental conditions on estimated exposure levels were assessed.

Some recommendations for future studies can be obtained from the findings of this study. For satellite tags, it is recommended to increase depth resolution or implement flexible programmable dive summary algorithms to avoid unnecessarily large depth uncertainties. Future studies could also consider developing 'acoustic smart tags', satellite tags with on-board processing for measuring the acoustic dose, such that the limited data can be transmitted through a satellite/ARGOS network. Until this is in place, it is highly recommended that studies looking at effects of anthropogenic sound on marine mammals using tags without acoustic sensors, measure sound speed profiles in situ at regular spatial and temporal intervals to sample the environmental variability, and deploy acoustic sensors elsewhere in the water column to estimate the accuracy of the modelling.

ACKNOWLEDGMENTS

We would like to thank everyone involved in the Jan Mayen field work, including Captain Christian Harboe-Hansen and the rest of the ship's crew, Lars Kleivane, Sander van IJsselmuide, Tomoko Narazaki, Miguel Neves dos Reis, Eilidh Siegal, Mike Williamson, Charlotte Curé, Stacy DeRuiter, Peter Tyack, Sacha Hooker, Eva Hartvig, Naomi Boon, Joanna Kershaw, Dave Moretti and Ron Morrissey. AvBB thanks Mathieu Colin for his valuable advice on the use of BELLHOP. The research described in this paper was supported by US Office of Naval Research, US Strategic Environmental Research and Development

Program (SERDP RC-2337), the French Ministry of Defence (DGA), and the Netherlands Ministry of Defence.

APENDIX A: OCEANOGRAPHIC CONDITIONS IN JAN MAYEN AREA DURING AND AROUND THE EXPOSURE EXPERIMENT

A. Copernicus Marine Environment Monitoring Service (CMEMS) predictions of sound speed profiles

In order to interpret the spatial and temporal variability of the observed SSPs, a Copernicus Marine Environment Monitoring Service (CMEMS) (von Schuckmann et al., 2017) hind-cast was carried out (on 13 November 2017) for 9 and 10 June 2016. The CMEMS hind-cast has a resolution of 12 hours, and a spatial grid of (1/12)º. The sea surface potential temperature map suggests that the source transmission location was in the middle of a strong temperature transition region between warmer salty waters, and colder, less saline waters (Fig. A1).

