Received: 4 October 2024 | Revised: 22 August 2025 | Accepted: 17 September 2025

DOI: 10.1002/ecs2.70501

ECOSPHERE
AN ESA OPEN ACCESS JOURNAL

ARTICLE

Coastal and Marine Ecology

Behavioral responses of goose-beaked whales (*Ziphius cavirostris*) to simulated military sonar

B. L. Southall[1,2] | **R. S. Schick**[1,2] | **W. R. Cioffi**[1,2] | **S. L. DeRuiter**[3] | **H. J. Foley**[2,4] | **C. M. Harris**[5] | **A. E. Harshbarger**[2] | **J. E. Joseph**[6] | **T. Margolina**[6] | **D. P. Nowacek**[2,7] | **N. J. Quick**[2,8] | **Z. T. Swaim**[2] | **L. Thomas**[5] | **D. M. Waples**[2] | **D. L. Webster**[9] | **J. H. Wisse**[2] | **A. J. Read**[2]

[1]Southall Environmental Associates, Inc., Aptos, California, USA

[2]Duke University Marine Laboratory, Nicholas School of the Environment, Beaufort, North Carolina, USA

[3]Department of Mathematics and Statistics, Calvin University, Grand Rapids, Michigan, USA

[4]Northeast Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Woods Hole, Massachusetts, USA

[5]Centre for Research into Ecological and Environmental Modelling, University of St Andrews, St Andrews, UK

[6]Oceanography Department, Naval Postgraduate School, Monterey, California, USA

[7]Pratt School of Engineering, Beaufort, North Carolina, USA

[8]School of Biological and Marine Sciences, University of Plymouth, Plymouth, UK

[9]Bridger Consulting, Bridger Consulting Group, Bozeman, Montana, USA

Correspondence
B. L. Southall
Email: brandon.southall@sea-inc.net

Funding information
US Fleet Forces Command Marine Species Monitoring Program; Naval Facilities Engineering Command Southeast, Grant/Award Numbers: N62470-10-D-3011, N62470-15-D-8006

Handling Editor: Hunter S. Lenihan

Abstract

We report direct measurements of changes in diving and movement behavior for 53 goose-beaked whales (*Ziphius cavirostris*) in relation to experimentally controlled mid-frequency (3–4 kHz) active sonar (MFAS) signals. These signals simulate powerful Navy sources that have been associated with multiple mortal stranding events for this species. We deployed a multi-scale combination of tags to monitor individual whales, including 50 long-duration (weeks), coarse-resolution satellite-transmitting tags and 3 short-duration (hours), high-resolution archival depth, orientation, and acoustic tags. We evaluated behavioral responses during 13 experimental trials (9 MFAS; 4 no-MFAS controls), resulting in 72 exposure events; some individuals were exposed in multiple trials. Whales were exposed at known and modeled horizontal ranges from ~2 to >200 km and from below ambient noise levels to received levels (RLs) up to ~142 dB re: 1μPa (root-mean-square [RMS]). We investigated changes in diving and movement behavior separately, with a suite of metrics, descriptive evaluations, and statistical tests. We observed similar patterns and probabilities of behavioral changes for control trials and the lowest RL conditions (<100 dB). Above 100 dB RLs, increasingly prevalent and consistent responses occurred, including extended deep dives, prolonged periods between deep dives, directed spatial movement away from the source, and cessation of echolocation. Aspects of these cryptic responses typically persisted for hours following exposure but did not result in broad-scale habitat abandonment. Our study builds upon experimental and observational studies conducted on sonar testing ranges and expands our understanding of the response of this species to MFAS in a region where operational sonar use occurs far less commonly than on Navy testing ranges. These data are directly applicable in the conservation and effective management of this sensitive, protected species.

KEYWORDS
behavior, controlled exposure experiment, Navy sonar, response, *Ziphius*

INTRODUCTION

Under some circumstances, exposure to tactical mid-frequency (1–10 kHz) sonar signals used in antisubmarine warfare has resulted in the stranding, and subsequent mortality, of deep-diving beaked whales (Family: *Ziphiidae*) around the world (e.g., D'Amico et al., 2009; Evans & England, 2001; Filadelfo et al., 2009; Miller et al., 2022). Most of these events have included goose-beaked whales (*Ziphius cavirostris*; hereafter *Ziphius*). The number of documented stranding events is small (on the order of a few dozen) relative to global sonar use, although this may not reflect the true level of mortality (e.g., Faerber & Baird, 2010). Further, the degree to which military mid-frequency active sonar (MFAS) sources may have sublethal effects on beaked whales remains poorly understood.

Beaked whales have been identified as being particularly sensitive to various human disturbances for decades (Southall et al., 2007), with responses documented to nonmilitary sounds as well, including vessel noise (Aguilar Soto et al., 2006; Pirotta et al., 2012) and scientific echosounders (Cholewiak et al., 2017). Given these observations, uncertainties, and associated attention, debate, and litigation, various international research and monitoring studies have been conducted (e.g., Bernaldo de Quirós et al., 2019; Falcone et al., 2017; Foley et al., 2021). Questions relating more broadly to the probability and magnitude of behavioral, physiological, and auditory impacts of various human noise sources on marine mammals have been the subject of intense research and management attention for several decades (Duarte et al., 2021; NAS, 2016; NRC, 2005; Southall, 2017; Southall et al., 2021; Southall, DeRuiter, et al., 2019; Southall, Finneran, et al., 2019).

Increasing evidence suggests that unusual beaked whale stranding events are mediated by their behavioral responses to active sonar (e.g., Southall et al., 2007, 2021) and that these responses are consistent with those observed in response to other potential threats, including predators (see Harris et al., 2018; Miller et al., 2022). Various behavioral response studies (BRS) have investigated the response of beaked whales to sonar, including observational studies in which animals are monitored using various means during ongoing MFAS operations (e.g., Falcone et al., 2017; Joyce et al., 2020; Moretti et al., 2014). Given the highly contextual nature of behavioral responses to acoustic stimuli in marine mammals (Ellison et al., 2012, 2018), experimental approaches established in classic exposure-response and ethological studies (see McGregor, 2013) have been adapted to quantify relationships between specific exposure conditions and response type and probability. An experimental form of BRS is a controlled exposure experiment (CEE), in which noise exposure is systematically manipulated in a consistent manner.

Here we apply this approach to *Ziphius* off Cape Hatteras, NC, USA, building on a series of prior CEEs that presented MFAS to various marine mammal species around the world (DeRuiter et al., 2013, 2017; Goldbogen et al., 2013; Isojunno et al., 2016; Kvadsheim et al., 2017; Miller et al., 2012, 2015; Southall et al., 2012; Southall et al., 2023; Southall, DeRuiter, et al., 2019; Stimpert et al., 2014; Tyack et al., 2011; Wensveen et al., 2019). Together, these studies developed a variety of sophisticated field and analytical tools to evaluate the relationship between exposure and behavioral response (see Harris et al., 2016; Southall et al., 2016). Despite their methodological differences, the combined overall results of these experimental studies, conducted with more than a dozen cetacean species, including four species of beaked whales, suggest that beaked whales respond to lower received levels (RLs) of these signals than most other cetaceans (see Southall et al., 2007, 2016, 2021). Responses documented to date have included cessation of echolocation and foraging behavior, changes to diving behavior, and horizontal avoidance from sound sources, which may persist for hours to days following the cessation of MFAS signals (e.g., DeRuiter et al., 2013; Falcone et al., 2017; Joyce et al., 2020; Moretti et al., 2014; Tyack et al., 2011). It is important to note that the pioneering beaked whale CEEs of DeRuiter et al. (2013), Stimpert et al. (2014), and Tyack et al. (2011) occurred on Navy ranges, in areas of intense use of active military sources. In contrast, Miller et al. (2015) and Wensveen et al. (2019) studied beaked whale responses in remote locations where MFAS rarely, if ever, occurs.

The generalized sensitivity of beaked whales to acoustic disturbance was first noted in the development of marine mammal noise exposure criteria (Southall et al., 2007), but the nature of beaked whale responses to MFAS in the context of antipredator behavior was first identified by Tyack et al. (2011), who experimentally measured responses to both MFAS and killer whale (*Orcinus orca*) signals. This work, together with the observational and experimental studies discussed above, has been framed in relation to the risk-disturbance hypothesis (Harris et al., 2018; Miller et al., 2022). This hypothesis postulates that animals perceive and respond to various forms of human disturbance in the same general manner as they respond to predation risk (Frid & Dill, 2002). Killer whales are an important predator of beaked whales (Wellard et al., 2016) and typically hunt near the surface and detect marine mammal prey visually or by passive listening. Responses observed in some *Ziphius* exposed to experimental (DeRuiter et al., 2013) and incidental (Falcone et al., 2017) MFAS exposures—cessation of vocal behavior, changes to diving behavior, and movement away from the sound source—are consistent with an antipredator response that invokes acoustic crypsis and flight away from a perceived risk (see Harris et al., 2018; Miller et al., 2022).

These prior studies, however, have been limited by either small sample sizes or a lack of experimental controls, and were conducted either on Navy ranges with intense MFAS use or in areas where MFAS is not typically used. Here, we employ a similar experimental approach with *Ziphius* in an area where MFAS does occur, but at much lower levels than on dedicated sonar training ranges (see Van Parijs et al., 2024). The continental shelf break off Cape Hatteras, NC, USA, supports a high density of beaked whales, due to a unique interaction of geographical and oceanographic features (McLellan et al., 2018). We built on previous work which investigated the distribution and behavior of this species in this study area (Cioffi et al., 2021; Foley et al., 2021; McLellan et al., 2018; Quick et al., 2020; Shearer et al., 2019; Stanistreet et al., 2017) to tune our experimental approach. Given the increasing interest in evaluating the fitness consequences of nonlethal exposure to individuals and placing these in the context of populations (e.g., Pirotta, 2022), we aimed to better understand patterns and variations in response by expanding the sample size of whales exposed to MFAS and evaluating responses over longer periods than previous CEEs. We developed sound propagation modeling methods to estimate MFAS RLs (Schick et al., 2019, 2024) and employed several types of biologging tags to quantify behavior on multiple scales (as in Cioffi et al., 2023; Quick et al., 2020; Wensveen et al., 2019).

The current study is part of a larger research project that includes studying responses to full-scale, tactical SONAR (SQS-53C MFAS systems) used on US naval surface vessels. Here we present and evaluate responses to exposure trials conducted using lower power signals designed to simulate spectral and temporal features of operational MFAS and produced by purpose-built scientific transducers (see DeRuiter et al., 2013; Southall et al., 2012; Southall, DeRuiter, et al., 2019). We seek to address whether *Ziphius* in our study area is more or less sensitive to MFAS signals in terms of diving and movement behavior as in other areas, including Navy ranges, and to consider these responses within the context of the risk-disturbance hypothesis described by Harris et al. (2018) and Miller et al. (2022). The multi-scale design allowed us to extend the spatial and temporal extent at which we addressed these questions, and to provide some evaluation of the time course of potential responses. Specifically, we investigated whether MFAS exposure induced *Ziphius* to exhibit: (1) *changes to diving behavior*, including the duration of long dives and cessation of echolocation behavior; (2) *changes to surfacing behavior*, including increased surface intervals and the duration of shallow dives; and (3) *spatial avoidance*. Our overall objectives were to identify the occurrence and magnitude of responses associated with MFAS exposures, inform our understanding of the potential consequences of such responses, and describe contextual aspects of response type and probability with which to derive exposure-response functions for future use in regulatory applications.

METHODS

To describe the nature of these complex experiments, we present the methods in two parts. The first section describes the experimental methodology and the second outlines the suite of analytical methods used to evaluate changes in diving behavior, horizontal avoidance, and social responses. An expansive supplementary materials document (Appendix S1) provides additional methodological details.

Overall experimental design and approach

Field site, vessel configuration, and vessel operations

Our study was conducted between May and September of 2017–2022 over the outer continental slope waters off Cape Hatteras in an area where previous studies of

Ziphius have been conducted (see fig. 1 in Foley et al., 2021; Shearer et al., 2019). We typically employed an 8.9-m aluminum hull vessel (R/V *Richard T. Barber*) to locate groups of *Ziphius* and deploy satellite-transmitting tags (hereafter "satellite tags"; described below) days to weeks prior to exposure trials. On the days in which exposure trials were conducted, field teams relocated potential "focal" individuals or groups using remote satellite transmissions via Service Argos or by directly intercepting transmissions with vessel-based Argos Goniometers (Woods Hole Group, Bourne, MA, USA). In most cases, once a focal individual or group was relocated, it was tracked using conventional focal follow methods in which surface observations were made at ranges of several hundred meters until animals dove; during these focal follows, we avoided close approaches (as in DeRuiter et al., 2013; Miller et al., 2015; Stimpert et al., 2014; Tyack et al., 2011). These focal follow observations provided verified surface locations to track fine-scale movement, provided information on group size and social composition, allowed images to be collected for later photo identification, and provided locational data to support real-time modeling of MFAS signals to position sound sources. Remote vessel-based focal follows were maintained consistently before, during, and after exposure trials that included MFAS as well as in no-MFAS "control" experiments.

When field conditions allowed (less than half of the total experimental trials), additional tagging effort was attempted to deploy short-term (in hours) archival tags (digital acoustic tags; Johnson & Tyack, 2003, hereafter "DTags"), which provide broadband acoustic recordings and detailed data on kinematics and diving behavior. The goal was to deploy a DTag on an additional individual either within a focal group with a satellite tag or a separate group or individual. If this was successful the DTag individual or group, whether it also had a satellite tag individual or not, became the focal group. In either case, the focal group was remotely observed, and no close approaches were made for a period of at least 2 h prior to the beginning of exposure trials. In most trials, there were multiple non-focal tagged individuals (up to 10) in the overall area, all of which were monitored using satellite tag transmissions; those within visual range were visually monitored as possible.

Custom sound propagation modeling tools (Margolina et al., 2018) were used in situ based on real-time locations of tagged animals to evaluate received MFAS conditions for all focal individuals/groups and as many instrumented non-focal individuals as possible. The goal of this modeling was (1) to target specific RLs for focal individuals/groups based on experimental objectives and (2) to ensure that no non-focal tagged animals RLs above the maximum exposure conditions allowed in our permits and protocols. Schick et al. (2019) compared RLs of MFAS signals measured with calibrated recorders to modeled values from early field trials, demonstrating the efficacy of this approach (see also Schick et al., 2024).

A second vessel served as a platform for the simulated MFAS sound source. In 2017–2019, we used ~15–20-m charter fishing vessels as the sound source platform, but starting in 2020, we employed the R/V *Shearwater*, a 23.2-m research vessel operated by the Duke University Marine Laboratory. While drifting and not under power, the experimental source was lowered and either activated during MFAS exposures or not activated in no-MFAS control trials. As well as being the sound source platform, this vessel served as an observation platform to ensure trials were conducted according to authorized protocols. It also provided visual observations to support the focal follow vessel, but at ranges of hundreds of meters or more from tagged individuals and without any deliberate close approaches. As with the focal follow vessel, the source vessel was operated identically during all trials, including MFAS exposures and no-MFAS controls, so the potential effects of these vessels on the whales' behavior would be as consistent as possible.

Multi-scale tagging approach

We used extensive baseline data on the diving and movement behavior of *Ziphius* in our study area (Cioffi et al., 2021; Foley et al., 2021; Shearer et al., 2019) to inform the configuration of tag types and optimize settings (Cioffi et al., 2023). Dive patterns for *Ziphius* in this area include deep, long, and presumed foraging dives (>800 m; >33 min; Shearer et al., 2019). Shorter, shallower, presumed non-foraging dives (<800 m; <33 min) occur between deep dives; these have been referred to as "bounce" dives in some previous studies, but we use the term "shallower" hereafter. Periods between deep dives are referred to as inter-deep-dive intervals (IDDIs). We characterized deep dives, shallower dives, and IDDIs using both satellite tags and DTags.

We deployed satellite-linked SPLASH10-292, SPLASH10-F-333, or SPLASH10-333 depth recorders with the extended depth option and location-transmitting tags (Wildlife Computers, Redmond, WA, USA) in the LIMPET configuration (Andrews et al., 2008). Each satellite tag (not more than one per individual) was programmed to collect surface positions (with reported error) and either dive summary records (with categorical summaries of dive data) or a depth time-series sampled over one of several possible selected time windows (see Cioffi et al., 2023). Data from these tags were obtained through satellite-uplink Argos transmissions and augmented in the field with Argos goniometers. Goniometers were used in the field to locate animals in real time and to augment movement modeling and enhance modeled RL estimates during analyses (see Schick et al., 2024).

