Aquatic Mammals 2019, *45*(6), 675-690, DOI 10.1578/AM.45.6.2019.675

Accounting for Positional Uncertainty When Modeling Received Levels for Tagged Cetaceans Exposed to Sonar

Robert S. Schick,[1] Matthew Bowers,[2] Stacy DeRuiter,[3] Ari Friedlaender,[4,5] John Joseph,[6] Tetyana Margolina,[6] Douglas P. Nowacek,[7] and Brandon L. Southall[4,5]

[1]*Marine Geospatial Ecology Lab, Nicholas School of the Environment, Duke University, Durham, NC 27708, USA*
E-mail: rss10@duke.edu
[2]*Department of Fish, Wildlife and Conservation Biology, Colorado State University, Fort Collins, CO 80523, USA*
[3]*Department of Mathematics and Statistics, Calvin University, Grand Rapids, MI 49546, USA*
[4]*Institute for Marine Sciences, Department of Ecology and Evolutionary Biology, University of California Santa Cruz, Santa Cruz, CA 96060, USA*
[5]*Southall Environmental Associates, Aptos, CA 95003, USA*
[6]*Oceanography Department, Naval Postgraduate School, Monterey, CA 93943, USA*
[7]*Nicholas School of the Environment & Pratt School of Engineering, Duke University Marine Lab, 135 Duke Marine Lab Road, Beaufort, NC 28516, USA*

Abstract

Exposure to anthropogenic sound can have a range of negative behavioral and physical effects on marine species and is of increasing ecological and regulatory concern. In particular, the response of marine mammals, and notably the family of cryptic deep-diving beaked whales, to military sonar is a timely and complex issue. To make inference on aspects of response by individual whales to noise of any type, it is critical to either measure or systematically estimate what received levels (RLs) the animal actually experienced. Various tools and techniques exist to monitor RLs and associated responses, each with advantages and disadvantages. Most behavioral response studies to date have used relatively short-term (hours to a few days), high-resolution acoustic tags that provide direct RL measurements. Because of their short duration, these tags do not allow for assessments of longer-duration baseline behavior before and following a disturbance that may tell us more about the nature of response within a broader context for tagged individuals. In contrast, longer-duration (weeks to months), satellite-transmitting tags lack high-resolution kinematic data and the ability to directly measure RL. Herein, we address these issues and efforts to derive robust statistical RL characterizations using animal movement and fine-scale, site-specific sound propagation modeling for longer-duration tags in the context of a behavioral response study off Cape Hatteras, North Carolina. In the autumn of 2017, we tagged nine Cuvier's beaked whales and three short-finned pilot whales and conducted controlled exposure experiments using simulated and operational military mid-frequency active sonar. We used sound propagation modeling methods and modeled positions of individual animals to estimate RLs in four dimensions and to statistically describe uncertainty within volumes of water space where animals were predicted to occur during exposure periods. By properly accounting for positional error in this study, it is clear that previous studies using single median RL estimates drastically underestimate the full range of plausible values; ranges in estimated RLs here often exceeded 40 dB. We also demonstrate how ancillary data from visual focal follows of tagged individuals can significantly narrow estimated RL ranges. Further, we compared measured RLs on a calibrated acoustic tag to modeled RLs at the same position to evaluate our volumetric modeling results. While satellite tags record data over longer time frames, their substantial geospatial error coupled with the unique deep-diving behavior of beaked whales means that estimates of RL can vary broadly and, consequently, that single point estimates from less robust approaches may be substantially in error. Accounting for this uncertainty using robust statistical modeling is critical to fairly characterize variance and effectively assess exposure-response relationships.

Key Words: behavioral response studies, beaked whales, uncertainty, received level, controlled exposure experiment, satellite tag

Introduction

Exposure to human noise, including military sonar systems, can cause varying degrees of disturbance and physical harm in different marine species (Filadelfo et al., 2009; Southall et al., 2016; Bernaldo de Quirós et al., 2019). The underlying mechanisms influencing the probability of disturbance, injury, and/or mortality are complex and not fully understood. For cetaceans exposed to military sonar, responses across various species range from no discernible response, to avoidance and cessation of feeding, to mortal strandings of few to many animals (Southall et al., 2016; Bernaldo de Quirós et al., 2019). Observed and potential impacts, as well as extensive scientific uncertainty, across a range of species have fueled scientific, regulatory, and conservation interest in the impacts of anthropogenic noise on marine mammals (National Research Council [NRC], 2005; Hatch et al., 2016; National Academy of Sciences, Engineering, and Medicine, 2017).

Consequently, considerable research has been conducted over the past two decades on how sonar affects marine mammals. This work has focused heavily on the family of cryptic deep-diving whales known as beaked whales given their disproportionate representation in stranding events (Bernaldo de Quirós et al., 2019) and their heightened behavioral sensitivity to sonar (Tyack et al., 2011; DeRuiter et al., 2013) as well as other noise sources (Aguilar Soto et al., 2006). As a result of this research, the following information has been gathered about beaked whales: (1) their baseline diving patterns from both short-term, high-resolution tags (e.g., Tyack et al., 2011; DeRuiter et al., 2013) and longer-term, lower-resolution satellite tags (Falcone et al., 2017; Shearer et al., 2019); (2) their distribution and abundance (e.g., Yack et al., 2013); (3) their behavioral responses to disturbance in controlled conditions (see Southall et al., 2016); and (4) documented and potential impact of disturbance on vital rates (e.g., Claridge, 2013; New et al., 2013). Despite these and other focused studies, we still lack important details on how animals respond to known sonar exposures. Key research needs include understanding how contextual factors (e.g., spatial relationships between source and receiver) influence behavioral response probability (DeRuiter et al., 2013) and how available data from relatively short-term tag deployments may relate to patterns of behavior and disturbance over longer periods. To better understand population consequences of disturbance—to both individuals and populations—behavioral responses of whales to known disturbance must be measured within the context of longer-term behavioral sampling. In addition, data on key features of the exposure, including noise received levels (hereafter RLs, reported throughout the article in dB re 1 μPa root-mean-square) and frequencies, remain an important consideration.

One way to obtain longer data records is to use tags that remain on the animal for longer periods, thereby allowing researchers to record movement and diving behavior over longer time frames and larger spatial extents in areas where they may be exposed to disturbance in opportunistic uncontrolled conditions (Tyack et al., 2011; Falcone et al., 2017). While they have the advantage of extended monitoring duration and realistic exposure scenarios, these kinds of observational studies have some limitations. For example, their uncontrolled nature means that contextual variables relevant to response probability related to both sources and receivers generally vary within and between exposure instances in ways that are inconsistent and unknown. Further, most modern satellite tags currently lack the ability to record sound, meaning RLs must be estimated based on the geospatial locations of both the animal and the sound source. Thus, tradeoffs ensue in using longer-duration, lower-resolution, non-acoustic tags. What you may gain in the overall duration and spatial resolution, you lose in both temporal resolution and data richness. Therefore, when using satellite tags to evaluate any response to disturbance, one may observe movements over long time periods but lack key information as to the RLs of exposure the animal experienced from either experimental or incidental noise exposures. However, determining RLs is critical to deriving exposure-response relationships for different species and contexts (Tyack et al., 2011; Southall et al., 2016; Harris et al., 2018). To complicate matters, the positional observations (i.e., the *x,y* positions that are recorded) are made with substantial error, and their frequency is subject to a number of factors ranging from tag placement to animal behavior. The satellite tags use the Argos system for communicating positional information, which depends on recording the Doppler shift between an Argos satellite and the tag on the animal. To get an observation made between the tag and the satellite, the animal must be at the surface; and there must be a sufficient number of satellites available that can detect and localize the tagged animal. Animals that are deep divers (e.g., beaked whales) are at the surface for shorter periods of time, thus reducing the number of chances of successful communication between the animal and orbiting satellites. When the links are successful, the observations are still recorded with uncertainty, and addressing this positional uncertainty has been the subject of much research in the animal biotelemetry field. Prior to 2008, each position was assigned an ordinal location quality code (e.g., 3, 2, 1, 0, A, B, and Z); following 2008, Argos provides error ellipses with each location whereby recorded locations with larger ellipses have higher positional uncertainty (McClintock et al., 2014).

Consequently, appropriately incorporating uncertainty from positional observation error and other sources into sound exposure estimates is crucial for adequately quantifying a key aspect of noise exposure in relation to measured behavioral responses to sound. This is also a needed metric in considering the role of other contextual covariates (see Ellison et al., 2012), including spatial relationships such as how RL and source-animal range interact (e.g., Southall et al., 2019a; Wensveen et al., 2019). Accounting for this uncertainty relies on knowing where the animal is in the *x*, *y*, and *z* dimensions (Cox et al., 2006). Herein, we address this issue by using modern inferential techniques for animal movement (Hooten et al., 2017) in conjunction with modern sound propagation tools (Margolina et al., 2018) to estimate RLs during a study involving Cuvier's beaked whales (*Ziphius cavirostris*) and short-finned pilot whales (*Globicephala macrorhynchus*) off Cape Hatteras, North Carolina. A similar recent approach looks at the response of northern bottlenose whales (*Hyperoodon ampullatus*) off Norway (von Benda-Beckmann et al., 2019; Wensveen et al., 2019). The work we report on here is part of a larger experiment called the Atlantic Behavioral Response Study (Atlantic-BRS) which is simultaneously deploying short-term, high-resolution DTAGs (Johnson & Tyack, 2003) and longer-duration satellite tags within an experimental framework (Southall et al., 2016, 2018). In a given BRS, researchers place tags (typically short-term, high-resolution suction cup-attached DTAGs) on animals, expose them to sound (e.g., pseudo-random noise, predator vocalizations, and simulated or actual sonar) using controlled exposure experiments (CEEs), and then record behavioral changes from baseline conditions, as well as whether responses cease following the cessation of exposure. Experimental BRSs have taken place around the globe and have generated a large corpus of data on responses to sound (Southall et al., 2016).