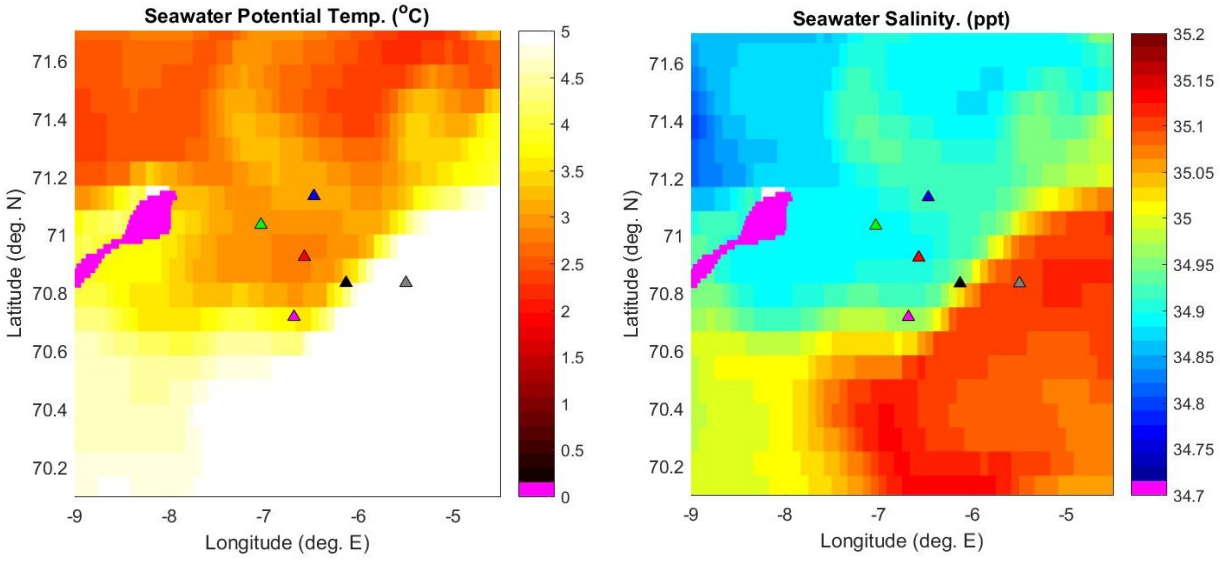
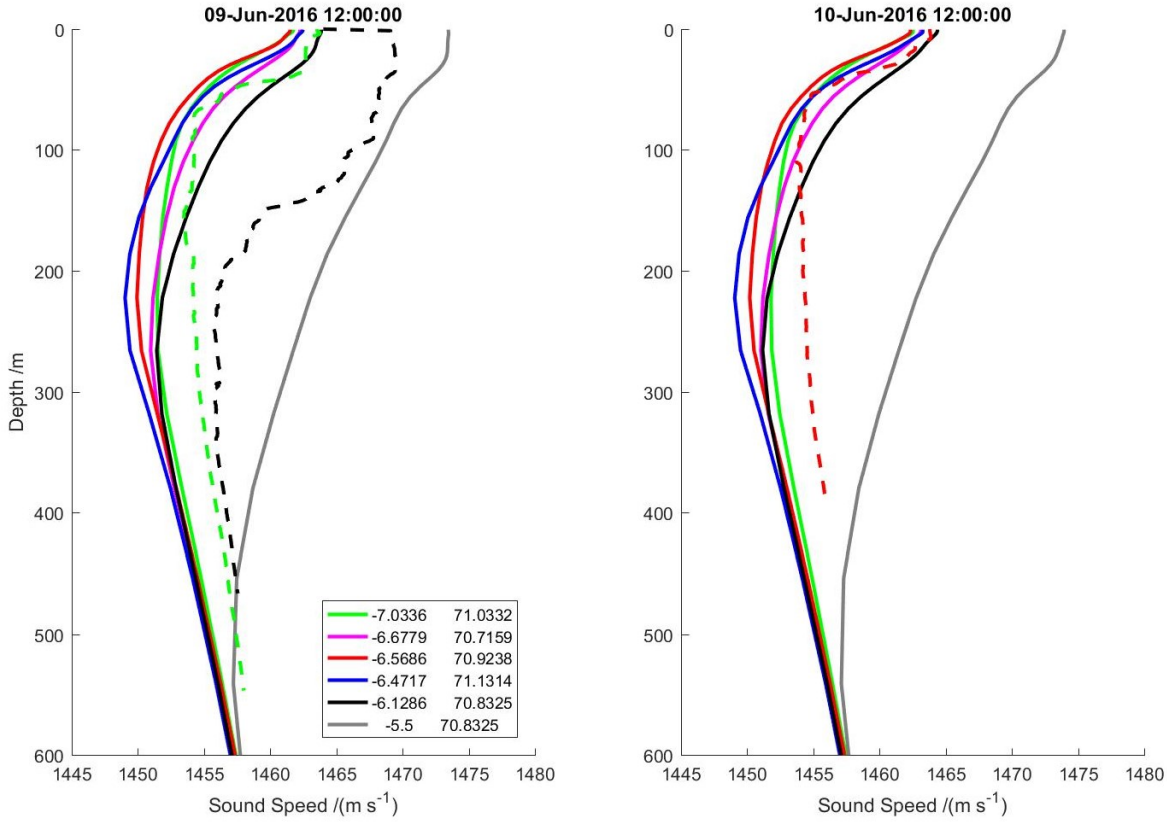


FIG. A1. Hind-casts seawater potential temperature, and salinity in the experimental area on 9 June. White triangles indicate locations where the SSPs were measured with the CTD.