We only included data from satellite-linked tags in our analysis if at least 24-h of baseline data was obtained prior to an exposure trial. In cases where multiple MFAS exposure trials occurred in a tag record, we excluded 24-h of data following any prior exposure from the baseline. The time course of responses to MFAS in beaked whales is not fully understood, but we selected a 24-h period because some earlier experiments suggest that acute responses begin to abate after several hours (DeRuiter et al., 2013; Wensveen et al., 2019).

We deployed version 3 DTags using similar approaches, sampling regimes (acoustics sampled at 240 kHz; accelerometers, magnetometers, and pressure sensors sampled at 250 Hz), settings, calibrations, and data processing as in DeRuiter et al., 2013; Southall et al., 2023; Southall, DeRuiter, et al., 2019. We monitored whales equipped with DTags through at least one deep dive (>33 min) and subsequent surface series prior to commencing exposure trials. A total of three DTags were deployed and included in exposures described here, one of which was deployed in a group with a satellite-tagged whale. These tags were all deployed more than 2 h prior to exposure trials and collected data for several hours following the end of exposure.

Satellite tag data movement and RL modeling

To account for the uncertainty in position estimates associated with the satellite tags, we implemented the following multistep procedure: First, where possible, we augmented the satellite-uplinked data with additional data obtained in the field during focal follows of individual whales. The augmented records included deployment location, visual observations of the tagged whale, and data on relative bearing and range to whales from the goniometer (see Schick et al., 2024). The positional estimates from Service Argos included uncertainty from which error ellipses were generated for all surface locations. Where possible, error ellipses were generated for augmented positions based on goniometer data using a predictive relationship from empirical field measurements developed by Borroni (2024).

As described in greater detail by Schick et al. (2024) and in Appendix S1: Section S1: Satellite tag data spatial modeling, horizontal range, and received level calculations, we fit a continuous-time correlated random walk model to the filtered and augmented Argos positions. We used the fitted model to impute 100 possible tracks to generate ensemble estimates of location that incorporated positional error. Within each of the 100 imputed tracks, we predicted locations at the start of each exposure trial relative to the known location of the sound source to determine the median and 95% CI for the horizontal distance of each whale and exposure trial.

We also predicted imputed locations for all tracks at 5-min intervals within each trial to conduct RL modeling (Margolina et al., 2018; Schick et al., 2019, 2024). We used sound propagation model output to estimate MFAS RLs at multiple imputed points, with derived methods to reject predicted locations that were impossible given dive relative to water depth (see Schick et al., 2024; Appendix S1: Section S1).

Simulated MFAS exposure experimental design

Animals were monitored with focal follows (described above) before and then during three 30-min phases: pre-exposure, exposure, and post-exposure. Between one and three focal whales equipped with either a satellite tag and/or a DTag were monitored during experimental trials, which also included between 0 and 10 non-focal tagged whales. Identical methods were used whether the exposure phase included MFAS or not; exposure phases without MFAS are referred to as control trials. We determined that no incidental MFAS was present on exposure trial days using a bottom-mounted passive acoustic recorder (High-frequency Acoustic Recording Packages; Wiggins & Hildebrand, 2016) deployed in the core of our study area. Several exposure trials included truncated exposure phases due to equipment or permit-related requirements, but all lasted for at least 15 min.

The custom sound propagation modeling tool (Margolina et al., 2018) was used to position the source vessel using specified received MFAS criteria. First, we specified a nominal target RL range for focal whales of 120–140 dB re: 1μPa (sound pressure level [SPL] root-mean-square [RMS], hereafter dB SPL). Second, we avoided positioning the source directly in the path of focal whales with a clear direction of travel (Ellison et al., 2012). Third, we avoided positioning sources in locations where model predictions indicated high variability in received sound fields around focal whales due to propagation conditions. Finally, we conducted trials at least 1 km away from other vessels. Mock MFAS exposures were modeled and identical vessel positioning and stationing were used for no-noise control trials.

We employed a 15-element, vertical line array to produce the experimental sound stimulus. We initiated MFAS exposures during a deep dive for focal whales to the extent that this could be predicted in the field. Source specifications, technical design features, and additional details are provided in Southall et al. (2012). Simulated MFAS signals were 1.6-s pulses with a sequence of three tones with fundamental frequencies ranging from 3.5 to 4.05 kHz. The signals were repeated every 25 s for a total of 72 pings over a 30-min exposure period. Transmissions were sequentially increased, as a condition of our research permit, from an initial source output level of 160 dB SPL at 1 m by 3 dB per ping to a maximum source level of 212 dB SPL at 1 m within 8 min. This maximum level was subsequently maintained for all remaining transmissions within a trial.

After exposures, we refined RL estimates in several ways. For whales equipped with DTags, we measured RLs directly from the tags (DeRuiter et al., 2013; Southall, DeRuiter, et al., 2019). For animals equipped with satellite tags, we estimated RLs using methods adapted from Schick et al. (2019) and included ancillary observed animal location data obtained in the field and depth data obtained from tags (as in Schick et al., 2024; additional methodological details given in Appendix S1: Section S1).

We summarized estimated RLs in two ways with common metrics relative to previous studies and future behavioral noise exposure criteria (Southall et al., 2021). First, we determined the distribution of median RL (dB SPL; in RMS as described above) values (and 95% CIs) for each animal for each 5-min time step. We then extracted the maximum of these median RLs for all 5-min time steps to categorize MFAS exposure for each whale. We present continuous values, but for the sake of broader comparisons, simple categorical RL bins were defined as follows: below ambient, <100 dB SPL, 100–120 dB SPL, or >120 dB SPL. Individuals for which MFAS exposures were determined to be below ambient noise conditions were excluded from subsequent analysis. Second, we calculated sound exposure level (in dB re: 1μPa^2-s; hereafter dB SEL) over the 1.6-s duration of each ping for up to 100 imputed locations for each 5-min time step and report median SEL (with 95% CIs). Finally, we calculated median and 95% CIs for cumulative SEL (hereafter dB cSEL) for each of up to 100 tracks by summing SELs for all pings during all 5-min time steps across each track. For each animal and trial, we present RLs as a single predicted exposure in dB SPL units, given the consistent use of this metric in both regulatory assessments and in many previous BRS. However, corresponding per ping SEL as well as cSEL values were calculated to enable future assessments using various exposure metrics (see Southall et al., 2021).

Assessment of behavioral response

We evaluated diving and movement behavior separately and applied several methods over different time windows to provide insights into the nature, magnitude, and time course of potential responses. We describe patterns of responses for four exposure categories: (1) no-MFAS controls; MFAS exposures at (2) "low" (<100 dB SPL); (3) "medium" (100–120 dB SPL); and (4) "high" (>120 dB SPL) RLs. Data and scripts to reproduce analyses and figures for noise exposure and behavioral response are available at: https://doi.org/10.5281/zenodo.17640058.

Diving behavior responses in satellite tag data

We explored potential changes in diving behavior using both exploratory and statistical assessments. First, we compared summary metrics of diving behavior before, during, and following MFAS and control (known absence of MFAS) experimental trials. We then applied a structured data exploration technique developed by Hewitt et al. (2022), which employs kernel density estimates of conditional distributions to compare baseline and exposure behavior along a set of dive metrics. Finally, we used generalized additive models (GAMs) to provide an inferential framework to test for differences between exposure and baseline conditions in metrics identified as important in the exploratory analyses.

Dive metrics, quantile scoring, and kernel density estimates of conditional distributions

As an initial exploratory assessment, we calculated a series of dive metrics over the 24-h immediately following a trial and then compared them to the distribution of each dive metric during baseline diving in non-exposure periods, following Tyack et al. (2011). This approach provides an overview of how unusual a certain behavior may be following exposure but is not an inferential test because it does not include temporal dependence or systematic treatment of error. Methods, dive profiles for all depth-sensing tags around exposure trials, depth metrics and quantile scores for each tag type, and synthesis results of this exploratory analysis are provided in Appendix S1: Section S2: Dive metrics, quantile scoring, kernel density estimates of conditional distributions.

We evaluated potential changes in diving behavior from MFAS exposure in two ways. The first was the structured exploratory data analysis tool described by Hewitt et al. (2022) to characterize diving behavior during trials and evaluate differences from baseline behavior, while explicitly addressing temporal dependence. This approach was only possible for the subset of depth time-series tags for which sufficient baseline and exposure data were collected ($n = 38$). The time-series depth data were summarized into

10 metrics (Table 1), with an additional 10 metrics having timing-dependent variations.

In this analysis, we used specified pre- and post-exposure windows to screen for changes in diving behavior. Based on earlier experiments in which temporal comparisons of fine-scale responses were evaluated (DeRuiter et al., 2013; Wensveen et al., 2019), we selected a 6-h window following exposure to assess potential acute responses. Within this window, we configured a set of nested windows, with longer 3-h windows from 0–3 to 3–6 h and shorter 1-h windows from 0–1 h through 5–6 h from the end of the exposure phase (see Hewitt et al., 2022).

Observed differences in these specified time windows before and after exposure (pre-post pairs) were then compared to hypothetical pairs of windows of the same length and placement in baseline diving periods, following a randomization procedure. Baseline diving data periods for this analysis were defined as time windows at least 24-h prior to the evaluated exposure trial and excluding any previous exposure periods and after 24-h had passed following exposure(s). We quantified possible responses following a randomization procedure to determine the quantile of the observed pre-post pair in the reference distribution of hypothetical pairs (for additional details, see Appendix S1: Section S2: Dive metrics, quantile scoring, kernel density estimates of conditional distributions). We identified candidate behavior changes as those below the 0.025 or above the 0.975 quantiles.

TABLE 1 Dive metrics considered by dive metric distribution quartile scoring, kernel density estimates, and generalized additive model (GAM) diving analysis aligned to show related metrics.

Dive metric distributions	Kernel density estimates	GAM diving analysis
IDDI duration Prior IDDI duration n shallower dives	Time on or near surface	IDDI duration
Mean shallower max depth IDDI maxdepth Dive maxdepth	Average depth	na
Dive duration	Time in deep dives	dive duration
Dive ascent duration	Time spent ascending	na
na	Time spent descending	na
na	Time spent without vertical movement	na
na	Total vertical distance traveled	na
na	Total ascent distance	na
na	Total descent distance	na
na	Total vertical direction changes	na

Abbreviations: IDDI, inter-deep-dive interval; na, not applicable.

GAMs for diving and surface behavior

To examine differences in deep-dive duration and IDDI durations between baseline and an exposure window, sets of GAMs were fitted using the *bam* function in the R package *mgcv* (Wood, 2017; Wood et al., 2016). For both sets of models, we used Gamma distributions with a log link. We elected to model deep-dive and IDDI durations because these metrics could be calculated for both tag programming regimes. Further, the results of previous studies with beaked whales and MFAS, as well as our exploratory analyses, suggested that these metrics were particularly likely to be affected by exposure to MFAS. We fitted one GAM per animal per metric (deep-dive or IDDI duration).

In addition to the exclusion of 24-h windows surrounding previous exposures from the baseline, we also removed exposure windows with data gaps if they included less than 6-h of continuous data within the 24-h exposure window. In cases in which this threshold was not met, the entire window was excised from the data. We used Akaike information criterion (AIC) to select between models with either no exposure effects, a simulated MFAS source effect, a control effect, or both a simulated MFAS source and control effect (depending on which types of exposures were present in that tag record). Note that not all models had both a control and simulated MFAS source exposure. Only AIC differences >2 were considered meaningful for the purposes of model selection. We constructed exposure effects as time-varying thin-plate spline smooths (with the penalty modification to allow shrinkage to zero) that operate only during the 24-h exposure windows. We constrained these smooths to pass through 0 at the end of the exposure window, reflecting our assumption that any response would have abated by this time. We used a first-order penalty ($m = 1$ in the smooth specification in *mgcv*) to reduce collinearity between the baseline and effect smooths as recommended by Pedersen et al. (2019) and to encourage the effect smooth to 0 before the end of the 24-h window if the data suggested an earlier return to baseline. We considered a response to have been detected if the best model included an exposure term. All models included a thin-plate spline smooth term for continuous running time, with a penalty modification to allow shrinkage to zero (Wood, 2003), and a cyclic spline smooth term for time of day. We used the built-in functionality of *bam* to account for possible autocorrelation in the residuals by setting the rho parameter to the estimated first-order autocorrelation in residuals from an initial fit that did not include autocorrelation.

Horizontal avoidance responses for individuals with satellite tags

To evaluate whether the animals exhibited horizontal avoidance responses following exposure to sonar, we used the continuous-time discrete-space movement model from Hanks et al. (2015). The model examines the residence time in each cell, the transitions from one cell to the next, and the covariates associated with each of these. To represent the potential effects of MFAS exposure, we computed the Euclidean distance to the source location at the start of the trial for each cell in a 500-m grid. Following Hanks et al. (2015), we used the distance to source as a gradient-based covariate, and we allowed the movement response of the animal with respect to the position of the ship to change over time.

For each animal and trial, we examined movements at a 15-min temporal resolution from 24 h before and after the onset of the exposure. The data were prepared using the R package *ctmcmove* (Hanks et al., 2015), and the model was fitted using the GAM function in the R package *mgcv* (Wood, 2017; Wood et al., 2016). We used a basis dimension of 40 and an adaptive P-spline smoother to allow for rapid changes in the smooth. We interpreted the results to indicate that the animal responded to the source if the time-varying coefficient for the distance to source covariate was both significant and negative in the 24 h following exposure. If the plot of the estimated change was clearly negative prior to the exposure, we did not include this as evidence of horizontal avoidance. Additional details of this approach to evaluate horizontal avoidance are provided in Appendix S1: Section S4: Horizontal avoidance analysis.

Fine-scale behavior responses in DTag records

We measured diving behavior and spatial orientation from DTags using standard filtering and processing methods (DeRuiter et al., 2013, 2023; Southall et al., 2023; Southall, DeRuiter, et al., 2019; R package version 0.1.0, https://animaltags.org/). We quantified the occurrence of echolocation clicks during baseline and exposure dives using customized MATLAB tools from Kootstra (2024). We conducted statistical analyses of diving behavior and potential changes using the approach developed by Michelot et al. (2021, 2023). Depth data were decimated to one sample every 3 s, after first applying a no-delay low-pass anti-alias filter (symmetric finite impulse response filter with length 900 and cutoff frequency 0.133 Hz). We computed the first difference of the depth measurements to reduce temporal autocorrelation in the time-series before modeling. Dives beyond 50 m were classified as either deep or shallower, as defined above. Dives were classified as exposure dives if a MFAS signal was present during that dive or baseline dives if there was no exposure. All three DTag exposure trials included MFAS transmissions.

To evaluate behavioral changes in DTag data, we fitted varying-coefficient stochastic differential equations (SDEs) (Michelot et al., 2021, 2023) to depth difference data in shallow and deep dives respectively. Each model described depth as a function of the proportion of dive time elapsed and employed a difference smooth to quantify changes in the model's diffusion parameter, which governs the variability of depth difference, between baseline and exposed dives. The model also included a random effect to account for individual differences.

Assessment of social responses

We evaluated social group configurations for focal groups before, during, and following exposure trials using standard photo identification methods with comparisons to a catalog of known individuals. This was only possible for a subset of focal groups due to the challenges of relocating animals after exposure (e.g., weather conditions; behavioral response of animals). Social response analysis methods and example results are presented in Appendix S1: Section S6: Assessment of social response.

RESULTS

Fieldwork occurred over approximately 100 days between June 2017 and August 2022. During this period, we deployed 65 satellite-linked tags and six DTags. Of these, 50 satellite-linked tags and 3 DTags yielded data during trials, either with exposure to simulated MFAS or as controls. Some individuals were included in more than one trial. We conducted 13 exposure trials, resulting in 72 total *Ziphius* exposure events, 69 of which included satellite-tagged individuals and three of which included whales with DTags. Of the 13 exposure trials, 9 included simulated MFAS with between 4 and 11 tagged whales and 4 were no-MFAS controls with between 2 and 10 individuals. A summary of tag deployments and trials is provided in Table 2. DTag exposure events are considered separately below.

For no-MFAS control trials, any individuals farther than 128.5 km (the outer extent for which RL modeling was performed) from the (inactive) sound source were excluded from subsequent response analyses. Median (and 95% CIs) modeled ranges and median (and 95% CIs) modeled RLs (SPLs) for simulated MFAS exposure trials are presented in Figure 1. The same range information from control trials is shown in Figure 2. Individual plots of modeled horizontal range for each exposure event and individual plots showing the distribution of RL values for each MFAS exposure trial at each 5-min timestep are included in Appendix S1: Section S1: Figures S1–S47.