Prior to conducting underwater sonar training and testing activity, the U.S. Navy (2018) carries out extensive computer-based simulations to determine how many individuals are "taken" (as defined under the military readiness definition of level of various impacts under the U.S. Marine Mammal Protection Act of 1972) as a function of exposure to sonar during an exercise. Simulated animals function as dosimeters that log cumulative exposure to all noise sources over the duration of an exercise, based on some characteristics of typical animal behavior and known features of sonar operations. Subsequently, the simulation applies exposure-response relationships that relate characteristics of exposure (e.g., noise RL) to the probability of response based on documented behavioral changes (or lack thereof) (Southall et al., 2007; Miller et al., 2014; Harris et al., 2018; U.S. Navy, 2018) to determine if an animal has been "taken" according to specified criteria. These takes can be purely auditory—that is, was the animal predicted to receive enough sound to exhibit either a temporary or permanent threshold loss of hearing? However, the large majority of predicted impacts are behavioral—for example, did the animal temporarily abandon an area or cease foraging following exposure? Much of the research that has informed these behavioral relationships comes from BRSs (Southall et al., 2016; Harris et al., 2018).

Behavioral changes of interest include changes in diving pattern, cessation of foraging, and horizontal avoidance or displacement—that is, whether the animal moves away from the source following exposure (Southall et al., 2016). To parameterize the relationships between exposure RL and response, we need to measure the magnitude of each with quantitative metrics of uncertainty. The crux of the problem addressed here relates to RL estimation given that (1) current long-duration satellite tag sensors such as those used here lack on-board hydrophones to obtain a direct RL measurement; and (2) the three-dimensional (3D) location of the animal is not precisely known. Therefore, we turn to sound propagation models to estimate RLs. Margolina et al. (2018) used a range-dependent parabolic equation acoustic propagation model (Collins, 1993) to predict RLs for animals exposed to simulated and actual SQS-53C mid-frequency (3 to 4 kHz) active sonar (MFAS) sources used by U.S. Navy ships. Sound propagation is highly complex and location specific, but using standard assumptions about source parameters (assumed to be observed with no error), oceanography, and sediment type, Margolina et al. (2018) have predicted RLs for animals *in situ* that have been validated using calibrated sensors, with maximum levels occurring within 3 dB of modeled values during sonar CEEs in California in several dozen instances involving four different marine mammal species at ranges up to 10 km.

By using satellite tags, we give up the fine-scale behavioral data as well as measured acoustic information but gain the ability to observe and track behavioral changes over longer time frames and broader spatial scales. This enables responses to disturbance experienced by each animal during CEEs to be placed into a much longer baseline of non-exposure periods. This also provides key insight into data obtained in much finer detail for individuals tagged with high-resolution acoustic tags. A key element of the Atlantic-BRS is to use these tags in a complementary manner, playing to the strengths of each to evaluate their respective limitations. As a first step to understanding the relationship between received exposure conditions and any movement response of the animal, we developed methods to estimate RL at the animal using robust 3D sound propagation models that also account for observational error in animals' positions.

Methods

Field Data Collection
During the field effort of the Atlantic-BRS off Cape Hatteras, North Carolina, in the Fall of 2017, we attached satellite tags to 12 animals—nine on Cuvier's beaked whales and three on short-finned pilot whales (Southall et al., 2018; see Table 1). The tags were SPLASH 10 tags (Wildlife Computers, Redmond, WA, USA), with the extended-depth-range option in the LIMPET configuration (Andrews et al., 2008). Tags were set to transmit every day for 21 h for beaked whales and 17 h for pilot whales. Details on tag settings and duration of individual deployments are in Baird et al. (2018). The tags lasted on average 33.9 d for beaked whales and 30.9 d for pilot whales. The tags recorded on average 2.6 fixes per day for beaked whales and 8.8 fixes per day for pilot whales (Baird et al., 2018). In addition to shallower diving patterns, pilot whales spend considerable time at or near the surface, and they have large dorsal fins for tag attachment. Each factor contributes to a higher number of fixes per day being recorded for pilot whales.

In addition to the satellite tags, we deployed DTAGs on one pilot whale and one beaked whale during CEEs in 2017. A DTAG was deployed on one individual (Gm17_234a) in a group of pilot whales at 15:48:37 UTC on 22 August 2017, and we followed the pod for approximately 4.45 h until the tag detached and was recovered at 20:16 UTC. During this 4.45-h period, we recorded the position of the focal follow boat, bearing to the group, and an estimate of range to the group at each surfacing. Results from the beaked whale tag (Zc17_234a) are not presented here because unlike Gm17_234a, this animal was not in a group with an animal equipped with a satellite tag; see Southall et al. (2018) for details.

During the fall field effort, two CEEs were conducted with whales of both species that were tagged with satellite tags and/or DTAGs—one involving simulated MFAS and one with an operational, full-scale MFAS. The simulated MFAS CEE (#17-01) was conducted on 22 August 2017 from a stationary commercial fishing vessel (F/V *Kahuna*, hereafter *Kahuna*) with both a beaked and pilot whale tagged with DTAGs and six beaked whales and three pilot whales being monitored with satellite tags. Following a pre-exposure baseline (no noise) period, an experimental sound source (15-element vertical line array [VLA]; see Southall et al., 2012, for details) projected 1.2 s signals simulating tactical MFAS every 25 s from 18:41 to 18:55 UTC, beginning at 160 dB re 1 μPa root-mean-square (hereafter SPL) with a 3 dB/ping ramp-up to full-power transmissions held at 212 dB SPL. The ship was positioned at 35.5457 N, -74.7699 W and was not under power during the CEE but drifting at approximately 3 kts to the northeast. The operational MFAS CEE (#17-02) coordinated with the *U.S.S. McFaul* (hereafter *McFaul*; DDG-74) took place on 12 September 2017 and involved seven beaked whales and three pilot whales monitored with satellite tags (no DTAGs). A 1-h exposure was conducted from 16:03:46 to 17:02:46 UTC, with individual MFAS signals occurring every 25 s at a constant source level of 235 dB SPL. The ship was initially positioned at 36.075 N, -74.2597 W, traveling on a true bearing of 210º at 8 kts.

Development of RL Methods
The positions of tagged animals are observed with variable error—less for animals with a DTAG that are simultaneously observed by visual monitoring and measured surface positions and much more for satellite-tagged individuals based on Argos-reported positions. To account for this error, we first estimated the position of the animal in the *x,y* plane at each 5-min interval during the CEE over the course of the deployment by fitting a continuous time movement model with an embedded Ornstein-Uhlenbeck (OU) process (Johnson et al., 2008) to each individual whale's observed track. Fitting was done with the 'crawl' package in *R* (Johnson & London, 2018). We used the error ellipse data recorded with each position (McClintock et al., 2014); and, prior to fitting the OU model, we removed outlier locations using the Douglas filter algorithm (Douglas et al., 2012). Model fitting was done using projected data; for this, we used an Albers equal area projection. Using the fitted OU model, we then predicted 100 tracks at 5-min resolution (Figure 1).

By estimating model parameters and using this ensemble of 100 tracks, we accounted for the positional uncertainty in the observation process. We also used ancillary information from the depth recorder on the satellite tag in the estimation process. Specifically, if the tag indicates a depth of 1,200 m, but an estimated *x,y* location was in 300 m of water, then we could presume this estimated position is incorrect and alter the track accordingly using the fix_path() function in the 'crawl' package, which uses a least cost algorithm to adjust the tracks around unsuitable locations. However, the tags were programmed to only transmit information on the maximum depth recorded. Therefore, we chose to be conservative while evaluating estimated positions in relation to bottom topography. This provided tracks that were more realistic while still reflecting the limitations in the way the tags were programmed.