744 Comparison of the measured SSPs from the CTD with modelled SSPs suggest that the range
745 of modelled SSPs was reasonable, but that the measured SSPs near the source transmission
746 location was more consistent with warmer and saline region further east of the front (Fig. A2).
747 This model-measurement mis-match provided useful insight for interpreting the differences
748 between the different CTD measurement locations but suggested that the hind-cast had
749 limited added-value in improvement the accuracy of the predicted sound propagation at the
750 time of the exposure.



751

752 FIG. A2. Hind-casts for CTD locations on 9-10 June. The black dashed line was measured
753 near source transmission (A) (on a different day), and the green and red dashed line measured
754 on 9 and 10 June closer to the northern recorder location (B and C). Colour coding is as in
755 Fig. 1. The grey SSP corresponded to the predicted SSP in warmer waters further to the east
756 of transmission site (Fig. A1).

757 **REFERENCES**

758 Antunes, A., Kvadsheim, P.H, Lam, F.P.A., Tyack, P.L., Thomas, L., Wensveen, P.J., and
759 Miller, P.J.O. (**2014**). "High thresholds for avoidance of sonar by freeranging long-finned
760 pilot whales (Globicephala melas)," Marine Pollution Bulletin 83 (1), 165–180.

761 Bourke, R.H., Paquette, R.G., and Blythe, R.F. (**1992**). "The Jan Mayen Current of the
762 Greenland Sea," J. Geophys. Res., 97(C5), 7241–7250, doi:10.1029/92JC00150.

763 Carter, M.I.D., Bennett, K. A., Embling, C.B., Hosegood, P.J., and Russell, D.J.F. (**2016**).
764 "Navigating uncertain waters: a critical review of inferring foraging behaviour from location
765 and dive data in pinnipeds," Movement Ecology, 4(1), 25, doi:10.1186/s40462-016-0090-9.

766 Cooke, S.J., Hinch, S.G., Wikelski, M., Andrews, R.D., Kuchel, L.J., Wolcott, T.G., and
767 Butler, P.J. (**2004**). "Biotelemetry: a mechanistic approach to ecology," Trends in Ecology &
768 Evolution, 19 (6), 334–343, doi:10.1016/j.tree.2004.04.003.

769 Colosi, J.A., Scheer, E.K., Flatté, S.M., Cornuelle, B.D., Dzieciuch, M.A., Munk, W.H., ... &
770 Metzger, K. (**1999**)." Comparisons of measured and predicted acoustic fluctuations for a
771 3250-km propagation experiment in the eastern North Pacific Ocean," The Journal of the
772 Acoustical Society of America, 105(6), 3202-3218.

773 Dekeling, R.P.A., Tasker, M.L., Van der Graaf, A.J., Ainslie, M.A, Andersson, M.H., André,
774 M., Borsani, J.F., Brensing, K., Castellote, M., Cronin, D., Dalen, J., Folegot, T., Leaper, R.,
775 Pajala, J., Redman, P., Robinson, S.P., Sigray, P., Sutton, G., Thomsen, F., Werner, S.,
776 Wittekind, D., and Young, J.V. (**2014**). "Monitoring Guidance for Underwater Noise in
777 European Seas, Part III: Background Information and Annexes," JRC Scientific and Policy
778 Report EUR 26556 EN, Publications Office of the European Union, Luxembourg, doi:
779 10.2788/2808.

780 Department of the Navy (**2013**). "Atlantic fleet training and testing: Fina environmental
781 impact statement/overseas environmental impact statement (FEIS/OEIS)" (Department of the
782 Navy, Washington, DC), p. 590.

783 DeRuiter, S.L., Southall, B.L., Calambokidis, J., Zimmer, W.M.X., Sadykova, D., Falcone,
784 E.A., Friedlaender, A.S., Joseph, J.E., Moretti, D., Schorr, G.S., Thomas, L., and Tyack, P.L.
785 (**2013**). "First direct measurements of behavioural responses by Cuvier's beaked whales to
786 mid-frequency active sonar," Biol. Lett. 9, 2- 6. doi:10.1098/rsbl.2013.0223.