Dive metric distributions, quantile scoring, and kernel density estimates of conditional distributions

Dive plots of all satellite-tagged individuals centered around exposure periods are provided in Appendix S1: Section S2: Figures S48–S86. Dive metric distributions for all satellite-tagged exposed individuals and exposures analyzed using quantile scoring assessments are provided in Appendix S1: Section S2: Figures S87–S123 and summarized across individuals (Appendix S1: Section S2: Figure S124).

For kernel density estimates, the proportion of excess detected behavior changes (i.e., the ratio of observed changes to the number expected by chance) increased with increasing RL exposure (Figure 3). Figure 3a,b highlights one animal (STag093 in MFAS exposure trial 19_03) whose time deep-diving was identified as changed during exposure. Figure 3b,d highlights an individual (STag096 in MFAS exposure trial 19_04) whose behavior did not differ for this metric. Summary results of excess detected changes across individuals by categorical RL bins are shown in Figure 3e. Detailed results plots for all individuals and exposure trials analyzed can be found in Appendix S1: Section S2: Figures S125–S146.

We detected more changes than expected by chance for MFAS exposure trials in the two highest RL categories. The two focal whales highlighted in Figure 3 exhibited very strong changes in diving behavior following exposure to MFAS signals. A few control animals (notably STag131) also exhibited excess differences. Percentile scores for individual metrics for each exposure event are summarized discretely (Figure 4).

When we examined metrics (all windows and lags collapsed) by exposure event (Figure 4), we found that diving changes in exposed animals were characterized by whales spending more time in deep diving or near the surface in the two highest RL bins. For the other two RL bins, we detected more excess changes for the no-MFAS control category than the lowest MFAS exposure treatment level. The two most typical changes detected in MFAS trials of any category were longer deep dives and longer IDDIs. In contrast, changes detected in no-MFAS controls included a larger proportion of shorter-than-typical deep dives and IDDIs. This suggests that behavioral changes detected in control trials were of a different nature than those observed during MFAS trials.

GAM analyses of diving behavior

We examined the 39 exposure events for which there was sufficient dive data available for GAM analyses on 34 tags (including both types of tag programming) for a total of 78 models. We detected changes in the medium and high sound exposure categories more frequently than in the low exposure category or control treatments (Figure 5). For deep-dive duration, no changes were detected in the controls, but changes were detected in 5.5% of low, 18.8% of medium, and 31.8% of high RL conditions. Similarly, a change in IDDI duration was detected in 5.5% of the controls and 0% of low, 12.5% of medium, and 22.7% of high exposure levels. The overall trend we observed was consistent with the dive metric quantile and kernel density assessments (see Table 3), although there was some variation in which methods detected behavioral change(s) for specific exposure events. Individual GAM analysis plots for all exposure and control trials analyzed can be found in Appendix S1: Section S3: Figures S147–S180.

Horizontal avoidance responses

We conducted horizontal avoidance analyses for all MFAS control and MFAS exposure events ($n = 65$) with satellite-tagged individuals for which sufficient baseline and post-exposure position data were available. Overall, we detected spatial avoidance in 42% of all experimental trials (MFAS and controls) during the 24-h period following exposure. This included relatively low proportions of avoidance in control trials (33%) and low exposure categories (29%). More individuals demonstrated horizontal avoidance at higher RLs: 50% in medium and 56% in high exposure conditions. Time-varying smooths for the effect of distance to source for two selected individuals are shown, along with a summary of the results for all

TABLE 2 Summary of satellite-tagged *Ziphius* during simulated mid-frequency active sonar (MFAS) and no-MFAS (control) exposure trials.[a]

Subject ID	Dive data type	Tag duration (days)	Exposure trial no.	Trial type	Trial date	Range (km) (95% CI)	RL (dB SPL) (95% CI)	RL (dB SPL) bin	RL (cSEL) (95% CI)
Tag060	Dive summary	34.4	17_01	MFAS	8/22/17	39.8 (32–54.4)	95.5 (81.5–121.5)	<100	107.5 (103.0–121.6)
Tag061	Dive summary	44.0	17_01	MFAS	8/22/17	17.4 (7.9–34.3)	115.7 (92.9–132.7)	100–120	123.6 (99.1–139.3)
Tag062	Dive summary	12.1	17_01	MFAS	8/22/17	3.6 (0.7–7.5)	136.3 (116.9–150.5)	>120	146.9 (124.9–161.1)
Tag063	Dive summary	30.1	17_01	MFAS	8/22/17	18.9 (5.6–35.8)	118.5 (87.6–139.3)	100–120	128.7 (102.1–146)
Tag064	Dive summary	35.0	17_01	MFAS	8/22/17	8.7 (1.4–16.5)	128.9 (110.1–146.1)	>120	138.7 (118.1–156.9)
Tag065	Dive summary	12.1	17_01	MFAS	8/22/17	2.1 (0.6–5.6)	136.3 (118.4–148.4)	>120	144.3 (130.7–160.7)
Tag069	Depth time-series	38.1	18_02	MFAS	5/25/18	20.6 (16.5–27.4)	101.8 (87.1–110.9)	100–120	118.1 (112.8–127.2)
			18_03	MFAS	5/30/18	3.8 (1.8–6.4)	141.6 (112.4–148.4)	>120	158.9 (153.6–163.7)
			18_05	Control	6/6/18	122.8 (99–146)	na	Control	na
Tag070	Dive summary	11.6	18_02	MFAS	5/25/18	17.0 (14.5–18.9)	116.5 (101.0–132.1)	100–120	133.7 (124–141.2)
			18_03	MFAS	5/30/18	99.6 (85.1–111.6)	87.6 (81.7–94.6)	<100	102.1 (101.1–107.9)
			18_05	Control	6/6/18	29.7 (24.1–36.5)	na	Control	na
Tag071	Depth time-series	25.5	18_07	Control	8/15/18	16.8 (2.4–29.5)	na	Control	na
			18_08	MFAS	8/27/18	45.1 (10.7–111.1)	84.4 (70.1–114.5)	<100	100.8 (91.0–126.0)
Tag072	Depth time-series	42.8	18_07	Control	8/15/18	26.3 (15–36.1)	na	Control	na
			18_08	MFAS	8/27/18	205.5 (194.3–219)	<Ambient[b]	na	na
Tag073	Depth time-series	43.4	18_07	Control	8/15/18	51.1 (30.5–71.6)	na	Control	na
			18_08	MFAS	8/27/18	58.3 (37.4–74.9)	78.4 (78.2–78.4)	<100	96.3 (95.8–96.9)
Tag075	Depth time-series	41.2	18_07	Control	8/15/18	42.5 (35.6–50.9)	na	Control	na
			18_08	MFAS	8/27/18	61.4 (54.3–74.5)	<Ambient[b]	na	na
Tag076	Dive summary	41.3	18_07	Control	8/15/18	65.4 (47.4–82.1)	na	Control	na
			18_08	MFAS	8/27/18	41.4 (34.6–46.2)	75.8 (73.8–80.6)	<100	93.5 (92.9–95.0)

TABLE 2 (Continued)

Subject ID	Dive data type	Tag duration (days)	Exposure trial no.	Trial type	Trial date	Range (km) (95% CI)	RL (dB SPL) (95% CI)	RL (dB SPL) bin	RL (cSEL) (95% CI)
Tag077	No dive data	23.7	18_07	Control	8/15/18	19.0 (16.3–22.3)	na	Control	na
			18_08	MFAS	8/27/18	33.2 (27.6–38.8)	77.1 (73.7–84.4)	<100	94.3 (93.3–96.8)
Tag078	Dive summary	23.1	18_07	Control	8/15/18	4.0 (1.3–9.5)	na	Control	na
			18_08	MFAS	8/27/18	14.3 (3.3–29.3)	106.2 (78.7–126.1)	100–120	124.8 (103.7–147.9)
Tag079	Dive summary	42.5	18_07	Control	8/15/18	32.3 (22.1–40.8)	na	Control	na
			18_08	MFAS	8/27/18	22.2 (4.3–40.1)	90.9 (74–118.6)	<100	115.4 (101.2–132.4)
Tag080	Dive summary	41.4	18_07	Control	8/15/18	34.2 (22.8–45.6)	na	Control	na
			18_08	MFAS	8/27/18	55.2 (37.7–72.6)	76.3 (73.6–82.2)	<100	93.7 (92.4–94.2)
Tag081	Dive summary	56.7	18_07	Control	8/15/18	30.8 (12.7–51.2)	na	Control	na
			18_08	MFAS	8/27/18	33.6 (15.9–51.2)	82.8 (71.7–108.8)	<100	99.5 (94.2–123.8)
Tag082	Depth time-series	52.6	19_01	MFAS	5/21/19	29.1 (13.9–46.6)	112.4 (95.6–125.8)	100–120	119.5 (112–129.4)
			19_02	MFAS	6/7/19	60.0 (52–66.9)	98.9 (85.4–109.7)	<100	116.4 (110.3–119.8)
Tag083	Depth time-series	39.2	19_01	MFAS	5/21/19	13.3 (4.7–21)	117.4 (102.0–136.8)	100–120	123.8 (110–138.2)
			19_02	MFAS	6/7/19	16.1 (9.3–24.5)	123.8 (91.1–130)	>120	141.1 (134.2–147)
Tag084	Depth time-series	44.1	19_02	MFAS	6/7/19	33.5 (16.2–52.6)	112.6 (91.2–126.6)	100–120	130.1 (112.7–143.1)
Tag085	Depth time-series	40.9	19_02	MFAS	6/7/19	25.1 (15.3–42.9)	116.5 (98.5–132.4)	100–120	135.9 (128.9–146.8)
Tag086	Depth time-series	13.1	19_02	MFAS	6/7/19	16.7 (11.1–23.6)	124.1 (108.4–135.2)	>120	140.5 (132.9–149.4)
Tag087	Depth time-series	21.0	19_02	MFAS	6/7/19	16.5 (4–40.6)	119.3 (104.0–138.9)	100–120	139.8 (133.2–146.7)
Tag088	Depth time-series	43.4	19_02	MFAS	6/7/19	53.4 (37.8–74.2)	108.0 (93.9–118.6)	100–120	125.1 (119.8–129)
Tag089	Depth time-series	27.5	19_02	MFAS	6/7/19	4.8 (2.5–8.3)	133.9 (119.0–142.5)	>120	151.0 (145.7–154.9)
Tag090	Depth time-series	16.0	19_03	MFAS	8/6/19	15.7 (5.7–25.4)	119.1 (100.4–134.7)	100–120	135.9 (124.7–149.6)
Tag091	Depth time-series	13.6	19_03	MFAS	8/6/19	68.8 (59.6–78.9)	95.5 (85.7–105.1)	<100	111.9 (108.7–114.6)

(Continues)

TABLE 2 (Continued)

Subject ID	Dive data type	Tag duration (days)	Exposure trial no.	Trial type	Trial date	Range (km) (95% CI)	RL (dB SPL) (95% CI)	RL (dB SPL) bin	RL (cSEL) (95% CI)
Tag092	Depth time-series	40.6	19_03	MFAS	8/6/19	13.7 (3.6–28.3)	122.8 (99.3–141.2)	>120	137.2 (120.2–157.2)
			19_04	MFAS	8/19/19	24.6 (5.8–48.2)	103.5 (75.9–134.6)	100–120	112.7 (95.1–138.4)
Tag093	Depth time-series	24.9	19_03	MFAS	8/6/19	4.3 (1.2–7.6)	134.8 (119.4–144.6)	>120	151.8 (146–156.4)
			19_04	MFAS	8/19/19	12.9 (3–25.6)	116.8 (90.3–145.2)	100–120	125.4 (102.4–161)
Tag095	Depth time-series	37.7	19_04	MFAS	8/19/19	3.7 (1.3–7.2)	134.3 (110.3–146.5)	>120	151.9 (137.8–158.2)
Tag096	Depth time-series	43.2	19_04	MFAS	8/19/19	3.4 (0.9–7.2)	133.8 (110.9–147)	>120	151.5 (136.7–158.8)
Tag097	Depth time-series	36.1	19_04	MFAS	8/19/19	3.7 (1.2–6.4)	124.0 (109.1–147.1)	>120	147.5 (134.4–156)
Tag098	Depth time-series	40.4	20_03	MFAS	8/19/20	41.8 (21.9–61.2)	95.7 (72.2–113.5)	<100	98.1 (91.6–112.4)
Tag101	Depth time-series	71.4	20_03	MFAS	8/19/20	40.9 (27.6–51.4)	90.5 (79.7–105.7)	<100	109.1 (102.4–119.4)
Tag102	Depth time-series	28.7	20_03	MFAS	8/19/20	75.7 (57.8–94.4)	82.8 (73.3–100.2)	<100	99.7 (93.3–108.1)
Tag103	Depth time-series	33.1	20_03	MFAS	8/19/20	86.4 (63.4–105.5)	78.5 (66.5–97.6)	<100	94.6 (89.5–105.3)
Tag105	Depth time-series	78.4	20_03	MFAS	8/19/20	9.8 (3.2–19.1)	123.6 (105.6–138.5)	>120	139.8 (132.1–154.6)
Tag106	Depth time-series	74.0	20_03	MFAS	8/19/20	119.1 (111.6–126.3)	71.0 (68.7–77)	<100	89.0 (87.7–90.7)
Tag107	Depth time-series	77.9	20_03	MFAS	8/19/20	36.3 (25.7–45.2)	84.1 (72.2–105)	<100	102.5 (94–112.5)
Tag108	Depth time-series	65.9	20_03	MFAS	8/19/20	7.5 (2–15.4)	126.3 (103.7–141.4)	>120	143.0 (122.4–163.7)
Tag109	Depth time-series	47.3	20_03	MFAS	8/19/20	132.3 (123.7–139.5)	<Ambient[b]	na	na
Tag110	Depth time-series	85.4	20_03	MFAS	8/19/20	4.3 (2.1–12.1)	133.7 (94.3–137.1)	>120	149.3 (142.3–151.7)
Tag111	Depth time-series	68.7	20_03	MFAS	8/19/20	41.4 (33.9–49.2)	89.3 (80–98.4)	<100	107.1 (102.5–112.2)
Tag129	Depth time-series	73.9	22_01	Control	8/4/22	3.9 (0.6–10)	na	Control	na
			22_02	Control	8/5/2022	196.4 (187.6–204.8)[c]	na	na	na
Tag130	Depth time-series	74.8	22_01	Control	8/4/2022	209.4 (205.9–213.2)[c]	na	na	na
			22_02	Control	8/5/22	40.1 (30.9–50.9)	na	Control	na

TABLE 2 (Continued)

Subject ID	Dive data type	Tag duration (days)	Exposure trial no.	Trial type	Trial date	Range (km) (95% CI)	RL (dB SPL) (95% CI)	RL (dB SPL) bin	RL (cSEL) (95% CI)
Tag131	Depth time-series	9.1	22_01	Control	8/4/22	171.0 (167.6–175.0)[c]	na	na	na
			22_02	Control	8/5/22	8.8 (2.3–15.9)	na	Control	na
Tag132	Depth time-series	76.4	22_01	Control	8/4/22	190.8 (187.0–194.6)[c]	na	na	na
			22_02	Control	8/5/22	27 (22.6–32)	na	Control	na
Tag133	Depth time-series	66.5	22_02	Control	8/5/22	30.1 (22.4–39)	na	Control	na
Tag134	Depth time-series	39.7	22_02	Control	8/5/22	34.6 (25.4–41.9)	na	Control	na

Abbreviation: na, not applicable.

[a]Tag and exposure trial types for each individual exposure event are shown along with median horizontal source-whale range (in kilometers; with 95% CI), median estimated MFAS received level (RL) (root-mean-square [RMS] sound pressure level [SPL] in dB re: 1μPa; with 95% CI), RL (dB SPL) bins (<100; 100–120; >120 dB), and median estimated cumulative sound exposure level (cSEL in dB re: $1\mu Pa^2$-s; with 95% CI).

[b]Denotes individuals for which fewer than five of 100 imputed track points had RLs above frequency-band specific ambient noise; exposure event is excluded from subsequent response analyses.

[c]Denotes individuals for no-MFAS controls conducted greater than 128.5 km from mock sound sources; exposure event is excluded from subsequent response analyses.

individuals (Figure 6). Plots for all individuals and exposure trials analyzed can be found in Appendix S1: Section S4: Figures S181–S244.

Detailed views of fine-scale movements of two example focal animals during exposure trials highlighted in Figure 6 are presented below (Figure 7). In both cases, we observed directed movement away from the source immediately following the MFAS exposures. Importantly, both movements were against the direction of the Gulf Stream, which flows at speeds of up to 10 km/h. Similar views for other focal individuals and exposure trials, not all of which demonstrated such clear and sustained avoidance, are found in Appendix S1: Section S4: Figures S245–S252.