Given the known locations of the source in each CEE and the positions of animals estimated from either tag data or visual observations in the field during exposure, we estimated RLs at 5-min intervals for each CEE (RL also given in dB SPL). To estimate RLs for individuals exposed to MFAS during CEEs in a way that fully accounts for positional uncertainty, we modeled sound propagation from the known location of the source and then co-located the propagated sound with each of the 100 estimated animal positions at each 5-min interval during the CEE. Specifically, we used the fitted OU model from the 'crawl' package and predicted 100 tracks at a 5-min temporal resolution. Within each track, we selected the closest point in time to the start of the CEE and additional points for each 5-min interval for the duration of the exposure. The position of the ship with MFAS is accurately known from GPS locations, and the characteristics of the sound source are known. Using this information, we ran sound propagation models (Collins, 1993; Margolina et al., 2018) using continuous propagated acoustic energy through the entire water column. The propagation model used here (Margolina et al., 2018) is a range-dependent acoustic propagation model that allows us to estimate RL through the water column for a known source level. We chose a 10-m resolution for the depth layers, which produced a vector of RLs over 10-m bins for each of 100 positions at each 5-min interval. Because of the varied bathymetry depths, the length of the RL vectors varied in the *z* dimension—shorter vectors in shallower areas and longer vectors in deeper areas. In a given *x, y, z* bin, any measured RLs below 60 dB SPL were set to NA—that is, they were excluded from the summary statistics under the conservative assumption that this would not

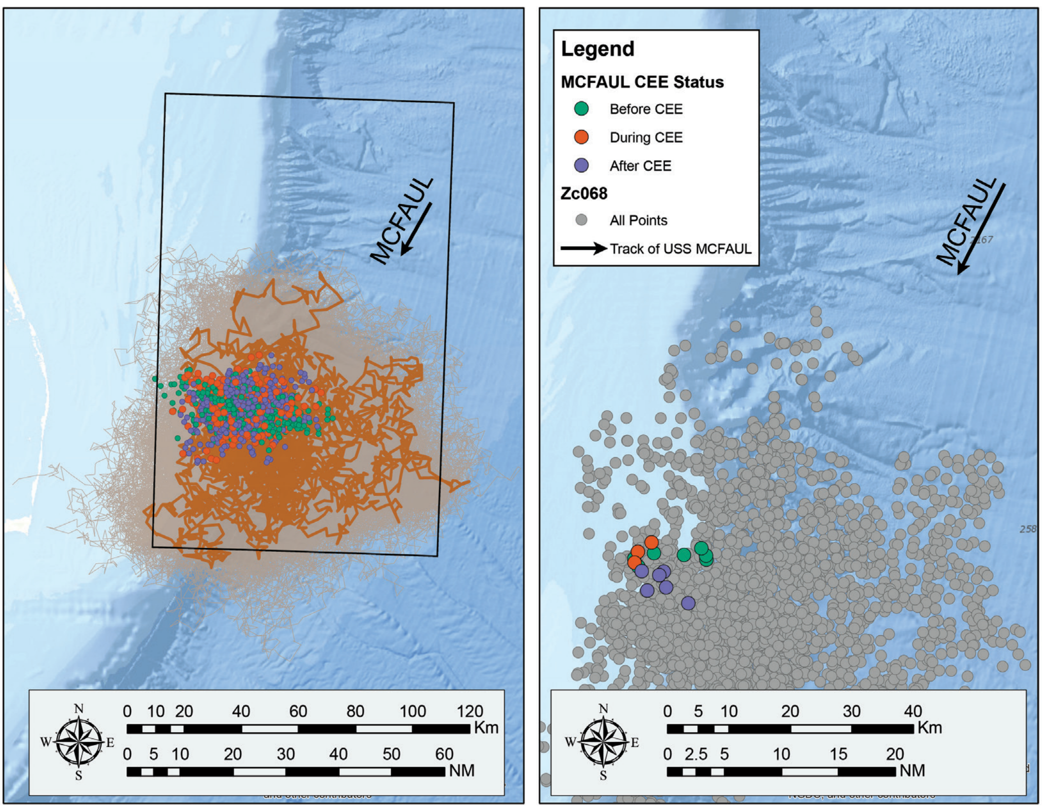

Figure 1. Movements of Zc068 in conjunction with the CEE from the *U.S.S. McFaul* (hereafter *McFaul*). Left panel shows 100 estimated tracks in light orange, with one example track highlighted in dark orange. These tracks represent the entire track; colored points correspond to imputed positions from each of 100 tracks for the hour before (green), during (orange), and after (purple) the CEE. Right panel zooms in on the area of the exposure and shows points from the highlighted track. In the right panel, the gray color indicates all the positions from one estimated track; colors of positions before, during, and after the CEE are as in the left panel.

be heard above ambient. Our analytical process builds up a distribution of RLs at each point in time—both at the surface and through the water column. We calculated summary statistics (mean and ±2 standard deviations [SD]) for the RLs for each individual, and we aggregated to wider depth bins to graphically summarize the distribution of RLs through the water column.

Evaluation of Methods Using Ancillary Data

When available, ancillary data can help narrow the positional uncertainty in a recorded movement track, which, in turn, should narrow the uncertainty around the RL. A unique situation existed for one CEE (MFAS CEE #17-01) in that a DTAG was attached to an animal (Gm17_234a) within a small (n = 6), consistently tight social group that included pilot whale Gm182 that had a satellite tag. This enabled us to directly compare the modeled RL estimates for Gm182 with measured RLs on Gm17_234a. It also enabled us to use the GPS positions from the focal follow boat to reduce positional uncertainty in the path of Gm182. (Note that to account for the possibility that the presence of the focal follow boat influences behavior, we have two levels of control. First, during a CEE, the focal follow boat is present throughout the before, during, and after periods; the constant presence of the boat should, in theory, minimize any additional change in behavior while in the during period. Second, we use experimental controls whereby we have "during" sequences where the boat is present but no sound is played. Southall et al. [2016] have shown that the use of these controls in previous BRSs result in no response to the presence of the focal follow boat.) To reduce positional uncertainty, we constructed a dead-reckoned track for Gm17_234a by processing accelerometer, magnetometer, and depth data with a Kalman filter. This path was then corrected with the observed GPS surface locations from the focal follow using the *R* package 'BayesianAnimalTracker' (Liu, 2014). We then assumed the points from the dead-reckoned track to be known (i.e., observed without error) and included them with the observed data from Gm182 prior to fitting the OU model in the 'crawl' package. We repeated the RL process described above with these estimated tracks during the scaled-source CEE from the *Kahuna*. By including a raw track and a track augmented with the user points from the dead-reckoned track, we can compare the RLs on the animal to determine how the observation error associated with the Argos system affects our estimation of RL on the animal.

Finally, as a means of evaluating sound propagation model predictions, we compared predicted RLs during CEEs with known source and receiver locations, with calibrated RLs measured on Gm17_234a as well as a calibrated, bottom-mounted passive acoustic sensor. For the whale on which DTAGs were deployed (Gm17_234a), RLs measured on the whale provide a direct measurement against which model results for satellite-tagged whales in this CEE can be compared. Note that for this comparison period, these two whales (Gm17_234a and Gm182) were consistently and repeatedly observed simultaneously at the surface within several meters of one another. We assumed that these whales were diving synchronously and, therefore, compared RLs from depth bins that coincided with observed depths on Gm17_234a. Additionally, for CEE 17-01, we used the same sound propagation modeling methods to predict RLs at the known location of a nearby (within 1.5 nmi) bottom-mounted (at 1,000 m) passive acoustic recorder (high-frequency acoustic recording package [HARP]; see Hildebrand et al., 2018) that received simulated MFAS signals during the CEE at calibrated RLs.

Results

We used robust means of characterizing animal locations within this complex and dynamic environment. Our model results demonstrate the impact of the positional uncertainty on predicted RL for one individual beaked whale (Zc068) and one individual pilot whale (Gm182) (Figures 1 & 2). Both of these animals were focal individuals for CEE 17-02 from the *McFaul*, with Gm182 also included in the second CEE with the *Kahuna*. Summary statistics for estimated exposure RL characteristics for each CEE are given in Table 1.

RL Estimation

One hundred modeled positions from the 100 imputed tracks for beaked whale Zc068's position at the start of the *McFaul* CEE were distributed broadly over the shelf, the shelf break, and into significantly deeper waters (Figure 3). Based on these positions, predicted RLs using propagation modeling are substantially variable. Estimated RLs experienced by the animal during the CEE vary depending on whether the animal was assumed to be over the shelf or in deeper waters (Figures 3 & 4; Table 1). In particular, the results in the shallowest depth bins indicate a bimodal distribution of RLs corresponding to positions on or off the shelf (Figure 4).