787 Ellison, W.T., Southall, B.L., Clark, C.W. and Frankel, A.S. (**2012**). "A new context-based
788 approach to assess marine mammal behavioural responses to anthropogenic sounds,"
789 Conserv. Biol. 26, 21-28.

790 Falcone, E.A., Schorr, G.S., Watwood, S.L., DeRuiter, S.L., Zerbini, A.N., Andrews, R.D.,
791 Morrissey, R.P., Moretti, D.J. (**2017**). "Diving behaviour of Cuvier's beaked whales exposed
792 to two types of military sonar," R. Soc. open sci. 4: 170629. doi:10.1098/rsos.170629.

793 Finette, S. (2006). "A stochastic representation of environmental uncertainty and its coupling
794 to acoustic wave propagation in ocean waveguides," The Journal of the Acoustical Society of
795 America, 120(5), 2567-2579.

796 Frasier, K.E., Wiggins, S.M., Harris, D., Marques, T.A., Len Thomas, L., and Hildebrand,
797 J.A. (**2016**). "Delphinid echolocation click detection probability on near-seafloor sensors," J.
798 Acoust. Soc. Am. 140(3), 1918–1930.

799 Gassmann, M., Wiggings, S.M., and Hildebrand, J.A. (2015). "Three-dimensional tracking of
800 Cuvier's beaked whales' echolocation sounds using nested hydrophone arrays," J. Acoust.
801 Soc. Am. 128 (4), 2483-2494.

802 Goldbogen, J.A., Southall, B.L., DeRuiter, S.L., Calambokidis, J., Friedlaender, A.S., Hazen,
803 E.L., Falcone, E.A., Schorr, G.S., Douglas, A., Moretti, D.J., Kyburg, C., McKenna, M.F.,
804 and Tyack, P.L. (**2013**). „Blue whales respond to simulated mid-frequency military sonar,"
805 Proceedings of the Royal Society B 280(1765):20130657.

806 Götz, T., and Janik, V. (**2011**). "Repeated elicitation of the acoustic startle reflex leads to
807 sensitization in subsequent avoidance behaviour and induces fear conditioning," BMC
808 Neuroscience 2011, 12, 30.

809 Harris, C.M., Thomas, L., Falcone, E.A., Hildebrand, J., Houser, D., Kvadsheim, P.H., Lam,
810 F.-P.A., Miller, P.J.O., Moretti, D.J., Read, A.J. et al., (**2017**). "Marine mammals and sonar:
811 dose-response studies, the risk-disturbance hypothesis and the role of exposure context," J.
812 Appl. Ecol. 55, 396–404. Doi:10.1111/1365-2664.12955.

813 Harris, C.M., Sadykova, D., DeRuiter, S.L, Tyack, P.L.., Miller, P.J.O., Kvadsheim, P.H.,
814 Lam, F.P.A., and Thomas, L. (**2015**). "Dose response severity functions for acoustic
815 disturbance in cetaceans using recurrent event survival analysis," Ecosphere 6(11):236.
816 Doi:10.1890/ES15-00242.1.

817 Heathershaw, A. D., C. E. Stretch, and S. J. Maskell. (**1991**). "Coupled ocean-acoustic model
818 studies of sound propagation through a front," The Journal of the Acoustical Society of
819 America 89 (1), 145-155.

820 Hooker, S., and Whitehead, H. (**2002**). "Click characteristics of northern bottlenose whales
821 (Hyperoodon amplullatus)," Marine Mammal Sci. 18(3), 69–80.

822 Houser, D.S., Martin, S.W., and Finneran, J.J. (**2013**). "Behavioural responses of California
823 sea lions to mid-frequency (3250-3450 Hz) sonar signals," Marine Environmental Research
824 92 (2013) 268-278.