Fine-scale behavioral responses in DTag records

We deployed three DTags (DTag17_234 in exposure trial 17_01, DTag19_218 in exposure trial 19_03, and DTag20_232 in exposure trial 20_03); all three whales were focal individuals during experimental trials. Dive profiles, measurements of MFAS RLs, heading, and minimum specific acceleration (MSA) are shown for each whale (Figure 8). All three individuals produced echolocation clicks during baseline deep foraging dives but ceased echolocation following MFAS exposure and for the duration of tag deployments. We overlaid baseline and exposed dive profiles for each of the three DTag records (Figure 9). The shifts in depth rates during exposure were summarized with difference smooths from the SDE output (Michelot et al., 2021, 2023), which showed larger differences from baseline in deep dives than in shallow dives (Figure 9). More detailed descriptions of individual behavior and response analyses for each of these three DTag whales are provided in Appendix S1: Section S5: Fine-scale behavior responses in DTag records.

Multi-scale integration of STag093 and DTag19_218

A synoptic comparison of data on differing time scales from STag093 and DTag19_218, two whales in the same group on the day of exposure trial 19_03, is shown in Figure 10.

The dive records from these two individuals showed a high degree of synchrony and consistency. The satellite tag record captured the horizontal displacement that occurred shortly after the MFAS exposure and continued throughout the following dive cycle. The animal moved back toward the pre-exposure location during the next dive cycle. The DTag record included the exposure dive and two subsequent dives. Clear, sustained increases in MSA (depicted in Figure 8b for this individual) were likely at least partially caused by increased fluking related to the horizontal

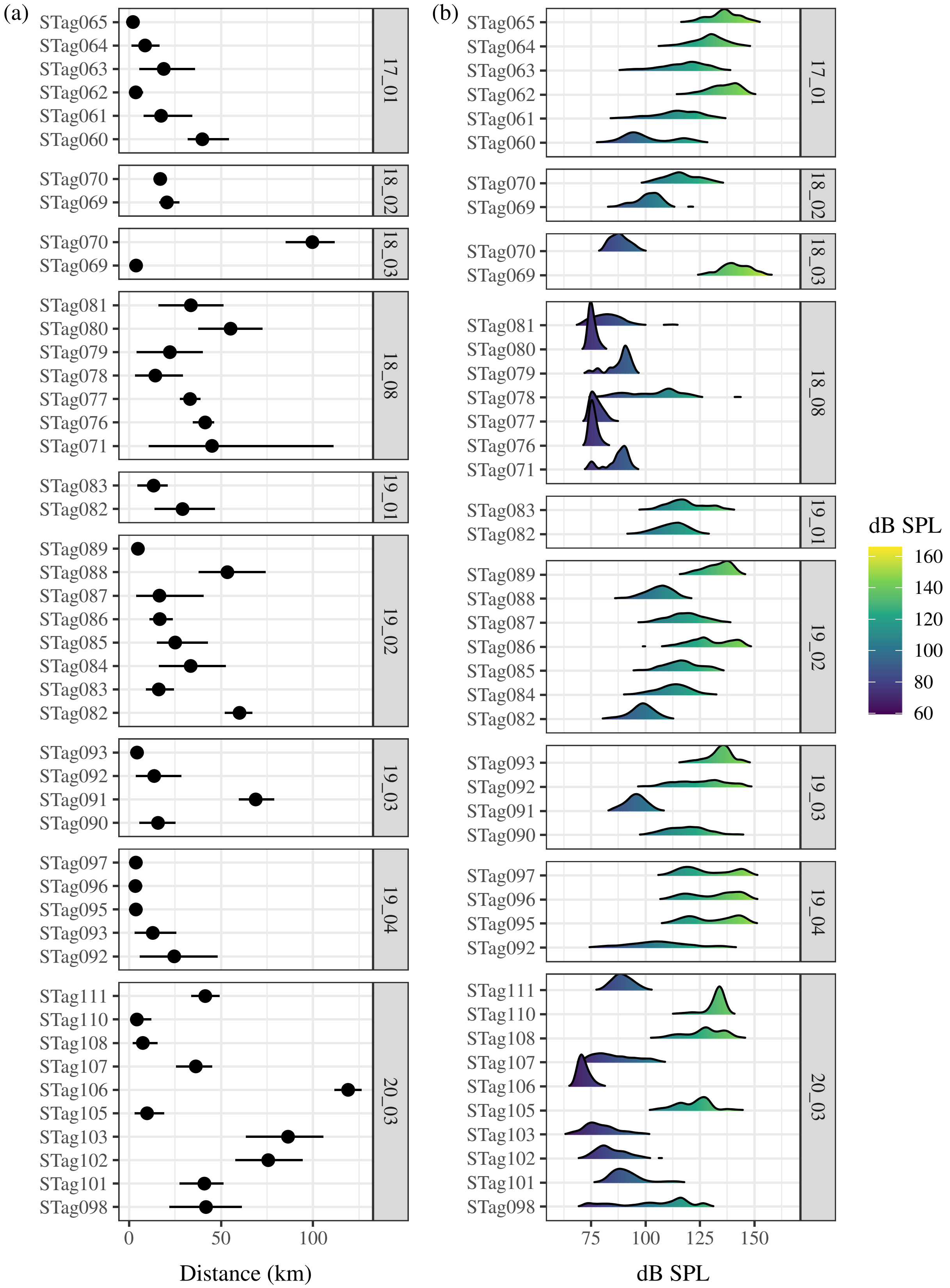

FIGURE 1 Summary plots of (a) modeled horizontal range (median: black dots; 95% CIs: black bars) from satellite-tagged (STag) *Ziphius* to sound source locations at the start of each mid-frequency active sonar (MFAS) exposure trial (trial IDs in gray side panel) and (b) modeled received level (RL) (dB re: 1μPa sound pressure level (SPL) distributions for all imputed track positions during the loudest 5-min interval for corresponding trials. Respective STag identification numbers for tagged individuals are given.

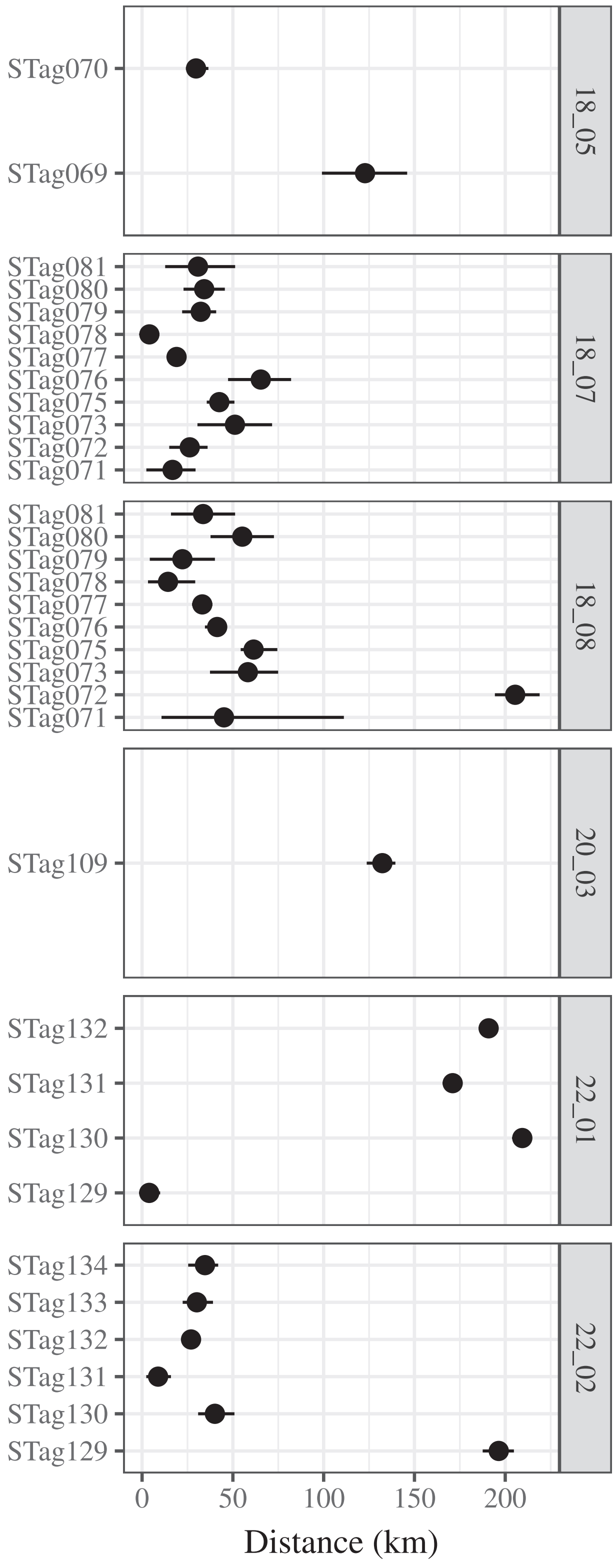

FIGURE 2 Summary plots of modeled horizontal range (median: black dots; 95% CIs: black bars) from satellite-tagged (STag) *Ziphius* to sound source locations at the start of each no-mid-frequency active sonar (MFAS) (control) exposure trial (trial IDs appear in gray side panel). Respective STag identification numbers for tagged individuals are given.

displacement observed in the satellite tag record. These two animals were photographed the following day, still together in the same group (see Appendix S1: Section S6: Figure S265).

Assessment of social response

In 13 instances, we were able to resight tagged individuals before, during, and after an exposure trial or before and after an exposure trial (see Appendix S1: Section S6: Figures S253–S264). In four cases we were able to capture group composition the day before, the day of, and the day after an exposure trial (see Appendix S1: Section S6: Figure S265). In this small dataset we documented some changes in social group composition over these periods but did not observe large effects such as complete dissolution or dramatic enlargement of groups. It is possible, however, that there are subtler effects that could be detected with a larger sample size and a formal comparison to and better understanding of baseline fission-fusion dynamics.

Summary of evidence for behavioral responses

Summary results providing strength of evidence for detection of a response in each exposure event are provided in Table 3. The results are presented by increasing median maximum RL and categorically by RL bins for potential diving response and for spatial avoidance. We used a structured categorical assessment for quantile scoring (described above in *Methods*: *Dive metrics, quantile scoring, and kernel density estimates of conditional distributions*; additional details in Appendix S1: Section S2: none, 0 parameters below 0.025 or above 0.975 quantile; limited, 1 parameter; moderate, 2 parameters; strong, >3 parameters) and kernel density estimates of conditional distributions (from *GAM analyses of diving behavior*, above; none, <2 excess changes detected; limited, 2–3 excess changes; moderate, 3–4 excess changes; strong, >4 excess changes). Results for responses detected for diving and spatial avoidance changes are presented as binary YES/NO summaries for dive and surface GAM and horizontal avoidance analysis. Individual DTag results all of which exhibited variable but clear behavioral changes as focal animals in the highest RL exposure condition are summarized descriptively and quantitatively (see Figures 8 and 9; Appendix S1: Section S5: Fine-scale behavior responses in DTag records).

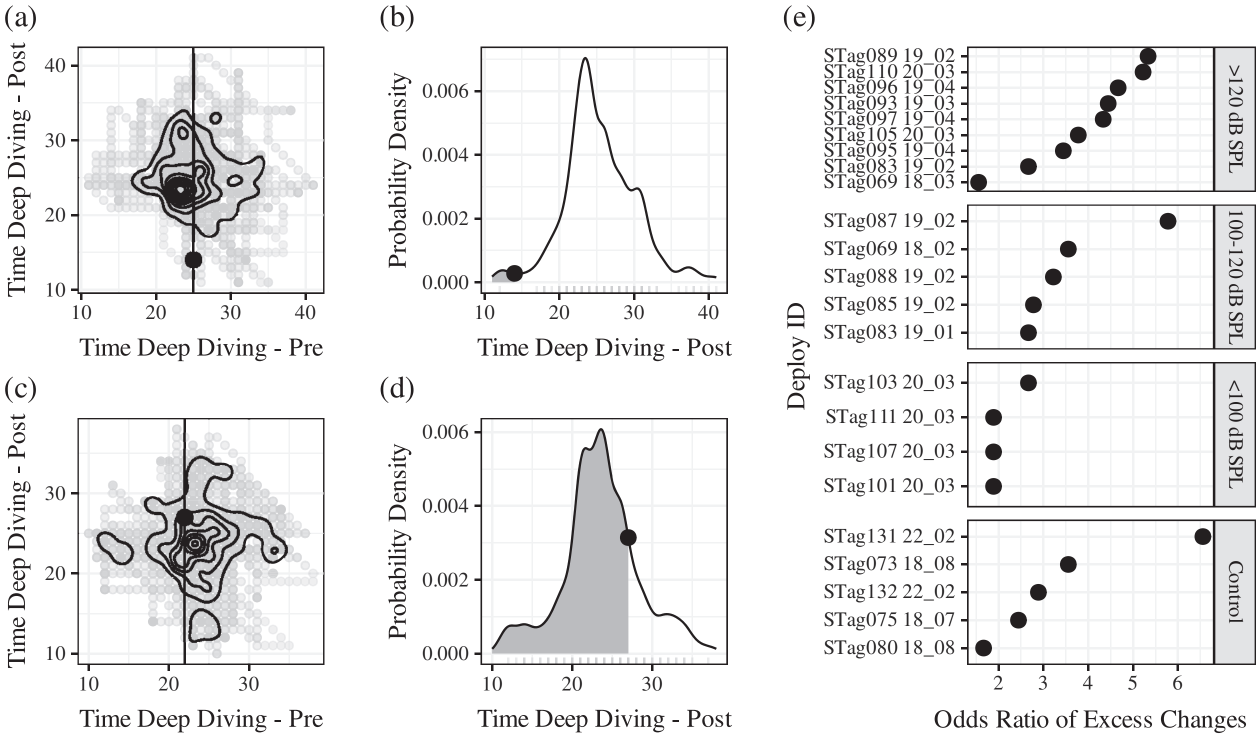

FIGURE 3 Kernel density estimates (2-D) for all pairs in pre versus post-exposure periods for one diving metric are shown for STag093 (in mid-frequency active sonar [MFAS] exposure trial 19_03) (a) and STag096 (in 19_04) (c) at the time of exposure. The 1D distributions in panels (b) and (d) correspond to the black vertical lines in panels (a) and (c). This vertical line corresponds to the value of the statistic at the time of the exposure, and the black dot along the line denotes the observed value. By comparing this observed pair against the distribution of random pre-post pairs, we can calculate a probability and use a critical threshold to determine when a change is detected (shaded gray area in panels (b) and (d)). (e) A summary plot showing the number of excess changes in all metrics (as % of expected by chance) for four received level (RL) bins (from bottom: no-MFAS controls; <100, 100–120, >120 re: 1μPa sound pressure level (SPL).

SYNTHESIS

We evaluated the behavioral responses of 53 *Ziphius* during 13 exposure trials (9 MFAS; 4 no-MFAS controls), resulting in 72 total exposure events. Here, we evaluate the results in terms of modeled exposure conditions (range and RL), observed changes in diving behavior, and patterns of spatial (horizontal) avoidance of sound sources during exposure trials. Modeled horizontal ranges for individuals at the start of exposure trials ranged from ~2 to 205 km for MFAS exposures and ~4 to 209 km for no-MFAS control sequences. Three individuals were excluded from analyses because modeled MFAS RLs were below estimated ambient noise conditions with the criteria we applied. Four individuals were excluded from control sequences because modeled horizontal distances from the source were too great. The other 65 individuals were included in at least some assessments. Given our experimental approach, focal individuals were exposed at shorter ranges (<5 km) than non-focal individuals. We did not specifically manipulate and test for differences in response for different range and RL conditions, but the data we present here include greater range and lower RL conditions, given that most previous CEEs have focused on a small number of focal animals close to sound sources.