In contrast, the position of the pilot whale during this CEE was observed with less error (Figure 2); that is, because the depth of the maximum dive was shallower than the beaked whale, and because its estimated position was off the shelf, the distribution of RLs is more concentrated at the surface and varies less through the water

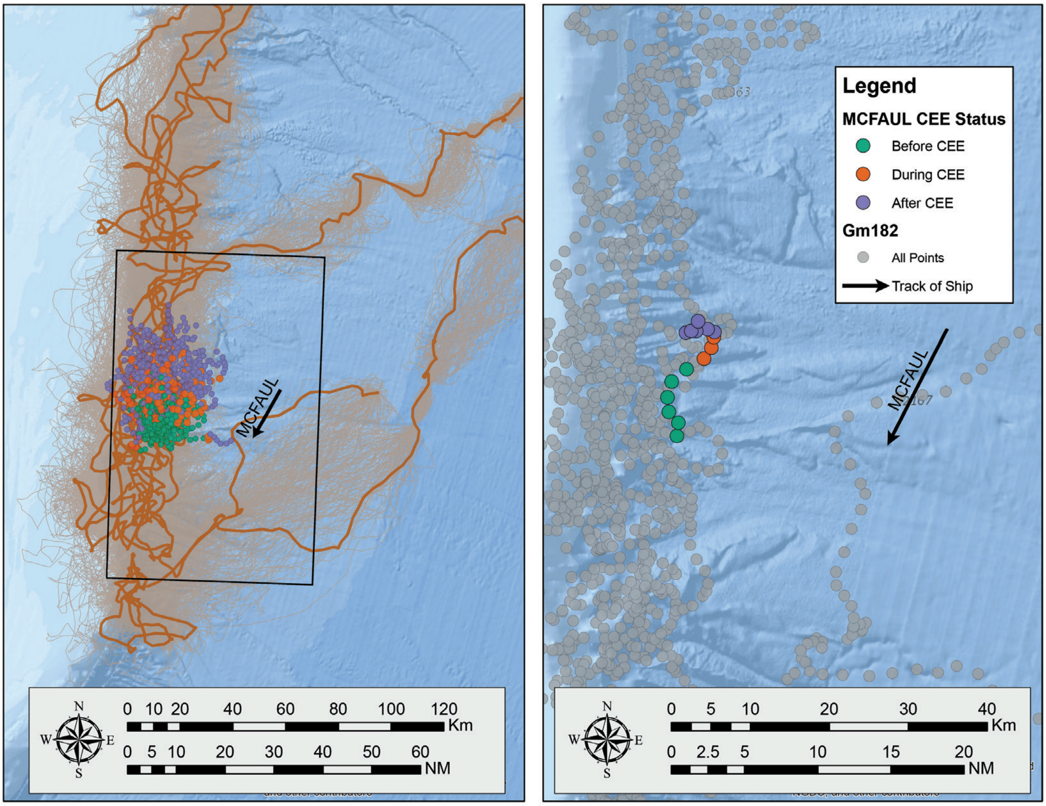

Figure 2. Movements of Gm182 in conjunction with the CEE from the *McFaul*. Left panel shows 100 imputed tracks (light orange) with an example track highlighted in dark orange. Colored dots indicate 100 estimated positions of Gm182 before (green), during (orange), and after (purple) the CEE exposure with the *McFaul*. Right panel zooms in on the region to show the position of Gm182 for one of 100 imputed tracks during the exposure. In the right panel, the gray color indicates all the positions from one estimated track; colors of positions before, during, and after the CEE are as in the left panel.

column (Figure 5). In particular, the positional uncertainty of the beaked whale in the shallowest depths means that estimates of RL are broad (from ~60 to ~110 dB SPL; Figure 4). If we knew the exact depths to which the animal was diving, the estimates around the RL could be significantly narrowed. For example, if we knew from the dive measurements that an animal was at 1,500 m, but the estimated *x,y* position was at 150 m depth, then we could assume this estimated position is implausible. Excluding these impossible positional estimates would narrow the uncertainty in the RL. As compared to the beaked whales, the pilot whales on average have more observed data, which means greater positional certainty and, thus, narrower estimates of RL (Figure 5).

Ancillary Data

The inclusion of ancillary georeferenced points significantly narrowed the uncertainty in estimates of *x,y* position for Gm182 (Figure 6), which translated to significantly narrower RL estimates during CEEs (Figure 7). In all three 5-min periods of the simulated CEE, the RL for Gm182 was both higher and less variable when the ancillary data were included (Figure 7). Further, presuming that Gm182 is diving synchronously with Gm17_234a, estimates of uncertainty around the RL in the *z* dimension narrow as well (Table 1). For example, with the ancillary data incorporated, the mean RL over 100 estimated positions was 12 dB SPL higher (133.1 vs 121.0 dB SPL), and the range was 16 dB SPL lower (119.6 to 146.6 dB SPL vs 99.4 to 142.6 dB SPL; Table 1). What this indicates is that while satellite tags offer greater duration, their positional data can be greatly enhanced with the inclusion of ancillary data. This means that the inclusion of ancillary data from the focal follow boat reduces the uncertainty that goes into assessing the exposure-response relationship at the core of this work. Lastly, it means that more

Table 1. Summary statistics from all the whales possibly exposed to sound during two CEEs in 2017—*Kahuna* and *U.S.S. McFaul*. Statistics denote the estimated distance (m) of the animal to the source as well as the median estimated received levels (RLs; dB re 1 µPa) over the course of each exposure. 95% confidence intervals (CIs) are shown for both distance and RL. Gm181 and Gm183 were very far from the source, and their estimated RLs were not above ambient. For Gm182, we indicate the RLs modeled with and without the ancillary positional information from the focal follow vessel—see text for further details. For Gm182, note the broader 95% CI when the ancillary information is not included.

Exposure	Animal	Max depth (m)	Distance to ship/source (m)	2.5%	97.5%	RL	2.5%	97.5%
McFaul	ZcTag 060	1,983	150,773.0	134,503.4	167,838.9	80.1	60.9	99.4
	ZcTag 061	800	74,003.6	66,412.6	81,663.9	104.3	88.5	120.2
	ZcTag 063	1,279	54,965.7	45,715.5	69,249.4	109.7	86.9	132.6
	ZcTag 064	1,439	106,983.5	94,750.9	120,433.1	91.1	71.1	111.1
	ZcTag 066	800	86,876.8	83,488.5	90,723.8	93.6	75.9	111.3
	ZcTag 067*	1,247	59,742.2	33,636.2	87,179.4	113.5	86.6	140.5
	ZcTag 068	1,599	79,869.5	70,560.3	89,655.2	95.6	73.0	118.1
	GmTag 181	75	491,140.8	490,239.7	492,127.6	NA	NA	NA
	GmTag 182	663	37,049.7	30,607.4	44,121.4	116.6	93.1	140.1
	GmTag 183	75	292,883.5	291,392.7	294,336.8	NA	NA	NA
Simulated MFAS source	ZcTag 060	679	22,985.4	48,477.1	73,986.4	95.5	65.9	125.1
	ZcTag 061	1,183	24,560	14,472.1	36,464.1	102.2	77.2	127.1
	ZcTag 062	1,727	16,797.6	3,358.6	36,384.1	112.5	89.9	135.1
	ZcTag 063	2,319	16,459.7	5,563.1	28,365	112.3	93.4	131.2
	ZcTag 064	800	19,386.2	7,722.2	32,784.7	112.0	86.1	137.9
	ZcTag 065	1,247	6,994	1,276.9	14,887.4	124.2	104.3	144.2
	GmTag 181	75	5,181.3	957.7	11,520.2	113.4	91.0	135.9
	GmTag 182	25	1,549.1	1,305.4	1,794.4	133.1	119.6	146.6
	GmTag 182 – No focal follow	75	5,782.9	1,425.3	11,447.3	121.0	99.4	142.6
	GmTag 183	567	35,553.1	30,519.1	40,848.6	104.0	85.7	122.4

*Animal's tag had pressure transducer issues.

information about the position of the animal in the water column will also narrow the uncertainty in RL.

Finally, we compared modeled RLs for Gm182 using the methods of Margolina et al. (2018) against measured RLs on the DTAG for Gm17_234a during the CEE of the *Kahuna* for each defined time period. This modeling accounted for the incremental escalation of source levels for the simulated MFAS during this CEE. Modeled RLs were compared with measured RLs at the first (18:41 GMT) and second (18:46 GMT) time intervals. For the final time period (18:51 GMT), the closest received ping was compared with the modeled values, but three additional pings during the 18:50 to 18:52 GMT period when the source was within ~100 m of the 18:51 GMT location were also considered to better evaluate variability in RLs over multiple pings relative to model predictions. At 18:41 GMT, the MFAS source level was 160 dB SPL, the mean modeled RL for Gm182 using ancillary data (i.e., the focal follow positions) was 85.3 dB SPL (min = 76.1 dB SPL; max = 101 dB SPL), and the measured RL on Gm17_234a was 95.6 dB SPL. At 18:46 GMT, the MFAS source level was 196 dB SPL, the mean modeled RL for Gm182 using ancillary data was

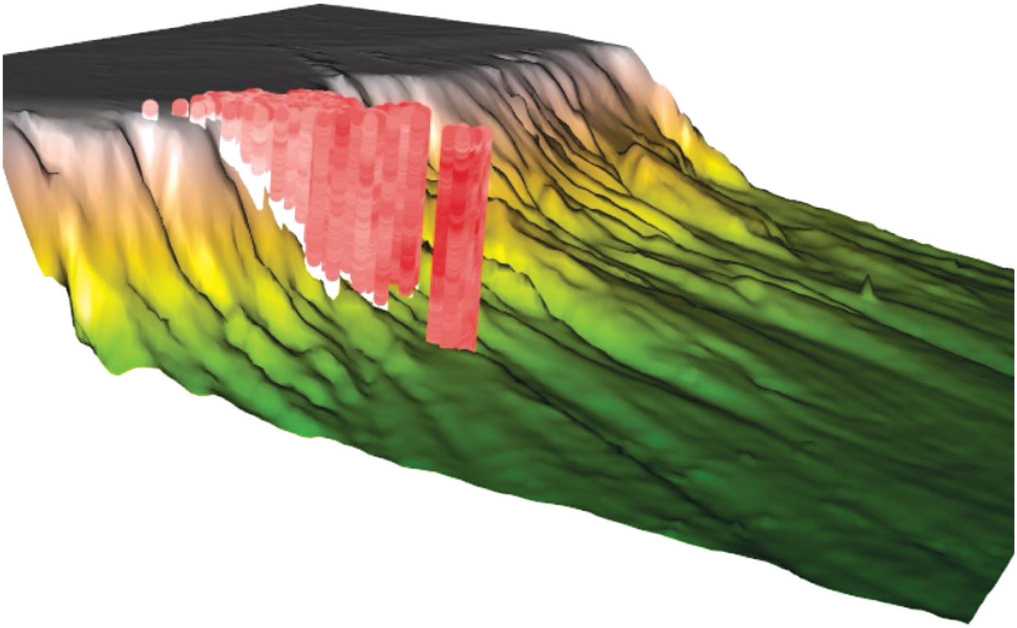

Figure 3. One hundred estimated positions in *x*, *y*, and *z* for Zc068 in relation to approximate bottom topography around the shelf break. Red colors denote received level (RL), with darker red corresponding to higher RLs. The relationship between positional uncertainty and complex bottom topography means that certain estimated locations on the shelf have many fewer depth bins through the water column than those farther offshore. RLs range from 60 dB (light pink) to 137.6 dB (dark red).