825 ISO (**2017**). ISO 18405 Underwater Acoustics—Terminology (International Organization for
826 Standardization, Geneva, 2017).

827 Johnson, M.P. and Tyack, P.L. (**2003**). "A digital acoustic recording tag for measuring the
828 response of wild marine mammals to sound," IEEE J. Ocean. Eng. 28, 3-12.

829 Katsnel'son, B. G., Lynch, J., and Tshoidze, A.V. (**2007**). "Space-frequency distribution of
830 sound field intensity in the vicinity of the temperature front in shallow water," Acoustical
831 Physics 53, no. 5 (2007): 611-617.

832 Langrock, R., Marques, T.A., Baird, R.W., and Thomas, L. (**2013**). "Modeling the Diving
833 Behaviour of Whales: A Latent-Variable Approach with Feedback and Semi-Markovian
834 Components," Journal of Agricultural, Biological, and Environmental Statistics, DOI:
835 10.1007/s13253-013-0158-6.

836 Lermusiaux, P.F., Xu, J., Chen, C.F., Jan, S., Chiu, L.Y., and Yang, Y.J. (**2010**). Coupled
837 ocean–acoustic prediction of transmission loss in a continental shelfbreak region: Predictive
838 skill, uncertainty quantification, and dynamical sensitivities. IEEE Journal of Oceanic
839 Engineering, 35(4), 895-916.

840 Lynch, J. F., Newhall, A. E., Sperry, B., Gawarkiewicz, G., Fredricks, A., Tyack, P., ... and
841 Abbot, P. (**2003**). Spatial and temporal variations in acoustic propagation characteristics at the
842 New England shelfbreak front. IEEE journal of oceanic engineering, 28 (1), 129-150.

843 Madsen, P.T. (2005). Marine mammals and noise: problems with root mean square sound
844 pressure levels for transients. Journal of the Acoustical Society of America, 117, 3952–3957.

845 Madsen, P.T., Johnson, M., Miller, P.J.O., Aguilar Soto, N., Lynch, J., and Tyack, P. (2006).
846 Quantitative measures of air-gun pulses recorded on sperm whales (Physeter macrophalus)

using acoustic tags during controlled exposure experiments. Journal of the Acoustical Society of America, 120, 2366–2379.

Martin, S.W., Martin, C.R., Matsuyama, B., and Henderson, E.E. (**2015**). "Minke whales (*Balaenoptera acutorostrata*) respond to navy training," J. Acoust. Soc. Am. 137 (5), 2533–2541.

Marques, T.A., Thomas, L., Ward, J., DiMarzio, N., and Tyack, P.L. (**2009**). "Estimating cetacean population density using fixed passive acoustic sensors: An example with Blainville's beaked whales," J. Acoust. Soc.Am. 125(4), 1982–1994.

McCarthy, E., Moretti, D., Thomas, L., DiMarzio, N., Morrissey, R., Jarvis, S.,Ward, J., Izzi, A., and Dilley, A.. (**2011**). "Changes in spatial and temporal distribution and vocal behaviour of Blainville's beaked whales (Mesoplodon densirostris) duringmultiship exercises with mid-frequency sonar," Mar. Mamm. Sci. 27, E206–E226, doi:10.1111/j.1748-7692.2010. 00457.x.

Miller, P.J.O., Kvadsheim, P.H., Lam, F.P.A., Tyack, P.L., Curé, C., DeRuiter, S.L., Kleivane, L., Sivle, L.D., van IJsselmuide, S.P., Visser, F., Wensveen, P.J., von Benda-Beckmann, A.M., Martín López, L.M., Narazaki, T., and Hooker, S.K. (**2015**). "First indications that northern bottlenose whales are sensitive to behavioural disturbance from anthropogenic noise. R. Soc. Open Sci. 2, 140484, doi:10.1098/rsos. 140484.

Manzano-Roth, R., Henderson, E.E., Martin, S.W, Cameron, M., Matsuyama, B.M. (**2016**). "Impacts of U.S. Navy Training Events on Blainville's Beaked Whale (Mesoplodon densirostris) Foraging Dives in Hawaiian Waters," Aquatic Mammals 2016, 42(4), 507-518, doi:10.1578/AM.42.4.2016.507.