We achieved target RLs (120–140 dB SPL) for focal individuals during MFAS exposure trials, with median maximum RLs ranging from 124.0- to 141.6 dB SPL. Non-focal individuals were exposed to levels from below ambient to 71.0- to 123.8 dB SPL. There was a roughly equal number of exposed individuals by RL category ($n = 18$ for controls; $n = 18$ for the low RL bin; $n = 14$ for the medium RL bin; and $n = 15$ for the high RL bin). This enabled us to evaluate the probability of potential responses across a range of target values with multiple individuals in each exposure trial. Our multi-scale design provided several unique insights. Whales equipped with DTags were always identified as focal individuals and thus, by design, were exposed in the high RL category. The kinematic records from these individuals provided fine-scale details of the behavioral response in this exposure range. The DTags also provided our only insights into echolocation

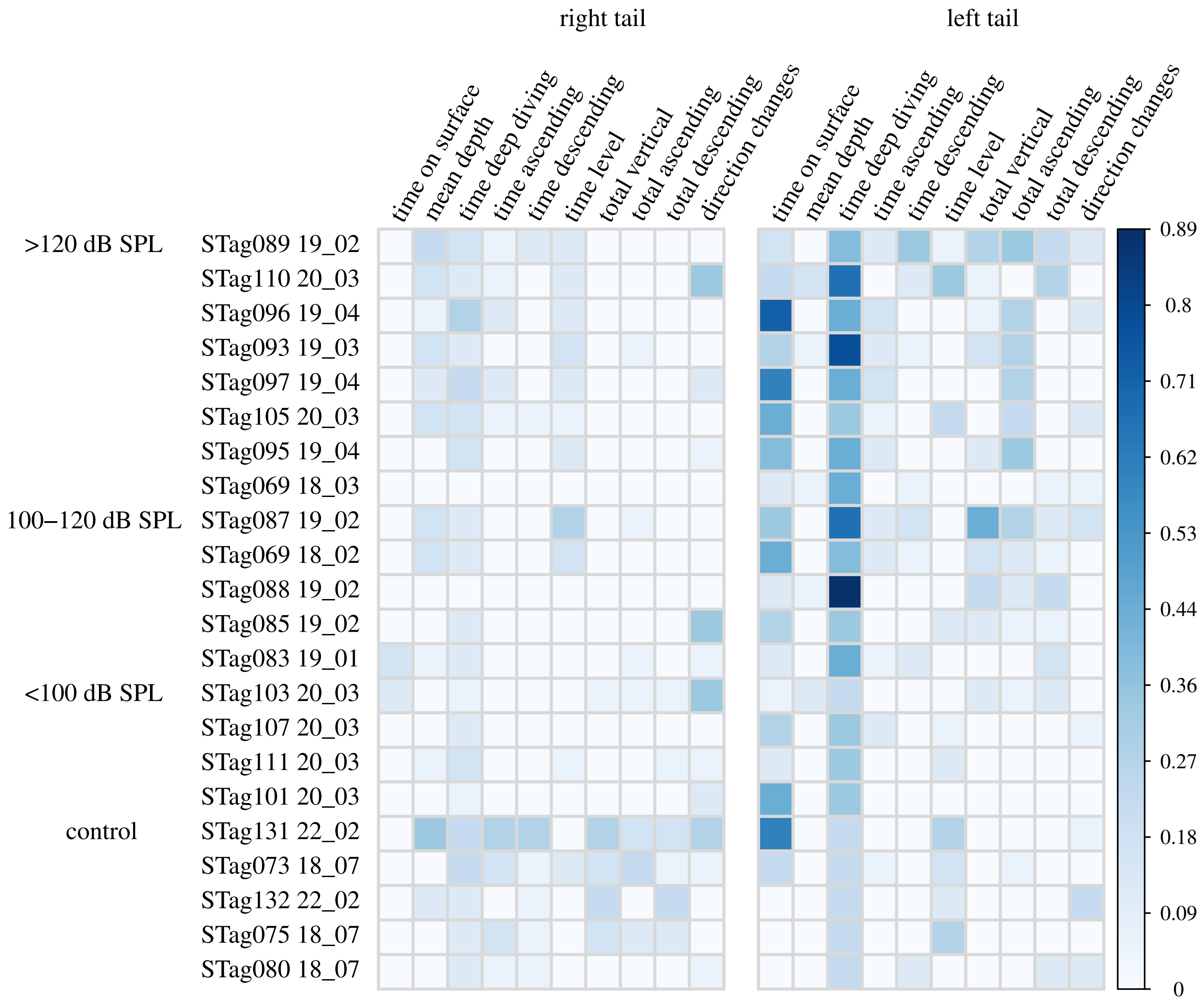

FIGURE 4 Matrix representation of each exposure event and a summary of detected changes for the kernel density estimates of conditional distributions. Each metric was collapsed for all lags, windows, and timing dependence and the percentage of changes detected (outside the central 95 percentiles) was calculated. Right and left tail changes are separated for each metric. Rows are sorted by received level (RL) bins (from bottom: no-mid-frequency active sonar (MFAS) controls; <100, 100–120, >120 dB re: 1μPa sound pressure level (SPL)) and columns are sorted by the changes detected across all exposure events.

behavior, which was consistently and abruptly terminated following MFAS exposure.

Descriptive analyses of diving behavior identified several clear patterns of behavioral changes in response to MFAS exposure. We documented substantial variability in the probability of response, as some individuals in the highest RL category exhibited few responses and some individuals in the control and low RL categories exhibited more responses. Perhaps this is not unexpected given the large sample size and existence of contextual factors outside the control of our experimental design. Despite this variation, we typically observed similar patterns of behavioral changes in control trials and the lowest exposures, with moderate to strong indications of changes in diving behavior in only 1 of 10 individuals in control trials and 2 of 8 in low MFAS RL exposures. We observed greater indications of changes in diving behavior in individuals for higher level exposures (4 of 7 individuals for moderate, and 8 of 11 individuals for high RL conditions). The most common responses we observed included longer IDDIs, increases in the number of shorter dives prior to the next deep dive, and increases in deep dive duration during exposure. In the highest exposure category, deep dives were about 29 min longer during exposure and IDDIs following exposure were about 187 min longer. Individuals exposed to MFAS above 100 dB SPL were more

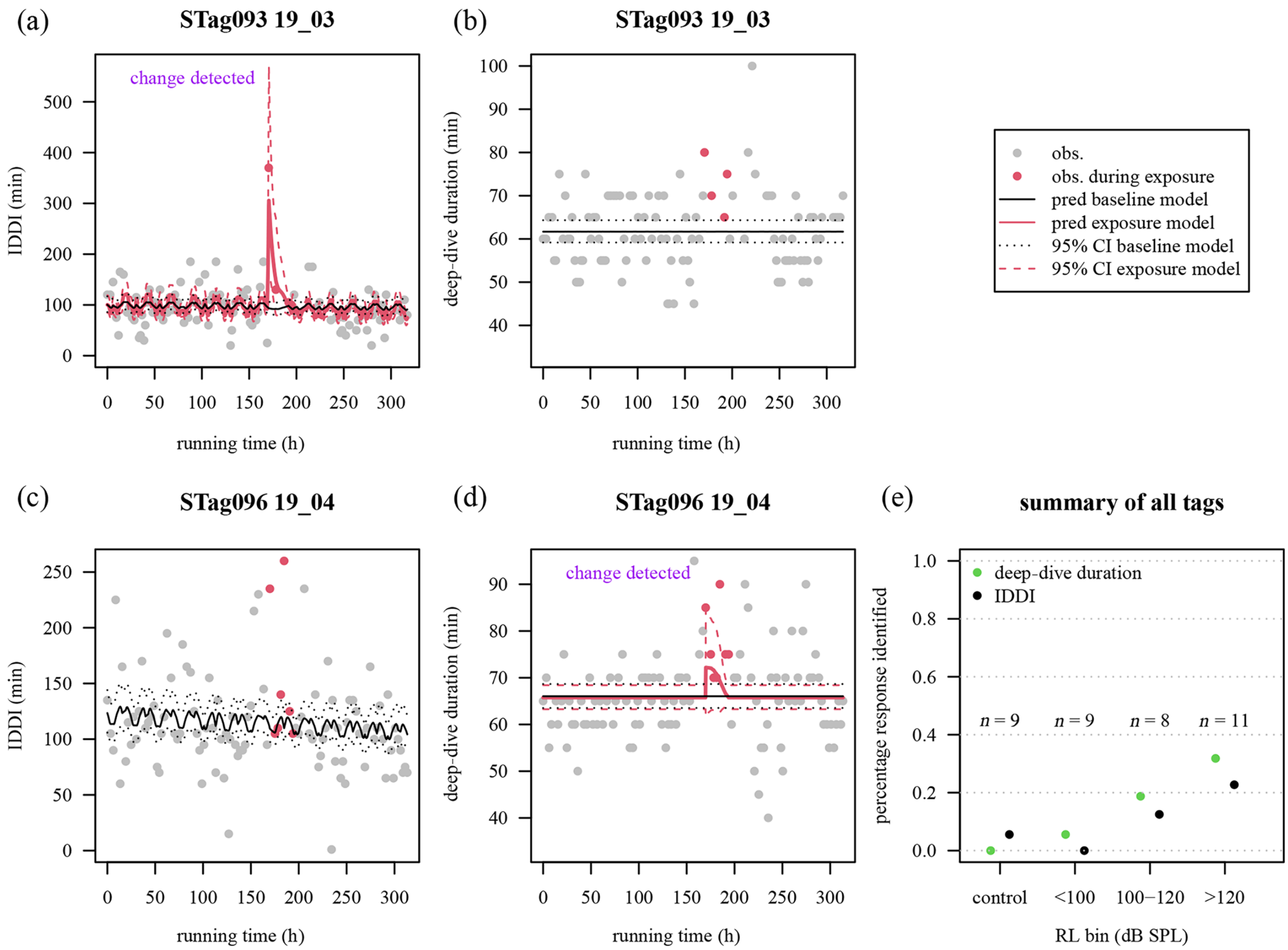

FIGURE 5 (a, c) Inter-deep-dive interval (IDDI) duration (for STag093 in mid-frequency active sonar [MFAS] exposure trial 19_03 and STag096 in MFAS exposure trial 19_04) and (b, d) deep-dive duration (for same respective individuals and exposure trials) plots with estimated smooths. Plots are from the baseline model (1) including only running time and time of day smooths and (2) including a variable for exposure when that was selected as the best model. Purple annotation indicates if the best model (selected by Akaike information criterion [AIC]) included an exposure term, which is interpreted as a change associated with exposure trials. (e) A summary of the percentage of individuals for which changes were detected by dive type and received level (RL) bin (no-MFAS control, MFAS exposures <100, 100–120, >120 dB re: 1μPa sound pressure level (SPL)). Observed (obs.) actual values are indicated before (gray), during (red), and after (gray) exposure trials.

likely to alter diving behavior by extending deep dives during exposure and forestalling additional deep dives (see Appendix S1: Section S2: Figure S124).

We also observed variation in both baseline behavior and response for whales with DTags. Nevertheless, all three whales with DTags illustrated clear changes in dive behavior and cessation of echolocation during and after MFAS exposure. One of these whales did not appear to exhibit a horizontal avoidance response, but the other two demonstrated sustained directed avoidance for hours following exposure (see especially Figure 7), sustained by intensive swimming (see: Figure 8b,c).

Finally, the multi-scale perspective (see: Figure 10), allowed us to demonstrate instantaneous responses to exposure in DTag19_218, and an unusual occurrence of directed horizontal avoidance and cessation of deep diving (STag_093) for animals in the same social group. These observations, while descriptive, were consistent across many exposure events here and in previous experiments with beaked whales and MFAS (DeRuiter et al., 2013; Stimpert et al., 2014; Tyack et al., 2011). They are also consistent with records of satellite-tagged *Ziphius* in observational studies of full-scale MFAS exposure (Falcone et al., 2017).

Kernel density estimates of conditional distributions for dive metrics suggested low levels of behavioral changes in control and low exposure conditions, compared to the response for medium and high RLs (see Figure 3, Table 3). During control and low RL exposures,

TABLE 3 By exposure-event summary of all dive behavior and avoidance analyses.

				Dive behavior analysis				Avoidance analysis
MFAS RL bin	Subject ID	Expos. trial no.	MFAS RL (dB SPL)	Quantile scoring	Kernel density assessment	Dive time GAM	Surface IDDI GAM	24-h window
Control	STag069	18_05	na	None	na	na	na	No
Control	STag070	18_05	na	na	na	na	na	No
Control	STag071	18_07	na	na	na	na	na	No
Control	STag072	18_07	na	na	na	na	na	No
Control	STag073	18_07	na	None	Mod	No	No	**Yes**
Control	STag075	18_07	na	None	Limited	No	No	No
Control	STag076	18_07	na	None	na	No	No	**Yes**
Control	STag077	18_07	na	na	na	na	na	No
Control	STag078	18_07	na	None	na	No	No	No
Control	STag079	18_07	na	na	na	na	na	**Yes**
Control	STag080	18_07	na	None	None	No	No	No
Control	STag081	18_07	na	na	na	na	na	No
Control	STag129	22_01	na	na	na	na	na	No
Control	STag130	22_02	na	na	na	na	na	No
Control	STag131	22_02	na	None	**Strong**	No	No	**Yes**
Control	STag132	22_02	na	Limited	Limited	No	**Yes**	No
Control	STag133	22_02	na	None	na	No	No	**Yes**
Control	STag134	22_02	na	Mod	na	No	No	**Yes**
<100	STag106	20_03	71.0	na	na	na	na	No
<100	STag076	18_08	75.8	Limited	na	**Yes**	No	No
<100	STag080	18_08	76.3	na	na	na	na	No
<100	STag077	18_08	77.1	na	na	na	na	**Yes**
<100	STag073	18_08	78.4	na	na	na	na	No
<100	STag103	20_03	78.5	Mod	Limited	No	No	**Yes**
<100	STag081	18_08	82.8	na	na	na	na	**Yes**
<100	STag102	20_03	82.8	None	na	No	No	**Yes**
<100	STag107	20_03	84.1	None	None	No	No	No
<100	STag071	18_08	84.4	na	na	na	na	No
<100	STag070	18_03	87.6	na	na	No	No	No
<100	STag111	20_03	89.3	None	None	No	No	No
<100	STag101	20_03	90.5	None	None	No	No	**Yes**
<100	STag079	18_08	90.9	Mod	na	No	No	No
<100	STag060	17_01	95.5	Limited	na	No	No	No
<100	STag091	19_03	95.5	na	na	na	na	No
<100	STag098	20_03	95.7	na	na	na	na	No
<100	STag082	19_02	98.9	na	na	na	na	No
100–120	STag069	18_02	101.8	Limited	Mod	**Yes**	No	No
100–120	STag092	19_04	103.5	na	na	na	na	**Yes**
100–120	STag078	18_08	106.2	na	na	No	**Yes**	**Yes**

(Continues)

TABLE 3 (Continued)

MFAS RL bin	Subject ID	Expos. trial no.	MFAS RL (dB SPL)	Dive behavior analysis: Quantile scoring	Dive behavior analysis: Kernel density assessment	Dive behavior analysis: Dive time GAM	Dive behavior analysis: Surface IDDI GAM	Avoidance analysis: 24-h window
100–120	STag088	19_02	108.2	**Strong**	Mod	No	No	**Yes**
100–120	STag082	19_01	112.4	na	na	na	na	No
100–120	STag084	19_02	112.6	na	na	na	na	No
100–120	STag061	17_01	115.7	None	na	No	No	**Yes**
100–120	STag070	18_02	116.5	na	na	na	na	**Yes**
100–120	STag085	19_02	116.5	None	Limited	No	No	**Yes**
100–120	STag093	19_04	116.8	na	na	na	na	**Yes**
100–120	STag083	19_01	117.4	**Strong**	Limited	No	No	No
100–120	STag063	17_01	118.5	Mod	na	**Yes**	No	No
100–120	STag090	19_03	119.1	na	na	na	na	**Yes**
100–120	STag087	19_02	119.3	**Strong**	**Strong**	**Yes**	**Yes**	No
>120	STag092	19_03	122.8	na	na	na	na	**Yes**
>120	STag105	20_03	123.6	None	Mod	No	No	No
>120	STag083	19_02	123.8	na	na	na	na	No
>120	STag097	19_04	124.0	**Strong**	**Strong**	**Yes**	No	No
>120	STag086	19_02	124.1	na	na	na	na	No
>120	STag108	20_03	126.3	na	na	na	na	**Yes**
>120	STag064	17_01	128.9	Limited	na	**Yes**	**Yes**	**Yes**
>120	STag110	20_03	133.7	**Strong**	**Strong**	**Yes**	**Yes**	**Yes**
>120	STag096	19_04	133.8	Mod	**Strong**	**Yes**	No	**Yes**
>120	STag089	19_02	133.9	**Strong**	**Strong**	**Yes**	No	**Yes**
>120	STag095	19_04	134.3	Mod	Mod	No	No	No
>120	STag093	19_03	134.8	**Strong**	**Strong**	No	**Yes**	**Yes**
>120	STag062	17_01	136.3	Mod	na	**Yes**	**Yes**	**Yes**
>120	STag065	17_01	136.3	Mod	na	No	**Yes**	**Yes**
>120	STag069	18_03	141.6	None	None	**Yes**	No	**Yes**
>120	DTag17_234	17_01	131.4	SmoothSDE:: **Strong**	na	na	na	na
>120	DTag19_218	19_03	136.4	SmoothSDE:: **Strong**	na	na	na	na
>120	DTag20_232	20_03	139.2	SmoothSDE:: **Strong**	na	na	na	na

Note: Strength of evidence for quantile scoring and kernel density assessments is given categorically as defined in the text (none, limited, moderate [mod], strong; with increasingly dark shading) as defined in the text. Model-based detected changes for diving behavior (GAMs) and spatial avoidance analyses are given as a binary yes/no outcome. DTag analyses with smoothSDEs are presented using a similar strength of evidence assessment.
Abbreviations: GAM, generalized additive model; IDDI, inter-deep-dive interval; MFAS, mid-frequency active sonar; na, not applicable; RL, received level; SDEs, stochastic differential equations; SPL, sound pressure level in dB re: 1μPa.

there was little consistency in which behavioral variables differed from baseline, whereas at medium and high RLs, there was a consistent pattern of longer IDDIs following MFAS exposures.