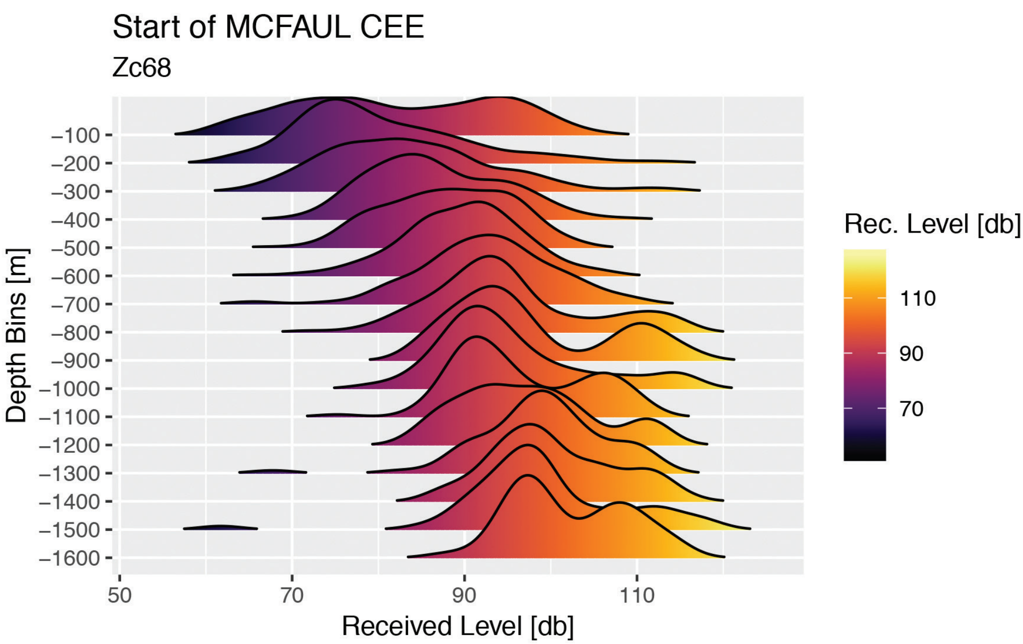

Figure 4. Estimated received level through the water column at the estimated position of Zc68 for the first 5 min of the *McFaul* CEE

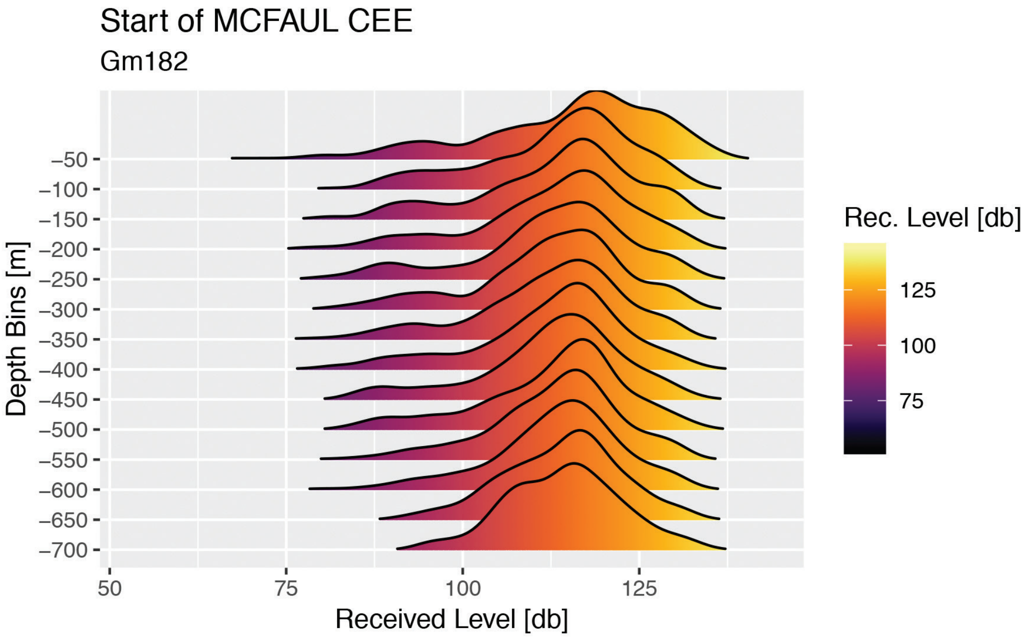

Figure 5. Estimated received level through the water column at the estimated position of Gm182 for the first 5 min of the *McFaul* CEE

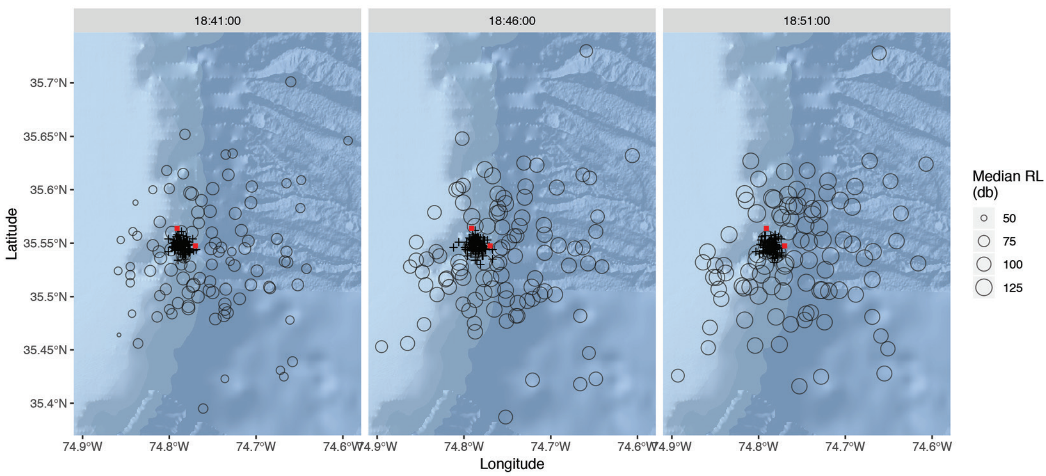

Figure 6. Positions of the *Kahuna* (red) and of Gm182 during the CEE. Open circles are 100 estimated positions and are symbol-coded according to the estimated RL. Crosses indicate 100 positions that are estimated using focal follow and DTAG data to estimate position of the animal (Gm17_234a)—see text for details.

126 dB SPL (min = 113 dB SPL; max = 160 dB SPL), and the measured RL on Gm17_234a was 133.2 dB SPL. At 18:51 GMT, the MFAS source level was 212 dB SPL, the mean modeled RL for Gm182 using ancillary data was 138 dB SPL (min = 128 dB SPL; max = 160 dB SPL), and the measured RL on Gm17_234a was 141.5 dB SPL. Measured RLs on Gm17_234a from 18:50 to 18:52 GMT ranged from 141.1 dB SPL to 149.7 dB SPL (mean = 143.8 dB SPL). Finally, for the 18:50 to 18:52 GMT period, modeled RLs for full-power MFAS signals (212 dB source level) were compared to measured RLs at the bottom-mounted HARP location (~1.5 nmi from