Moretti, D., Thomas, L., Marques, T., J., Harwood, Dilley, A., Neales, B., Shaffer, J., McCarthy, E., New, L., Jarvis, S., and Morrissey, R. (**2014**). "A risk function for behavioural

disruption of Blainville's beaked whales (Mesoplodon densirostris) from mid-frequency active sonar," PLoS ONE 9, e85064. doi:10.1371/ journal.pone.0085064.

Mork, K.A, Drinkwater, K.F., Jónsson, S. Valdimarsson, H., Ostrowski, M. (**2014**). "Watermass exchanges between the Norwegian and Iceland seas over the Jan Mayen Ridge using in-situ current measurements," Journal of Marine Systems 139, 227–240.

Nowacek, D.P., Johnson, M.P., and Tyack, P.L., (**2004**). "North Atlantic right whales (Eubalaena glacialis) ignore ships but respond to alerting stimuli,", Proc. R. Soc. Lond. B 2004 271 227-231, Doi:10.1098/rspb.2003.2570.

Pecknold, S., and Osler, J.C. (**2012**). "Sensitivity of acoustic propagation to uncertainties in the marine environment as characterized by various rapid environmental assessment methods," Ocean Dynamics, 62(2), 265-281.

Porter, M.B. and Bucker, H.P. (**1987**). "Gaussian beam tracing for computing ocean acoustic fields," J. Acoust. Soc. Am. 82, 1349-1359.

Rudels, B., Bjfrk, G., Nilsson, J., Winsord, P., Lakec, I., and Nohr, C., (**2005**). "The interaction between waters from the Arctic Ocean and the Nordic Seas north of Fram Strait and along the East Greenland Current: results from the Arctic Ocean-02 Oden expedition", Journal of Marine Systems 55, 1– 30.

Southall, B.L., Moretti, D., Abraham, B., Calambokidis, J., and Tyack, P.L. (**2012**). "Marine mammal behavioural response studies in Southern California: Advances in technology and experimental methods," Mar Technol Soc J 46, 48−59.

Southall1, B.L., Nowacek, D.P., Miller, P.J.O., Tyack, P.L. (**2016**). "Experimental field studies to measure behavioural responses of cetaceans to sonar," Endang Species Res Vol. 31, 293–315, doi: 10.3354/esr00764.

893 Schorr, G.S., Falcone, E.A., Moretti, D.J., and Andrews, R.D. (**2014**) "First long-term
894 behavioural records from Cuvier's Beaked whales (Ziphius cavirostris) reveal record-breaking
895 dives," PLoS ONE 9, e92633, doi:10.1371/journal.pone.0092633.

896 Shapiro, G., Chen, F., and Thain. R. (**2014**). "The effect of ocean fronts on acoustic wave
897 propagation in the Celtic Sea," Journal of Marine Systems 139, 217-226.

898 Tyack, P.L., Zimmer, W.M.Z., Moretti, D., Southall, B.L., Claridge, D.E., Urban, J.W., Clark,
899 C.W., D'Amico, A., DiMArzio, N., Jarvis, S., McCarthy, E., Morrissey, R., Ward, J., and
900 Boyd, I.L. (**2011**). "Beaked whales respond to simulated and actual navy sonar," PLoS ONE
901 6, e17009, doi:10.1371/journal.pone.0017009.

902 Tomkiewicz, S.M., Fuller, M.R., Kie, J.G., and Bates, K.K. (**2010**). "Global positioning
903 system and associated technologies in animal behaviour and ecological research.,"
904 Philosophical Transactions of the Royal Society B: Biological Sciences, 365(1550), 2163–
905 2176. Doi:10.1098/rstb.2010.0090.