Quantitative evaluations of diving behavior using GAMs for IDDI and deep-dive duration revealed similar patterns of behavioral changes following MFAS exposure (Figure 5). Individuals in control and low RL exposure trials

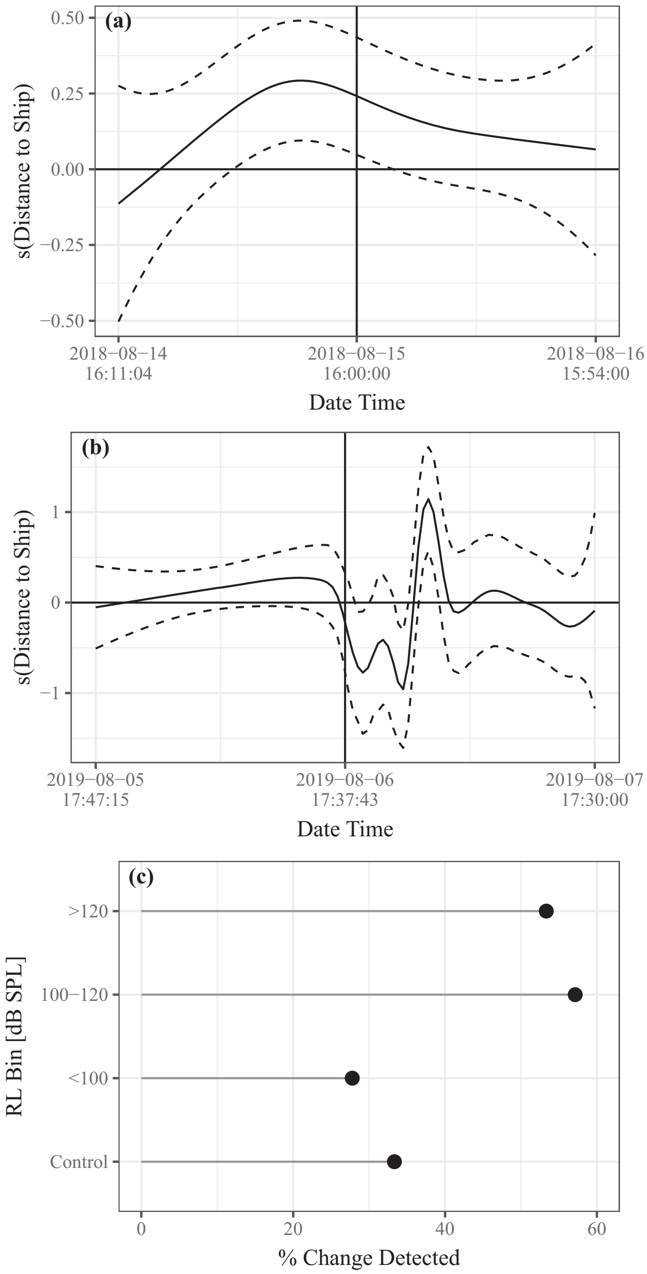

FIGURE 6 Plots of the time-varying smooth for the effect of the distance from the animal to the source ship (s(Distance to Ship)) for two animals: (a) STag072 in no-mid-frequency active sonar (MFAS) control exposure 18_07 and panel (b) STag093 in MFAS exposure trial 19_03. Each smooth is plotted over a 48-h period. The dashed black line is the 95% SE for the smooth. The vertical line indicates the start of the respective exposure trial. STag093 exhibited clear avoidance following the MFAS exposure trial (median received level [RL] was 134 dB SPL), whereas STag072 did not during the no-MFAS control. (c) The percentage of exposure trials with detected horizontal avoidance for all individuals within each RL bin (no-MFAS control, MFAS exposures <100, 100–120, >120 dB re: 1μPa sound pressure level (SPL)).

had a very low probability of changes in either dive metric. Individuals in medium and high RL exposure conditions had an increasing probability of changes in both metrics. For both medium and high RL categories, there was a slightly higher probability of longer deep dives than longer IDDIs. The proportion of individuals with detected changes in IDDI and deep-dive duration was lower than in other evaluations of diving behavior. However, this is likely a function of our limited power to detect a response, given assumptions in the GAM formulations and the goal of identifying a binary change metric in a single exposed dive. Nevertheless, these results underscore that the probability of a change in diving behavior (specifically longer IDDIs and dive durations) increased with increasing RL.

We evaluated the probability of horizontal avoidance using time-varying smooths relating distance to the source before and following exposure. Approximately 30% of individuals in both control and low RL MFAS conditions exhibited spatial avoidance of the sound source, whereas 50% or more individuals exhibited spatial avoidance in medium and high RL exposure conditions (Figure 6). This increasing probability of horizontal avoidance response with increased RLs was evident for many of the focal *Ziphius* with higher RLs, observed directly in the field (e.g., STag093 in 19_03; see Figure 7).

DISCUSSION

We present experimental results documenting the behavioral response to simulated MFAS signals for 53 *Ziphius* in 72 exposure events. Most animals exposed above 120 dB SPL demonstrated clear changes in diving and/or spatial avoidance of the source. This large sample size substantially expands our knowledge of this important issue. It also underscores the variation that exists in baseline behavior and, importantly, in behavioral responses to MFAS. We observed behavioral changes in some control (no-MFAS) trials and low RL MFAS exposures, suggesting that some individuals may have exhibited a response to either the presence of our research vessels or to some other unobserved factor. Importantly, however, we documented an increasing probability and severity of response as RLs increased.

Many individuals in the moderate and high RL conditions exhibited longer deep dives, followed by longer IDDIs, together with directed spatial avoidance of the source. All three DTag individuals ceased foraging behavior for hours following MFAS exposure. Overall, these findings are consistent with initial experimental observations of similar responses in another beaked whale species to both MFAS exposure and killer whale signals (Tyack et al., 2011) and experimental (DeRuiter et al., 2013) and observational (Falcone et al., 2017) studies with *Ziphius*. Our results add to the growing body of evidence supporting the risk-disturbance hypothesis (see Harris et al., 2018; Miller et al., 2022), which suggests

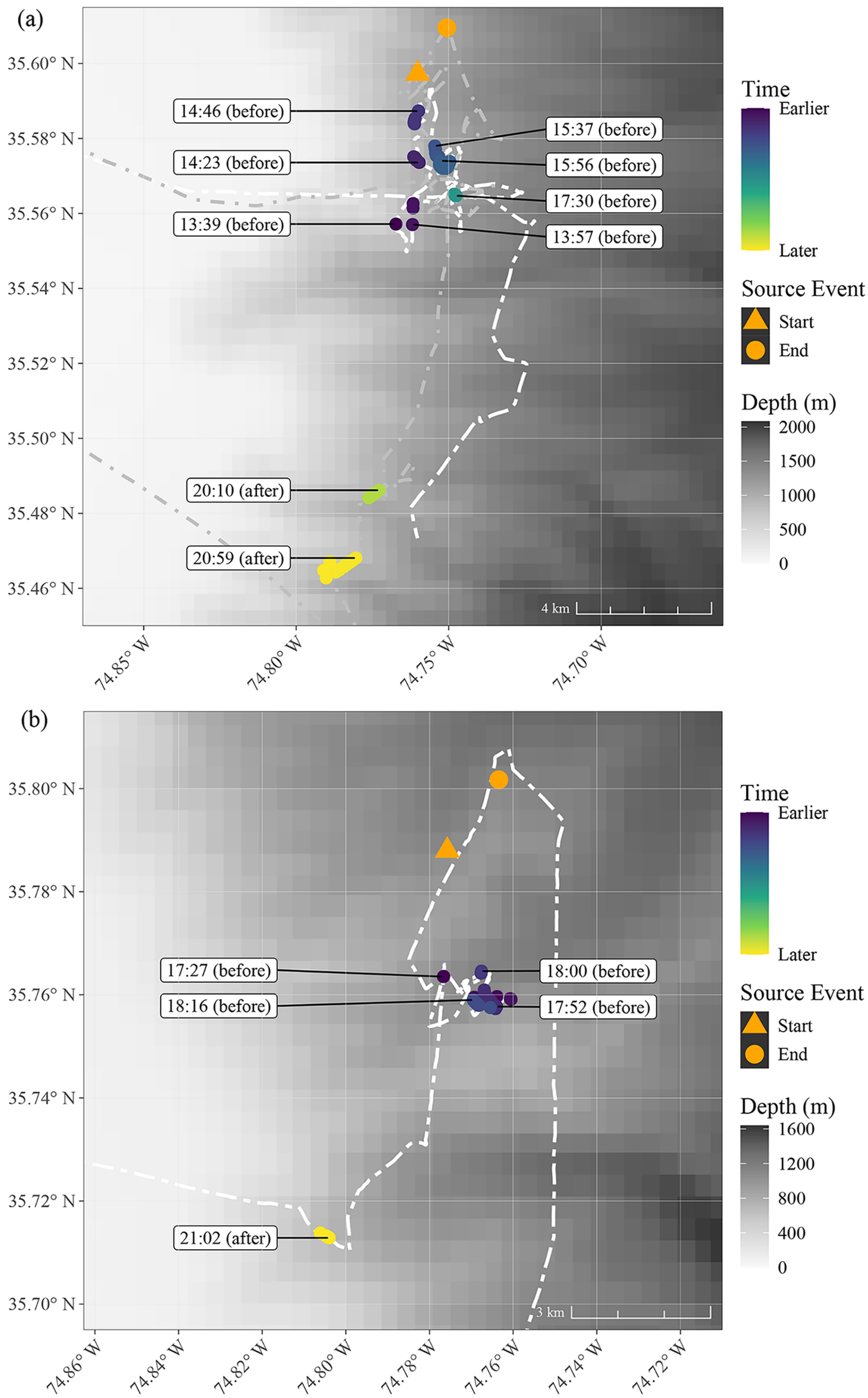

FIGURE 7 Time-annotated maps of (a) STag093 (focal) in the same group with DTag19_218 in mid-frequency active sonar (MFAS) exposure trial 19_03 and (b) STag096 (focal) in the same group with STag095, STag097 in MFAS exposure trial 19_04 showing focal follow positions before and after MFAS exposures. White dashed lines show the track of the primary follow boat. Gray dashed lines indicate track lines of a secondary follow boat. Points show position of boat(s) when photos were taken of the focal animal (or its group) during surfacings. Cooler colors denote surfacings prior to exposure trial; warmer colors denote after exposure trials. The start and end location of the vessel emitting simulated MFAS for the sound exposure trial are denoted by an orange triangle and circle, respectively.

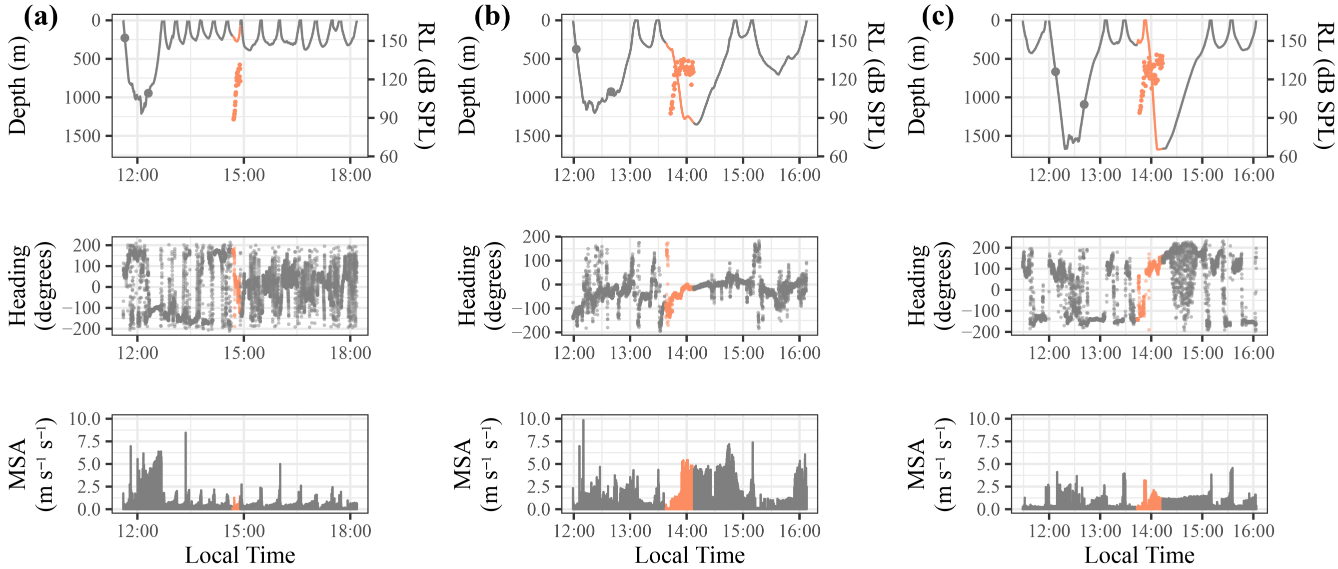

FIGURE 8 DTag records for *Ziphius* tagged during mid-frequency active sonar (MFAS) exposure trials, including (a) DTag17_234 in exposure trial 17_01, (b) DTag19_218 in exposure 19_03, and (c) DTag20_232 in exposure 20_03. Dive profiles (top row) are shown with exposure periods (highlighted in orange), start and end times of echolocation clicks (gray dots), and received level (RL) measurements (orange dots). Heading (middle row) and minimum specific acceleration (MSA; bottom row) are also shown; MFAS exposures for both are indicated with orange highlighting. SPL, sound pressure level in dB re: 1μPa.

that animals perceive and respond to human disturbance in the same manner as they respond to a predator (Frid & Dill, 2002). Nevertheless, most individuals appeared to return to baseline conditions relatively quickly following exposure (see Figure 10; all dive profiles available in Appendix S1: Section S2: Figures S48–S86). We observed changes in diving and spatial avoidance that lasted on the order of hours, but we did not witness any long-term abandonment of core habitat (Foley et al., 2021; Shearer et al., 2019).

We conducted our experiments in a region where actual MFAS is used only sporadically during relatively infrequent Navy training exercises. Thus, the context of our exposures was very different from that of previous studies with beaked whales conducted on Navy ranges (e.g., Curtis et al., 2021; DeRuiter et al., 2013; Falcone et al., 2017; Joyce et al., 2020; Moretti et al., 2014; Stimpert et al., 2014; Tyack et al., 2011) where animals are exposed to MFAS regularly, or studies conducted in areas where MFAS does not occur (Miller et al., 2015; Wensveen et al., 2019). Studies on Navy ranges, where operational sonar is used intensively, may not fully represent the response of beaked whales to MFAS because repeated exposure to these signals may dampen both the probability and intensity of response. Behavioral tolerance by animals in regions of intensive sonar use may change the nature of either behavioral response type or response probability, potentially making measurements on ranges less representative of the responses of animals in areas where MFAS occurs less frequently.

Interestingly, the overall patterns of changes in diving (notably increased IDDIs, dive duration) presented here are similar to those observed in *Ziphius* and other beaked whale species on Navy ranges (DeRuiter et al., 2013; Falcone et al., 2017; Stimpert et al., 2014; Tyack et al., 2011). Some similarities in spatial avoidance are also evident in our results relative to observational studies of beaked whales on ranges (Joyce et al., 2020; Moretti et al., 2014). These similarities provide further support for the generalized nature of response in the context of cryptic behavior and antipredator responses discussed above. When beaked whales respond to disturbance, there appears to be a consistent generalized pattern that is like how they respond to the presence of a predator.