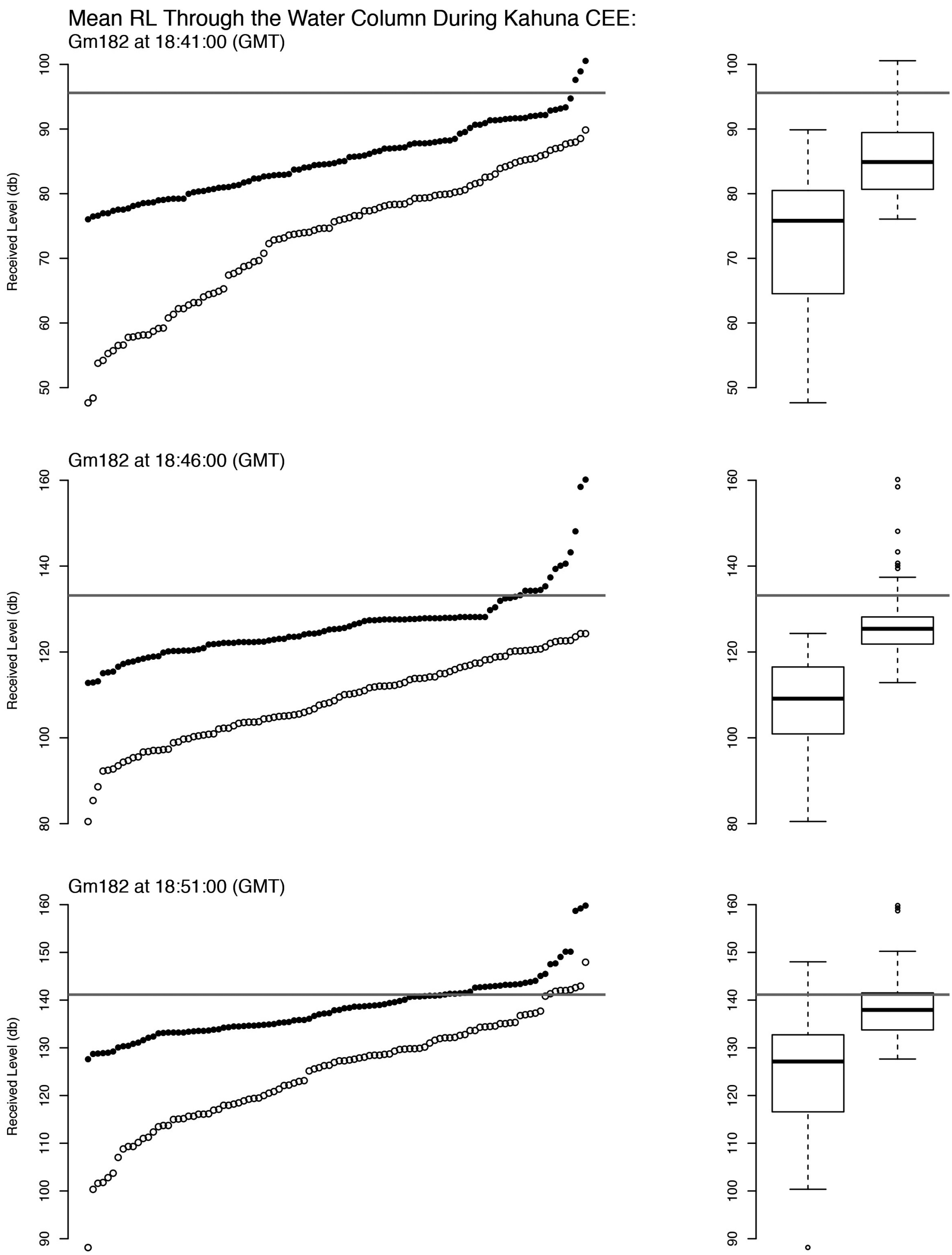

Figure 7. Comparison of RL on Gm182 with and without the "user points" from the DTAGged animal during three different 5-min intervals within a CEE. Filled black circles correspond to estimated locations that include the user points; open circles correspond to estimated locations that exclude the user points. Points are ordered from left to right by increasing RLs; box plots in right panel provide a graphical summary of these points. Horizontal blue line represents levels recorded by the DTAG on animal Gm17_234a. Note that there was a ramp-up for this CEE for which the source levels started at 160 dB and increased 3 dB/ping until it reached a maximum of 212 dB. Note also the max depth was assumed to be 25 m based on the DTAG dive record.

the source; 1,000 m depth). Predicted RLs at the HARP location for four pings during this period ranged from 114 to 117 dB SPL, while measured RLs ranged from 106 to 117 dB SPL.

Discussion

Our new analytical process demonstrates that by using modern, four-dimensional (4D) sound propagation models (Margolina et al., 2018), we can obtain reasonable estimates of RL for satellite-tagged animals during different types of CEEs (Table 1). Critically, these estimates of RL robustly characterize and fully account for positional error in the observed tracks. The direct validation of these model results to within < 3 dB of calibrated RL measurements (on animals and from passive acoustic recorders at fixed locations from multiple transmissions) demonstrates the validity of the modeling methods when position is known and their utility to characterize error associated with geospatial uncertainty. The sound propagation tools used here (Margolina et al., 2018) were originally developed to help plan CEEs by positioning experimental sources based on a current position of the animal and desired RL. For inference following the CEE, the tool was inverted and applied in a more conventional approach, propagating sound from the MFAS source to the receiver. That is, knowing the characteristics of the source, the bottom topography, and the oceanography of the system, the tool was used to estimate RLs at known locations.

By fully accounting for spatial uncertainty with modern, robust movement models, we can obtain reasonable median estimates with variance. This allows us to inform risk functions and properly propagate uncertainty into the next phase of inferential movement models that examine the relationship between observed movements, environmental variables, and sound exposure layers. For example, Hanks et al. (2015) propose a model that accounts for positional uncertainty while making inference on the variables that explain movement—both baseline and time-dependent responses to specific covariates. In terms of evaluating the sound propagation model, we have two unique *in situ* references that provide key insight. These included an animal equipped with a DTAG (Gm 17_234a) in the same group as an animal equipped with a satellite tag (Gm182) during a simulated CEE, as well as a bottom-mounted HARP at a known location. During each 5-min period (Figure 6), measured RLs on the DTAG were similar to the model estimates and within the calculated RL error, with the last period providing the closest match between recorded and estimated RLs (Figure 7). These results are consistent with earlier applications of these methods but notably occur in a more dynamic and complex oceanographic environment on the shelf break in areas with strong and complex current patterns and thermal gradients. These results confirm that this approach can generate an accurate representation of the RL on the animal as well as fair representations of variance in those estimates associated with positional uncertainty.

Behavioral response studies on marine mammals have sought to increase the duration over which animals are monitored before and after known exposure events. As such, there is a need to fully characterize aspects of exposure events that may not be directly measured (e.g., exposure RL in disturbance studies). Previous work on estimating noise exposure in instances where it is not measured directly has often been limited in the extent to which they account for the positional uncertainty inherent in satellite-transmitting tags and the resulting variability in estimated RL. Determining RLs at an animal during exposure to sound is important given the objective of deriving exposure-response probabilistic functions and the importance of RL within many regulatory assessments (Southall et al., 2007, 2019b; Hatch et al., 2016). Additionally, recent work has stressed the importance of considering RL in relation to multiple contextual factors, including behavioral state of the animal during exposure, spatial orientation of sound sources and receivers, prior exposure of the animal to sound, and the environmental context within which the exposure took place (Ellison et al., 2012; Goldbogen et al., 2013; Friedlaender et al., 2016; Southall et al., 2019a, 2019b; Wensveen et al., 2019). These factors are also critical to determining the type and magnitude of behavioral response.

With particular types of tags that accurately record sound (Johnson & Tyack, 2003; Szesciorka et al., 2016), we can measure RL directly on the animal. The use of these types of tags is one of several existing monitoring approaches—approaches that have distinct advantages and disadvantages that largely reflect the spatio-temporal scales at which they can record data on animal behavior. Long-term satellite tags enable scientists to observe behavioral responses in the movements of individuals at broader scales, but they do not provide any information on sound exposure or RL. Since RL is an important exposure variable within a regulatory framework, and since it is a necessary first step to understanding the multifaceted contextual response (e.g., Goldbogen et al., 2013), the method we describe to more accurately model animal position and sound exposure from satellite tag data offers a way to take advantage of the benefits of using these tags while still estimating and incorporating measures of RL into the analysis. Of the 19 animal-exposure events within the Atlantic-BRS effort in 2017, ten were to real MFAS sources (3 pilot whales and 7 beaked whales) and nine were to simulated MFAS sources (3 pilot whales and 6 beaked whales). Of the ten real MFAS exposures, none received an average RL higher than 117 dB SPL (Table 1), and two of the three pilot whales were far enough from the *McFaul* at the time of the CEE that model results indicate exposures would have been inaudible (Table 1).

One of our critical findings is that spatial uncertainty concerning the position of the animal at the time of the exposure can be large (*cf*. distances and credible intervals reported in Table 1; Figures 1 & 2). For example, the beaked whale closest to the *McFaul* at the start of the CEE was ZcTag067. The median estimate of distance to the ship was 59.7 km, with the min and max distances being 33.6 and 87.2 km, respectively. In turn, the median RL was 114 dB SPL, with a min and max of 87 and 141 dB SPL, a range of over 50 dB. Factoring this range of RL into subsequent exposure-response analysis provides important information about how the uncertainty in RL can influence the uncertainty in observed behavioral response. Thus, positional uncertainty leads to significant ambiguity in the RL during CEEs, which will propagate through to modeling the relationship between exposure and response. However, our ability to statistically characterize this variability in RLs at least provides an objective means of interpreting the results within the context of other exposure-response instances that a single number from a point source estimate that fails to account for geospatial uncertainty could.