906 Tougaard, J., Carstensen, J., Teilmann, J., Skov, H., Rasmussen, P. (**2009**). „Pile driving zone
907 of responsiveness extends beyond 20 km for harbor porpoises (Phocoena phocoena (L.)," (L)
908 J. Acoust. Soc. Am. 126 1, 11-14.

909 Urick, R.J. (**1975**). Principles of underwater sound for engineers. 2nd ed. New York:
910 McGraw-Hill.

911 von Benda-Beckmann, A.M., Lam, F.P.A., Moretti, D.J., Fulkerson, K., Ainslie, M.A., van
912 IJsselmuide, S.P., Theriault, J., and Beerens, S.P. (**2010**). "Detection of Blainville's beaked
913 whales with towed arrays," Appl. Acoust. 71, 1027–1035.

von Schuckmann, K., Le Traon, P.Y., Alvarez-Fanjul, E., Axell, L., Balmaseda, et al. (**2017**). "The Copernicus Marine Environment Monitoring Service Ocean State Report," Journal of Operational Oceanography, 9:sup2, s235-s320, Doi:10.1080/1755876X.2016.1273446.

Wahlberg, M., Beedholm, K., Heerfordt, A., and Møhl, B. (**2011**). "Characteristics of biosonar signals from the northern bottlenose whale, Hyperoodon ampullatus," J. Acoust. Soc. Am. 130 (5), 3077-3084.

Ward, J., Morrissey, R., Moretti, D., DiMarzio, N., Jarvis, S., Johnson, M., Tyack, P., and White, C. (**2008**). "Passive acoustic detection and localization of Mesoplodon densirostris (Blainville's beaked whale) vocalizations using distributed bottom-mounted hydrophones in conjunction with a Digital Tag (DTAG) recording," Can. Acoust. 1, 60–66.

Ward, J., Jarvis, S., Moretti, D., Morrissey, R., DiMarzio, N., Johnson, M., Tyack, P., Thomas, L., and Marques, T. (**2011**). "Beaked whale (Mesoplodon densirostris) passive acoustic detection in increasing ambient noise," J. Acoust. Soc. Am. 129 (2), 662–669.

Ward Shaffer, J., Moretti, D., Jarvis, S., Tyack, P., and Johnson, M. (**2013**). "Effective beam pattern of the Blainville's beaked whale (Mesoplodon densirostris) and implications for passive acoustic monitoring," Acoust. Soc. Am. 133 (3), 1770-1784.

Wensveen, P. J. (2012). The effects of sound propagation and avoidance behaviour on naval sonar levels received by cetaceans. MPhil Thesis, University of St Andrews. Retrieved from http://hdl.handle.net/10023/3194

Wensveen, P. J. (2016). Detecting, assessing and mitigating the effects of naval sonar on cetaceans. PhD Thesis, University of St Andrews. Retrieved from http://hdl.handle.net/10023/8684

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

939 Wensveen, P.J., Kvadsheim, P.H., Lam, F.P.A., von Benda-Beckmann, A.M., Sivle, L.D.,
940 Visser, F., Curé, C., Tyack, P.L., and Miller, P.J.O. (2017). "Lack of behavioural responses of
941 humpback whales (Megaptera novaeangliae) indicate limited effectiveness of sonar
942 mitigation," Journal of Experimental Biology, 220, 4150-4161, doi:10.1242/jeb.161232.

943 Wensveen, P.J., Isojunno, S., Hansen, R.R., von Benda-Beckmann, A.M., Kleivane, L., van
944 IJsselmuide, S., Lam, F.P.A., Kvadsheim, P.H., DeRuiter, S.L., Curé, C., Narazaki, T., Tyack,
945 P. L., Miller. P.J.O (2019). "Northern bottlenose whales in a pristine environment respond
946 strongly to close and distant navy sonar signals" Royal Soc. Proc. B (in press).

947 Zimmer, W., Harwood, J., Tyack, P., Johnson, M., and Madsen, P. (2008). "Passive acoustic
948 detection of deep-diving beaked whales," J. Acoust. Soc. Am. 124, 2823–2832.