Our results suggest a low probability of response to exposure below 100 dB SPL with an increasing probability of changes in diving and spatial avoidance above 100 and, especially, above 120 dB SPL. However, this appears to differ in interesting ways from previous studies across different beaked whale species in regions with differing levels of MFAS exposure. For instance, individuals from another beaked whale species in habitats with no MFAS, responded consistently and strongly to any signals with RLs above ambient noise levels (Miller et al., 2015; Wensveen et al., 2019). We observed less sensitivity for *Ziphius* to such low RLs. However, the responses we

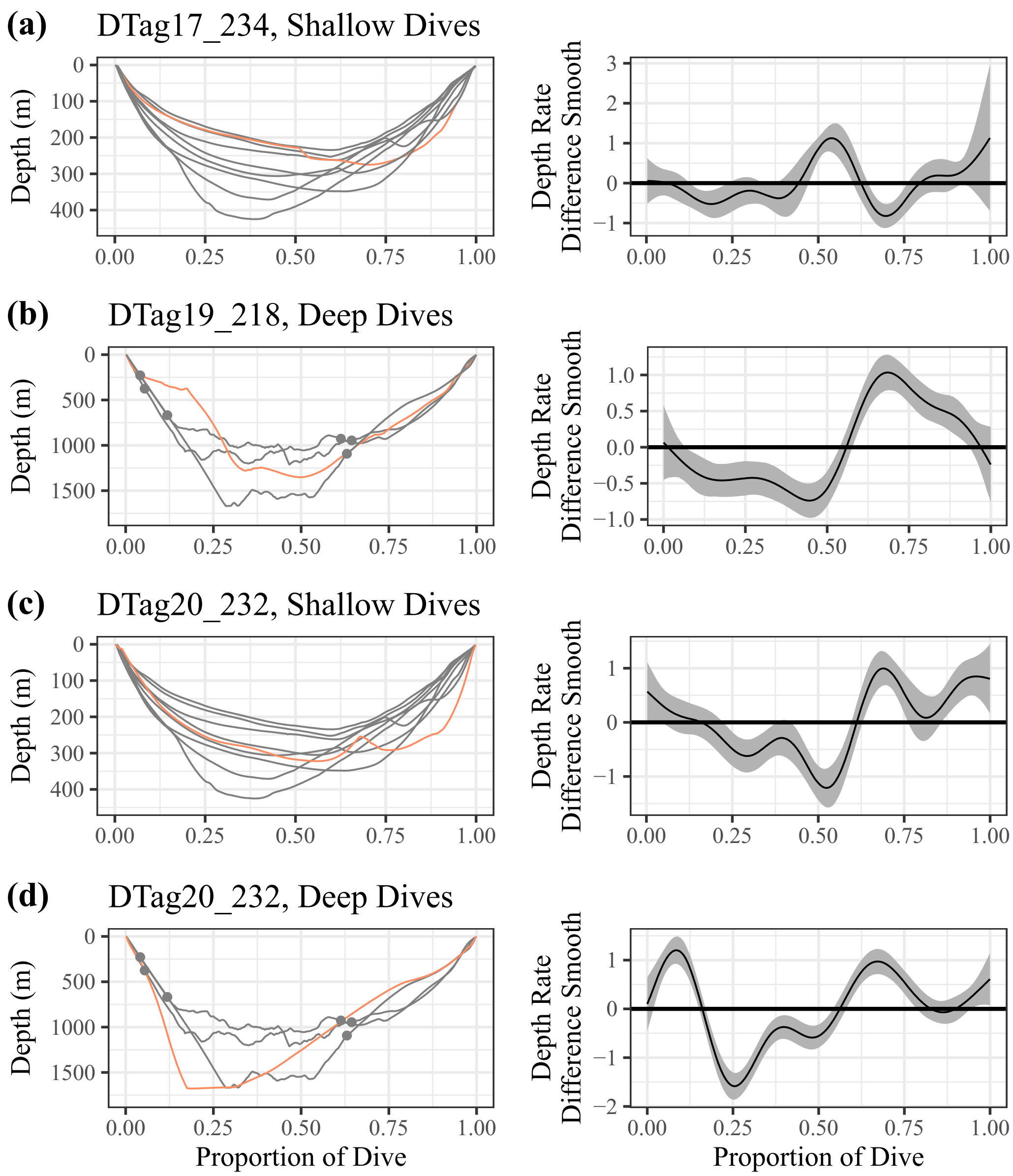

FIGURE 9 Baseline (gray) and mid-frequency active sonar (MFAS)-exposure (orange) dive profiles for three focal whales: (a) DTag17_234, (b) DTag19_234, and (c, d) DTag20_232 separated into respective shallow and deep dives (left panels). Dots indicate the start and end times of echolocation click production (if any; none occurred for any individual during exposure dives). Shifts in the volatility of depth rates during exposure are summarized with stochastic differential equation (SDE) output smooths (right panels).

observed differed in the opposite manner from the reduced likelihood of responses at higher RLs (>120 dB SPL) for other species of beaked whales from studies conducted with operational Navy MFAS on active sonar ranges (Joyce et al., 2020; Moretti et al., 2014). The effect of prior exposure on the likelihood of response remains an open question that has substantial implications for the conservation and management of beaked whales.

Our multi-scale approach provided response data over various temporal and spatial scales and across a wide range of RLs and ranges. This enabled us to view baseline behavior and exposure-response data through different lenses that enhanced the overall interpretation of response. However, this approach also came with limitations. A substantial challenge was the coarse nature of the dive data and the error inherent in spatial positions derived from satellite-transmitting dive tags. To meet these challenges, we adaptively programmed the tags (Cioffi et al., 2023), modified field approaches to obtain telemetered data directly, and developed methods to estimate RLs with precision (Schick et al., 2019, 2024). Overall, we conclude that these tags provide sufficient resolution to evaluate both avoidance and changes in diving behavior and that the coarser resolution and

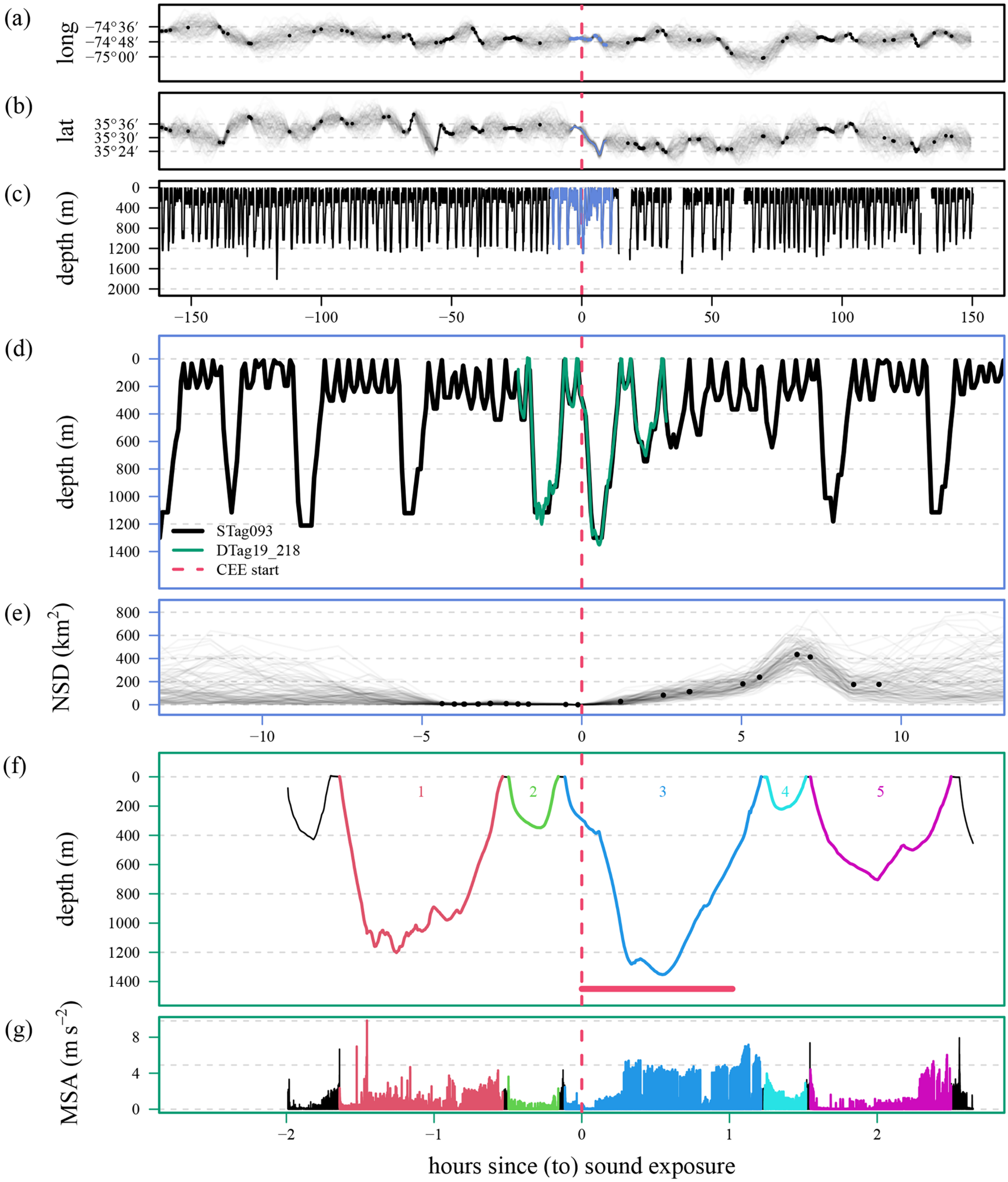

FIGURE 10 Multi-scale data presentation of STag093 and DTag19_218 deployed on animals in the same group at the time of mid-frequency active sonar (MFAS) exposure trial 19_03. Three nested scales are demonstrated centered on the initiation of the exposure (dotted vertical line): (a, b, c) 13 days of satellite tag data outlined in black, (d, e) an enlargement of 20 h of satellite tag data outlined in blue; blue highlighted in (a, b, c) indicate extent on previous scale; (f, g) 4.5 h of DTag dive data outlined in green in (d). Longitude and latitude were modeled from Argos positions and used to calculate net squared displacement (NSD) in reference to an origin at the point nearest in time to the start of the exposure trial. Gray lines in (a, b, e) indicate 100 imputed tracks derived from the crawl movement model. Black dots represent mean estimates at times when positional data was received. Minimum specific acceleration (MSA) is shown in panel (g).

relatively higher error are tolerable given the large sample sizes obtained here and the programming and analytical approaches applied. Finally, while there is considerable additional work yet to be done in assessing the effects of exposure on social groupings of beaked whales, we were able to begin to consider these consequential aspects of responses to MFAS.

This study represents the first step in quantifying the response of *Ziphius* to MFAS off Cape Hatteras. In ongoing research, we are extending CEEs to operational SQS-53C tactical MFAS produced in a controlled manner by coordinating directly with operational U.S. Navy surface vessels. Behavioral response results from these CEEs with operational sonars will yield key insights into contextual aspects of the probability of response (e.g., habituation, tolerance, response probability for different source range and RL combinations), as well as possible contextual issues related to how these sources are operated (stationary vs. mobile in realistic conditions for operational vessels). We anticipate that a comparative assessment of those results and those presented here will be integrated into risk function(s) relating exposure magnitude (e.g., RL) and response probability. Future work will include further evolutions of spatial avoidance analysis building from Hanks et al. (2015) and adaptive approaches to improve resolution and reduce error for positional and RL data (Schick et al., 2019, 2024). We are also currently investigating the effects of other MFAS signal characteristics, including similar waveforms but with continuous as opposed to pulsed transmissions. Finally, we are extending and enhancing field efforts to evaluate the potential impacts of MFAS on *Ziphius* social behavior, including additional effort to sample baseline focal group composition on consecutive days before, during, and after exposure trials to calculate fission-fusion rates, better understand demographics (e.g., age class, sex from biopsy samples) and evaluate the effects of MFAS exposure on specific age and sex classes, with a specific focus on reproductive females.

The responses of *Ziphius* to simulated MFAS signals documented in this study have implications for the management and conservation of this species. These data have direct application in exposure-response assessments and response probability functions that are increasingly required by regulatory agencies and used by navies for compliance and MFAS impact assessment. Furthermore, both the baseline and behavioral response data we report here have direct applications to inform population consequences of disturbance assessments (e.g., temporary cessation of foraging, longer IDDI) and in understanding the nature of previous MFAS-associated strandings in beaked whales.

AUTHOR CONTRIBUTIONS

B. L. Southall: Conceptualization; funding acquisition; investigation; methodology; analysis; project administration; supervision; validation; visualization; data curation; writing—original draft; writing—review and editing. **R. S. Schick:** Investigation; methodology; analysis; supervision; validation; visualization; data curation; writing—original draft; writing—review and editing. **W. R. Cioffi:** Investigation; methodology; analysis; supervision; validation; visualization; data curation; writing—original draft; writing—review and editing. **S. L. DeRuiter:** Methodology; analysis; validation; visualization; data curation; writing—original draft; writing—review and editing. **H. J. Foley:** Invistigation; analysis; writing—review and editing. **C. M. Harris:** Conceptualization; analysis; project administration; supervision; validation; writing—review and editing. **A. E. Harshbarger:** Investigation; analysis; visualization; writing—original draft; writing—review and editing. **J. E. Joseph:** Methodology; supervision; writing—review and editing. **T. Margolina:** Methodology; analysis; validation; visualization; data curation; writing—review and editing. **D. P. Nowacek:** Conceptualization; funding acquisition; investigation; methodology; analysis; supervision; validation; writing—review and editing. **N. J. Quick:** Investigation; methodology; analysis; supervision; validation; visualization; writing—review and editing. **Z. T. Swaim:** Investigation; methodology; validation; writing—review and editing. **L. Thomas:** Conceptualization; methodology; analysis; validation; data curation; writing—original draft; writing—review and editing. **D. M. Waples:** Investigation; analysis; validation; data curation; writing—review and editing. **D. L. Webster:** Investigation; methodology; analysis; validation; writing—review and editing. **J. H. Wisse:** Investigation; analysis; writing—review and editing. **A. J. Read:** Conceptualization; funding acquisition; investigation; methodology; project administration; supervision; validation; writing—original draft; writing—review and editing.

ACKNOWLEDGMENTS

This project would not have been possible without the hard work of many dedicated field personnel, students, analysts, and other direct collaborators. These include: Ari Friedlaender, Matt Bowers, Alex Shorter, Robin Baird, Will Sloger, Mark Wilson, Dave Moretti, Dave Anderson, John Hildebrand, Kait Frazier, Jen Dunn, Ron Morrisey, Stephanie Watwood, Megan McKenna, Kate Sutherland, and Danielle Alvarez. We also acknowledge additional technical and statistical support from: Kristin Southall, Théo Michelot, Richard Glennie, David Miller, Theoni Photopoulou, Saana Isojunno, Phillippe Bouchet, David Haas, Greg Schorr, Kenady Wilson, Matthew

Rutishauser, and Heather Baer. We appreciate the support from the Duke University Marine Laboratory Marine Operations and *R/V Shearwater* crew, including John Wilson, Tina Thomas, Mathew Dawson, Rachel Dudas, Brantley Acree, and Mikey Diehl. We acknowledge the following captains of charter fishing vessels that supported the project: Harold "Smitty" Smith (*F/V Spray*) Reed Meredith (*F/V Kahuna*), Cullen Malarney (*F/V Kahuna*), James "Jimmie" Jr. Horning (*F/V Hog Wild*), Krage Gardiner (*F/V Fin Galley*), and H. Drexel "Stormy" Harrington (*F/V Captain Tiki XIV*). Funding was provided by the US Fleet Forces Command Marine Species Monitoring Program through the Naval Facilities Engineering Command Atlantic under contract nos. N62470-10-D-3011 and N62470-15-D-8006 issued to HDR, Inc. We appreciate the logistical support and collaboration from key individuals in these entities including Joel Bell, Ron Filipowicz, Jessica Aschettino, Dan Engelhaupt, Jackie Bort, Christy Cowan, Laura Busch, Joe Atangan, Jene Nissen, and many others. This project also substantially and directly leveraged methods from related studies supported through the US Navy's Living Marine Resources (LMR) Program and the Office of Naval Research, Marine Mammals and Biology Program. Research was conducted under National Marine Fisheries Service scientific research permit numbers 17086 and 20605 to Robin W. Baird, 14809 and 22156 to Douglas P. Nowacek, and 16239 to Daniel T. Engelhaupt and National Marine Fisheries Service general authorization letter of confirmation numbers 19903 and 25471 to Andrew J. Read. All activities were approved by the Duke University Institutional Animal Care and Use Committee.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Data and scripts to reproduce analyses and figures for noise exposure and behavioral response (Southall et al., 2025) are available in a github repository: https://github.com/atlanticbrs/zc_sim_source_resp and archived with Zenodo: https://doi.org/10.5281/zenodo.17640058.