CEEs are complicated and multifaceted (Southall et al., 2016); within these studies, we collect a broad variety of data about the individual animals, their movements and dive behavior, and the sound itself. Herein, we found that using ancillary data associated with focal follows significantly narrowed the uncertainty regarding the position of the animal, which, in turn, narrowed the estimates of RL (Table 1; Figures 6 & 7). Specifically, we estimated the range of RLs to be over 43 dB using just the tag, and over 27 dB with the inclusion of the ancillary data. In either case, using just one median estimate would severely underestimate the uncertainty in the RL. This is a critical, if obvious, finding because it allows us to partially address one of the disadvantages associated with using satellite tags—namely, their relatively coarse reporting cycle and relatively low positional accuracy. Unlike conducting a CEE with a DTAG where one must maintain close proximity to the tagged whale to establish a reliable georeferenced track of the animal, with a satellite tag, such restrictions are lifted. Yet, although the animal may be in the vicinity of the source, because of diving behavior or satellite configuration, locations may not be recorded and transmitted during the CEE. If, as was the case with Gm182 during the CEE with the simulated MFAS, we are able to maintain visual contact with the whale, then our positional estimates greatly improve (Figure 6). In terms of reducing uncertainty, this has a positive, cascading effect on the estimates of RL (Table 1; Figure 7). Ultimately, this helps reduce the uncertainty in the exposure-response relationship(s) that are used in activity planning and the regulatory decision-making process. While the animal in this case was a pilot whale, this process of using ancillary data may be even more important in the deep-diving beaked whales for which the reporting and richness of the positional data is lower (Quick et al., 2019). Other ancillary data, like resights of a tagged animal during photo-identification field work, can be used as well (W. Cioffi, pers. comm., 7 January 2019). Each of these positions are recorded with minimal GPS error, and their inclusion improves estimation of the true movement paths of the animal.

Results from previous BRS work have demonstrated the importance of knowing where in the water column the animal is during the exposure (Tyack et al., 2011; DeRuiter et al., 2013; Goldbogen et al., 2013). For example, Goldbogen et al. (2013) documented the varying response of blue whales to MFAS exposure as a function of water column position, with animals both in a feeding state and in deeper water more likely to exhibit a behavioral response. This importance is due in part to the complexities of sound propagation through the water; for a given *x,y* position, an animal at different depths in the water column can experience dramatically different RLs. Couple the depth-related changes of sound propagation with the positional uncertainty that accompanies satellite tags, and the range of RL experienced by an animal can be large and varied (Figure 4). In areas of highly complex bottom topography and oceanography (e.g., off Cape Hatteras, North Carolina), this can result in a very different understanding of RL as a function of the animal's position in *x*, *y*, and *z*; this was especially apparent in the beaked whales (Figures 3 & 4) as compared to pilot whales (Figure 5). In particular, the uncertainty surrounding the RL at the surface is much higher than at depth for beaked whales (Figure 4); much of this discrepancy is related to the estimated positions of the beaked whale being very close to the shelf break (Figures 1 & 3). In contrast, estimates of the pilot whale's position at the start of the CEE were in an area of less bathymetric relief (Figure 5). Pilot whales, in general, are diving to much shallower depths than beaked whales, which highlights the need to know where in the water column the animal is at the start of the CEE in areas of high bathymetric relief.

We are currently experimenting with alternate tag configurations, which afford more continuous measurements of depth than the maximum depth per dive we recorded in 2017 (W. Cioffi & N. Quick, pers. comm., 24 November 2018; Quick et al., 2019). This should narrow our uncertainty of the possible volume of RL to which the animal was exposed. Because of limitations in the tag-to-satellite bandwidth, programming the tags to report in this way typically means that the additional depth data that are recorded result in shorter temporal extent of measurement. This has implications for the experimental design—for example, if bad weather limits field work during the approximately 2-wk period of high intensity depth recording, we may end up with better depth data but no data during the CEE. Thus, field logistics need to be factored into the tag settings. In addition, more data over longer periods should, in theory, provide more opportunities for inference at the level of the individual—both in terms of its underlying behavior and its response to disturbance. Thus, optimizing the tag settings within a CEE context remains an active and important area of research, as does accounting for and minimizing the associated uncertainty on multiple levels.

Acknowledgments

Funding was provided by the U.S. Navy's Marine Species Monitoring Program. There, we thank Joel Bell and Jene Nissen. We also thank Jessica Aschetino and Dan Engelhaupt from HDR. RSS received additional funding from the Office of Naval Research (N00014-17-1-2817). We thank Andy Read, Will Cioffi, Nicola Quick, Devin Johnson, and Josh London for helpful comments and discussion throughout this process; Robin Baird, Daniel Webster, Zach Swaim, Heather Foley, and Danielle Waples for their support in tag deployments and other field logistics; Will Harmon and Mark Wilson for help in our coordination with Navy vessels; John Hildebrand and Sean Wiggins for their collaboration with HARP analysis; and Liz Henderson and two anonymous reviewers whose comments strengthened this article. Tagging was undertaken under National Marine Fisheries Service Scientific Research Permit Nos. 17086 and 20605 (issued to Robin W. Baird), and tagging protocols were approved by the Cascadia Research Collective Institutional Animal Care and Use Committee.

Literature Cited

Aguilar Soto, N., Johnson, M., Madsen, P. T., Tyack, P. L., Bocconcelli, A., & Fabrizio Borsani, J. (2006). Does intense ship noise disrupt foraging in deep-diving Cuvier's beaked whale (*Ziphius cavirostris*)? *Marine Mammal Science*, *22*(3), 690-699. https://doi.org/10.1111/j.1748-7692.2006.00044.x

Andrews, R. D., Pitman, R. L., & Ballance, L. T. (2008). Satellite tracking reveals distinct movement patterns for Type B and Type C killer whales in the southern Ross Sea, Antarctica. *Polar Biology*, *31*, 1461-1468. https://doi.org/10.1007/s00300-008-0487-z

Baird, R. W., Webster, D. L., Swaim, Z. T., Foley, H. J., Anderson, D., & Read, A. J. (2018). *Spatial use by Cuvier's beaked whales and short-finned pilot whales satellite tagged off Cape Hatteras, North Carolina: 2017 annual progress report*. Prepared for U.S. Fleet Forces Command. Submitted to Naval Facilities Engineering Command Atlantic, Norfolk, Virginia.

Bernaldo de Quirós, B. Y., Fernandez, A., Baird, R. W., Brownell, R. L., Jr., Aguilar de Soto, N., Allen, D., . . . Schorr, G. (2019). Advances in research on the impacts of anti-submarine sonar on beaked whales. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1895), 20182533. https://doi.org/10.1098/rspb.2018.2533

Claridge, D. E. (2013). *Population ecology of Blainville's beaked whales (*Mesoplodon densirostris*)*. St. Andrews, UK: The University of St Andrews. Retrieved from http://research-repository.st-andrews.ac.uk/handle/10023/3741

Collins, M. D. (1993). A split-step Padé solution for the parabolic equation method. *The Journal of the Acoustical Society of America*, *93*, 1736. https://doi.org/10.1121/1.406739

Cox, T. M., Ragen, T. J., Read, A. J., Vos, E., Baird, R. W., Balcomb, K. C., . . . Benner, L. (2006). Understanding the impacts of anthropogenic sound on beaked whales. *Journal of Cetacean Research and Management*, *7*(3), 177-187. https://doi.org/10.1109/LPT.2009.2020494

DeRuiter, S. L., Southall, B. L., Calambokidis, J., Zimmer, W. M., Sadykova, D., Falcone, E. A., . . . Tyack, P. L. (2013). First direct measurements of behavioural responses by Cuvier's beaked whales to mid-frequency active sonar. *Biology Letters*, *9*(4), 20130223. https://doi.org/10.1098/rsbl.2013.0223

Douglas, D. C., Weinzierl, R., Davidson, S. C., Kays, R., Wikelski, M., & Bohrer, G. (2012). Moderating Argos location errors in animal tracking data. *Methods in Ecology and Evolution*, *3*(6), 999-1007. https://doi.org/10.1111/j.2041-210X.2012.00245.x

Ellison, W. T., Southall, B. L., Clark, C. W., & Frankel, A. S. (2012). A new context-based approach to assess marine mammal behavioral responses to anthropogenic sounds. *Conservation Biology*, *26*(1), 21-28. https://doi.org/10.1111/j.1523-1739.2011.01803.x

Falcone, E. A., Schorr, G. S., Watwood, S. L., DeRuiter, S. L., Zerbini, A. N., Andrews, R. D., . . . Moretti, D. J. (2017). Diving behaviour of Cuvier's beaked whales exposed to two types of military sonar. *Royal Society Open Science*, *4*(8). https://doi.org/10.1098/rsos.170629

Filadelfo, R., Mintz, J., Michlovich, E., D'Amico, A., Tyack, P. L., & Ketten, D. R. (2009). Correlating military sonar use with beaked whale mass strandings: What do the historical data show? *Aquatic Mammals*, *35*(4), 435-444. https://doi.org/10.1578/AM.35.4.2009.435

Friedlaender, A. S., Hazen, E. L., Goldbogen, J. A., Stimpert, A. K., Calambokidis, J., & Southall, B. L. (2016). Prey-mediated behavioral responses of feeding blue whales in controlled sound exposure experiments. *Ecological Applications*, *26*(4), 1075-1085. https://doi.org/10.1002/15-0783

Goldbogen, J. A., Southall, B. L., DeRuiter, S. L., Calambokidis, J., Friedlaender, A. S., Hazen, E. L., . . . Tyack, P. L. (2013). Blue whales respond to simulated mid-frequency military sonar. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1765). https://doi.org/10.1098/rspb.2013.0657

Hanks, E. M., Hooten, M. B., & Alldredge, M. W. (2015). Continuous-time discrete-space models for animal movement. *The Annals of Applied Statistics*, *9*, 145-165.

Harris, C. M., Thomas, L., Falcone, E. A., Hildebrand, J., Houser, D., Kvadsheim, P. H., . . . Janik, V. M. (2018). Marine mammals and sonar: Dose-response studies, the risk-disturbance hypothesis and the role of exposure context. *Journal of Applied Ecology*, *55*(1), 396-404. https://doi.org/10.1111/1365-2664.12955

Hatch, L. T., Wahle, C. M., Gedamke, J., Harrison, J., Laws, B., Moore, S. E., . . . Van Parijs, S. M. (2016). Can you hear me here? Managing acoustic habitat in U.S. waters. *Endangered Species Research*, *30*, 171-186. https://doi.org/10.3354/esr00722

Hildebrand, J. A., Frasier, K. E., Trickey, J. S., Thayre, B. J., Rafter, M., Wiggins, S. M., & Baumann-Pickering, S. (2018). *Marine mammal passive acoustic monitoring for the Atlantic Fleet using high-frequency acoustic recording packages: 2017 annual report*. Prepared for U.S. Fleet Forces Command. Submitted to Naval Facilities Engineering Command Atlantic, Norfolk, Virginia.

Hooten, M. B., Johnson, D. S., McClintock, B. T., & Morales, J. M. (2017). *Animal movement: Statistical models for telemetry data*. Boca Raton, FL: CRC Press. https://doi.org/10.1201/9781315117744

Johnson, D. S., & London, J. M. (2018). crawl: An *R* package for fitting continuous-time correlated random walk models to animal movement data. *Zenodo* website. https://doi.org/10.5281/zenodo.596464

Johnson, D. S., London, J. M., Lea, M-A., & Durban, J. W. (2008). Continuous-time correlated random walk model for animal telemetry data. *Ecology*, *89*, 1208-1215. https://doi.org/10.1890/07-1032.1

Johnson, M. P., & Tyack, P. L. (2003). A digital acoustic recording tag for measuring the response of wild marine mammals to sound. *Journal of Oceanic Engineering*, *28*(1), 3-12. https://doi.org/10.1109/JOE.2002.808212

Liu, Y. (2014). BayesianAnimalTracker: Bayesian melding of GPS and DR path for animal tracking. R *package, Version 1.2*. Retrieved from https://cran.r-project.org/package=BayesianAnimalTracker

Margolina, T., Joseph, J. E., & Southall, B. L. (2018). BRS sound exposure modeling tool: A system for planning, visualization and analysis. *OCEANS 2018 MTS/IEEE*, 1-4. https://doi.org/10.1109/OCEANS.2018.8604911

McClintock, B. T., London, J. M., Cameron, M. F., & Boveng, P. L. (2014). Modeling animal movement using the Argos satellite telemetry location error ellipse. *Methods in Ecology and Evolution*, *6*, 266-277. https://doi.org/10.1111/2041-210X.12311

Miller, P. J. O., Antunes, R. N., Wensveen, P. J., Samarra, F. I. P., Catarina Alves, A., Tyack, P. L., . . . Thomas, L. (2014). Dose-response relationships for the onset of avoidance of sonar by free-ranging killer whales. *The Journal of the Acoustical Society of America*, *135*(2), 975. https://doi.org/10.1121/1.4861346

National Academy of Sciences, Engineering, and Medicine. (2017). *Approaches to understanding the cumulative effects of stressors on marine mammals*. Washington, DC: National Academies Press.

National Research Council (NRC). (2005). *Marine mammal populations and ocean noise: Determining when noise causes biologically significant effects*. Washington, DC: National Academies Press.

New, L. F., Moretti, D. J., Hooker, S. K., Costa, D. P., & Simmons, S. E. (2013). Using energetic models to investigate the survival and reproduction of beaked whales (family Ziphiidae). *PLOS ONE*, *8*(7), e68725. https://doi.org/10.1371/journal.pone.0068725

Quick, N. J., Cioffi, W. R., Shearer, J. M., & Read, A. J. (2019). Mind the gap – Optimising satellite tag settings for time series analysis of foraging dives in Cuvier's beaked whales. *Animal Biotelemetry*, *7*(5). https://doi.org/10.1186/s40317-019-0167-5

Shearer, J. M., Quick, N. J., Cioffi, W. R., Baird, R. W., Webster, D. L., Foley, H. J., . . . Read, A. J. (2019). Diving behaviour of Cuvier's beaked whales (*Ziphius cavirostris*) off Cape Hatteras, North Carolina. *Royal Society Open Science*, *6*(2), 181728. https://doi.org/10.1098/rsos.181728

Southall, B. L., Nowacek, D. P., Miller, P. J., & Tyack, P. L. (2016). Experimental field studies to measure behavioral responses of cetaceans to sonar. *Endangered Species Research*, *31*, 293-315. https://doi.org/10.3354/esr00764

Southall, B. L., Moretti, D., Abraham, B., Calambokidis, J., DeRuiter, S. L., & Tyack, P. L. (2012). Marine mammal behavioral response studies in southern California: Advances in technology and experimental methods. *Marine Technology Society Journal*, *46*(4), 48-59. https://doi.org/10.4031/MTSJ.46.4.1

Southall, B. L., Baird, R. W., Bowers, M., Cioffi, W., Harris, C., Joseph, J., . . . Webster, D. L. (2018). *Atlantic Behavioral Response Study (BRS) – 2017 Annual progress report*. Prepared for U.S. Fleet Forces Command. Submitted to Naval Facilities Engineering Command Atlantic, Norfolk,

Virginia, under Contract No. N62470-15-D-8006, Task Order 50 (issued to HD).
Southall, B. L., Bowles, A. E., Ellison, W. T., Finneran, J. J., Gentry, R. L., Greene, C. R., Jr., . . . Nachtigall, P. E. (2007). Overview. *Aquatic Mammals*, *33*(4), 411-414. https://doi.org/10.1578/AM.33.4.2007.411
Southall, B. L., DeRuiter, S. L., Friedlaender, A., Stimpert, A. K., Goldbogen, J. A., Hazen, E., . . . Calambokidis, J. (2019a). Behavioral responses of individual blue whales (*Balaenoptera musculus*) to mid-frequency military sonar. *Journal of Experimental Biology*, *222*(5). https://doi.org/10.1242/jeb.190637
Southall, B. L., Finneran, J. J., Reichmuth, C., Nachtigall, P. E., Ketten, D. R., Bowles, A. E., . . . Tyack, P. L. (2019b). Marine mammal noise exposure criteria: Updated scientific recommendations for residual hearing effects. *Aquatic Mammals*, *45*(2), 125-232. https://doi.org/10.1578/AM.45.2.2019.125
Szesciorka, A. R., Calambokidis, J., & Harvey, J. T. (2016). Testing tag attachments to increase the attachment duration of archival tags on baleen whales. *Animal Biotelemetry*. https://doi.org/10.1186/s40317-016-0110-y
Tyack, P. L., Zimmer, W. M., Moretti, D., Southall, B. L., Claridge, D. E., Durban, J. W., . . . Boyd, I. L. (2011). Beaked whales respond to simulated and actual navy sonar. *PLOS ONE*, *6*(3), e17009. https://doi.org/10.1371/journal.pone.0017009
U.S. Navy. (2018). *Environmental impact statement/Overseas environmental impact statement Atlantic Fleet training and testing*. Norfolk, VA: U.S. Navy.
von Benda-Beckmann, A. M., Wensveen, P. J., Prior, M., Ainslie, M. A., Hansen, R. R., Isojunno, S., . . . Miller, P. J. O. (2019). Predicting acoustic dose associated with marine mammal behavioural responses to sound as detected with fixed acoustic recorders and satellite tags. *The Journal of the Acoustical Society of America*, *145*(1401). https://doi.org/10.1121/1.5093543
Wensveen, P. J., Isojunno, S., Hansen, R. R., von Benda-Beckmann, A. M., Kleivane, L., van IJsselmuide, S., . . . Miller, P. J. O. (2019). Northern bottlenose whales in a pristine environment respond strongly to close and distant navy sonar signals. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1899), 20182592. https://doi.org/10.1098/rspb.2018.2592
Yack, T. M., Barlow, J., Calambokidis, J., Southall, B., & Coates, S. (2013). Passive acoustic monitoring using a towed hydrophone array results in identification of a previously unknown beaked whale habitat. *The Journal of the Acoustical Society of America*, *134*(3), 2589-2595. https://doi.org/10.1121/1.4816585