REFERENCES

Aguilar Soto, N., M. P. Johnson, P. T. Madsen, P. L. Tyack, A. Bocconcelli, and J. F. Borsani. 2006. "Does Intense Ship Noise Disrupt Foraging in Deep-Diving Cuvier's Beaked Whales (*Ziphius cavirostris*)?" *Marine Mammal Science* 22: 690–99. https://doi.org/10.1111/j.1748-7692.2006.00044.x.

Andrews, R. D., R. L. Pitman, and L. T. Ballance. 2008. "Satellite Tracking Reveals Distinct Movement Patterns for Type B and Type C Killer Whales in the Southern Ross Sea, Antarctica." *Polar Biology* 31: 1461–68.

Bernaldo de Quirós, Y., A. Fernandez, R. W. Baird, R. L. Brownell, Jr., N. Aguilar de Soto, D. Allen, M. Arbelo, et al. 2019. "Advances in Research on the Impacts of Anti-Submarine Sonar on Beaked Whales." *Proceedings of the Royal Society B* 286(1895): 20182533.

Borroni, A. 2024. "Exposure and Behavioural Response of Satellite-Tracked Cuvier's Beaked Whales (*Ziphius cavirostris*) to Naval Traffic in the Western Mediterranean Sea." PhD thesis, University of Genova. https://hdl.handle.net/11567/1168596.

Cholewiak, D., A. I. DeAngelis, D. Palka, P. J. Corkeron, and S. M. Van Parijs. 2017. "Beaked Whales Demonstrate a Marked Acoustic Response to the Use of Shipboard Echosounders." *Royal Society Open Science* 4(12): 170940.

Cioffi, W. R., N. J. Quick, H. J. Foley, D. M. Waples, Z. T. Swaim, J. M. Shearer, D. L. Webster, et al. 2021. "Adult Male Cuvier's Beaked Whales (*Ziphius cavirostris*) Engage in Prolonged Bouts of Synchronous Diving." *Marine Mammal Science* 37(3): 1085–1100.

Cioffi, W. R., N. J. Quick, Z. T. Swaim, H. J. Foley, D. M. Waples, D. L. Webster, R. W. Baird, B. L. Southall, D. P. Nowacek, and A. J. Read. 2023. "Trade-Offs in Telemetry Tag Programming for Deep-Diving Cetaceans: Data Longevity, Resolution, and Continuity." *Animal Biotelemetry* 11(1): 23.

Curtis, K. A., E. A. Falcone, G. S. Schorr, J. E. Moore, D. J. Moretti, J. Barlow, and E. Keene. 2021. "Abundance, Survival, and Annual Rate of Change of Cuvier's Beaked Whales (*Ziphius cavirostris*) on a Navy Sonar Range." *Marine Mammal Science* 37(2): 399–419.

D'Amico, A., R. C. Gisiner, D. R. Ketten, J. A. Hammock, C. Johnson, P. L. Tyack, and J. Mead. 2009. "Beaked Whale Strandings and Naval Exercises." *Aquatic Mammals* 35(4): 452–472.

DeRuiter, S., M. Johnson, D. Sweeney, Y. McNamara-Oh, S. Fynewever, O. Tejevbo, T. Marques, Y. Wang, and O. Ogedegbe. 2023. "Tagtools: Work with Data from High-Resolution Biologging Tags." R Package Version 0.1.0. https://CRAN.R-project.org/package=tagtools.

DeRuiter, S. L., R. Langrock, T. Skirbutas, J. A. Goldbogen, J. Calambokidis, A. S. Friedlaender, and B. L. Southall. 2017. "A Multivariate Mixed Hidden Markov Model to Analyze Blue Whale Diving Behaviour during Controlled Sound Exposures." *Annals of Applied Statistics* 11: 362–392. https://doi.org/10.1214/16-AOAS1008.

DeRuiter, S. L., B. L. Southall, J. Calambokidis, W. M. X. Zimmer, D. Sadykova, E. A. Falcone, A. S. Friedlaender, et al. 2013. "First Direct Measurements of Behavioural Responses by Cuvier's Beaked Whales to Mid-Frequency Active Sonar." *Biology Letters* 9: 20130223. https://doi.org/10.1098/rsbl.2013.0223.

Duarte, C. M., L. Chapuis, S. P. Collin, D. P. Costa, R. P. Devassy, V. M. Eguiluz, C. Erbe, et al. 2021. "The Soundscape of the Anthropocene Ocean." *Science* 371(6529): eaba4658.

Ellison, W. T., B. L. Southall, C. W. Clark, and A. S. Frankel. 2012. "A New Context-Based Approach to Assess Marine Mammal Behavioral Responses to Anthropogenic Sounds." *Conservation Biology* 26(1): 21–28.

Ellison, W. T., B. L. Southall, A. S. Frankel, K. Vigness-Raposa, and C. W. Clark. 2018. "An Acoustic Scene Perspective on Spatial, Temporal, and Spectral Aspects of Marine Mammal Behavioral Responses to Noise." *Aquatic Mammals* 44(3): 239–243.

Evans, D. L., and G. R. England. 2001. *Joint Interim Report, Bahamas Marine Mammal Stranding Event of 15–16 March*

2000. Washington, DC: U.S. Department of Commerce and Secretary of the Navy.

Faerber, M. M., and R. W. Baird. 2010. "Does a Lack of Observed Beaked Whale Strandings in Military Exercise Areas Mean no Impacts Have Occurred? A Comparison of Stranding and Detection Probabilities in the Canary and Main Hawaiian Islands." *Marine Mammal Science* 26(3): 602–613.

Falcone, E. A., G. S. Schorr, S. L. Watwood, S. L. DeRuiter, A. N. Zerbini, R. D. Andrews, and D. J. Moretti. 2017. "Diving Behaviour of Cuvier's Beaked Whales Exposed to Two Types of Military Sonar." *Royal Society Open Science* 4(8): 170629.

Filadelfo, R., J. Mintz, E. Michlovich, A. D'Amico, P. L. Tyack, and D. R. Ketten. 2009. "Correlating Military Sonar Use with Beaked Whale Mass Strandings: What Do the Historical Data Show?" *Aquatic Mammals* 35(4): 435–444.

Foley, H. J., K. Pacifici, R. W. Baird, D. L. Webster, Z. T. Swaim, and A. J. Read. 2021. "Residency and Movement Patterns of Cuvier's Beaked Whales Ziphius Cavirostris off Cape Hatteras, North Carolina, USA." *Marine Ecology Progress Series* 660: 203–216.

Frid, A., and L. Dill. 2002. "Human-Caused Disturbance Stimuli as a Form of Predation Risk." *Conservation Ecology* 6(1): art11.

Goldbogen, J. A., B. L. Southall, S. L. DeRuiter, J. Calambokidis, A. S. Friedlaender, E. L. Hazen, E. A. Falcone, et al. 2013. "Blue Whales Respond to Simulated Mid-Frequency Military Sonar." *Proceedings of the Royal Society B: Biological Sciences* 280(1765): 20130657.

Hanks, E. M., M. B. Hooten, and M. W. Alldredge. 2015. "Continuous-Time Discrete-Space Models for Animal Movement." *Annals of Applied Statistics* 9(1): 145–165. https://doi.org/10.1214/14-AOAS803.

Harris, C. M., L. Thomas, E. A. Falcone, J. Hildebrand, D. Houser, P. H. Kvadsheim, F.-. P. A. Lam, et al. 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.

Harris, C. M., L. Thomas, D. Sadykova, S. L. DeRuiter, P. L. Tyack, B. L. Southall, A. J. Read, and P. J. Miller. 2016. "The Challenges of Analyzing Behavioral Response Study Data: An Overview of the MOCHA (Multi-Study OCean Acoustics Human Effects Analysis) Project." In *The Effects of Noise on Aquatic Life II*. edited by A. N. Popper and A. Hawkins, 399–407. New York, NY: Springer.

Hewitt, J., A. E. Gelfand, N. J. Quick, W. R. Cioffi, B. L. Southall, S. L. DeRuiter, and R. S. Schick. 2022. "Kernel Density Estimation of Conditional Distributions to Detect Responses in Satellite Tag Data." *Animal Biotelemetry* 10(1): 28.

Isojunno, S., C. Curé, P. H. Kvadsheim, F. P. A. Lam, P. L. Tyack, P. J. Wensveen, and P. J. O. M. Miller. 2016. "Sperm Whales Reduce Foraging Effort during Exposure to 1–2 kHz Sonar and Killer Whale Sounds." *Ecological Applications* 26(1): 77–93.

Johnson, M. P., and P. L. Tyack. 2003. "A Digital Acoustic Recording Tag for Measuring the Response of Wild Marine Mammals to Sound." *IEEE Journal of Oceanic Engineering* 28(1): 3–12. https://doi.org/10.1109/JOE.2002.808212.

Joyce, T. W., J. W. Durban, D. E. Claridge, C. A. Dunn, L. S. Hickmott, H. Fearnbach, K. Dolan, and D. Moretti. 2020. "Behavioral Responses of Satellite Tracked Blainville's Beaked Whales (*Mesoplodon densirostris*) to Mid-Frequency Active Sonar." *Marine Mammal Science* 36(1): 29–46.

Kootstra, C. P. 2024. "Deep Divers Reaching the Seafloor: Investigating Cuvier's Beaked Whale (*Ziphius cavirostris*) Foraging Behavior." International Master of Science in Marine Biological Resources (IMBRSea) thesis, Ghent University.

Kvadsheim, P. H., S. DeRuiter, L. D. Sivle, J. Goldbogen, R. Roland-Hansen, P. J. Miller, P. J. O. Miller, et al. 2017. "Avoidance Responses of Minke Whales to 1–4 kHz Naval Sonar." *Marine Pollution Bulletin* 121(1–2): 60–68.

Margolina, T., J. E. Joseph, and B. L. Southall. 2018. "BRS Sound Exposure Modeling Tool: A System for Planning, Visualization and Analysis." In *OCEANS 2018 MTS/IEEE Charleston*, 1–4. IEEE.

McGregor, P. K., ed. 2013. *Playback and Studies of Animal Communication*, Vol. 228. New York, NY: Springer Science & Business Media.

McLellan, W. A., R. J. McAlarney, E. W. Cummings, A. J. Read, C. G. Paxton, J. T. Bell, and D. A. Pabst. 2018. "Distribution and Abundance of Beaked Whales (Family *Ziphiidae*) off Cape Hatteras, North Carolina, USA." *Marine Mammal Science* 34(4): 997–1017.

Michelot, T., R. Glennie, C. Harris, and L. Thomas. 2021. "Varying-Coefficient Stochastic Differential Equations with Applications in Ecology." *Journal of Agricultural, Biological, and Environmental Statistics* 26: 446–463. https://doi.org/10.1007/s13253-021-00450-6.

Michelot, T., R. Glennie, L. Thomas, N. Quick, and C. M. Harris. 2023. "Continuous-Time Modelling of Behavioural Responses in Animal Movement." *The Annals of Applied Statistics* 17(4): 3570–88.

Miller, P. J., S. Isojunno, E. Siegal, F. P. A. Lam, P. H. Kvadsheim, and C. Curé. 2022. "Behavioral Responses to Predatory Sounds Predict Sensitivity of Cetaceans to Anthropogenic Noise within a Soundscape of Fear." *Proceedings of the National Academy of Sciences of the United States of America* 119(13): e2114932119.

Miller, P. J., P. H. Kvadsheim, F. P. A. Lam, P. L. Tyack, C. Curé, S. L. DeRuiter, F. P. Lam, et al. 2015. "First Indications that Northern Bottlenose Whales Are Sensitive to Behavioural Disturbance from Anthropogenic Noise." *Royal Society Open Science* 2(6): 140484.

Miller, P. J., P. H. Kvadsheim, F. P. A. Lam, P. J. Wensveen, R. Antunes, A. C. Alves, F. Visser, L. Kleivane, P. Tyack, and L. D. Sivle. 2012. "The Severity of Behavioral Changes Observed during Experimental Exposures of Killer (*Orcinus orca*), Long-Finned Pilot (*Globicephala melas*), and Sperm (*Physeter macrocephalus*) Whales to Naval Sonar." *Aquatic Mammals* 38(4): 362.

Moretti, D., L. Thomas, T. Marques, J. Harwood, A. Dilley, B. Neales, J. Shaffer, et al. 2014. "A Risk Function for Behavioral Disruption of Blainville's Beaked Whales (*Mesoplodon densirostris*) from Mid-Frequency Active Sonar." *PLoS One* 9(1): e85064.

National Academies of Sciences (NAS), Engineering. 2016. "Approaches to Understanding the Cumulative Effects of Stressors on Marine Mammals." https://doi.org/10.17226/23479.

National Research Council, and National Academies Press. 2005. *Marine Mammal Populations and Ocean Noise: Determining when Noise Causes Biologically Significant Effects*. Washington, DC: National Academies Press.

Pedersen, E. J., D. L. Miller, G. L. Simpson, and N. Ross. 2019. "Hierarchical Generalized Additive Models in Ecology: An Introduction with Mgcv." *PeerJ* 7: e6876.

Pirotta, E. 2022. "A Review of Bioenergetic Modelling for Marine Mammal Populations." *Conservation Physiology* 10(1): coac036.

Pirotta, E., R. Milor, N. Quick, D. Moretti, N. Di Marzio, P. Tyack, I. Boyd, and G. Hastie. 2012. "Vessel Noise Affects Beaked Whale

Behavior: Results of a Dedicated Acoustic Response Study." *PLoS One* 7: e42535. https://doi.org/10.1371/journal.pone.0042535.

Quick, N. J., W. R. Cioffi, J. M. Shearer, A. Fahlman, and A. J. Read. 2020. "Extreme Diving in Mammals: First Estimates of Behavioural Aerobic Dive Limits in Cuvier's Beaked Whales." *Journal of Experimental Biology* 223(18): jeb222109.

Schick, R. S., M. Bowers, S. DeRuiter, A. Friedlaender, J. Joseph, T. Margolina, D. P. Nowacek, and B. L. Southall. 2019. "Accounting for Positional Uncertainty when Modeling Received Levels for Tagged Cetaceans Exposed to Sonar." *Aquatic Mammals* 45(6): 675–690. https://doi.org/10.1578/AM.45.6.2019.675.

Schick, R. S., W. R. Cioffi, H. J. Foley, J. Joseph, N. A. Kaney, T. Margolina, Z. T. Swaim, L. Zheng, and B. L. Southall. 2024. "Estimating Received Level in Behavioral Response Studies through the Use of Ancillary Data." *The Journal of the Acoustical Society of America* 156(6): 4169–80.

Shearer, J. M., N. J. Quick, W. R. Cioffi, R. W. Baird, D. L. Webster, H. J. Foley, Z. T. Swaim, D. M. Waples, J. T. Bell, and A. J. Read. 2019. "Diving Behaviour of Cuvier's Beaked Whales (*Ziphius cavirostris*) off Cape Hatteras, North Carolina." *Royal Society Open Science* 6(2): 181728.

Southall, B. L. 2017. "Noise." In *Encyclopedia of Marine Mammals*, 3rd ed., edited by B. Würsig and H. Thiewesson, 699–707. New York: Academic Press.

Southall, B. L., A. N. Allen, J. Calambokidis, C. Casey, S. L. DeRuiter, S. Fregosi, A. S. Friedlaender, et al. 2023. "Behavioural Responses of Fin Whales to Military Mid-Frequency Active Sonar." *Royal Society Open Science* 10(12): 231775.

Southall, B. L., A. E. Bowles, W. T. Ellison, J. J. Finneran, R. L. Gentry, C. R. Greene, Jr., D. Kastak, et al. 2007. "Marine Mammal Noise Exposure Criteria: Initial Scientific Recommendations." *Aquatic Mammals* 33: 411–521.

Southall, B. L., S. L. DeRuiter, A. Friedlaender, A. K. Stimpert, J. A. Goldbogen, E. Hazen, C. Casey, et al. 2019. "Behavioral Responses of Individual Blue Whales (Balaenoptera Musculus) to Mid-Frequency Military Sonar." *Journal of Experimental Biology* 222(5): jeb190637.

Southall, B. L., J. J. Finneran, C. Reichmuth, P. E. Nachtigall, D. R. Ketten, A. E. Bowles, W. T. Ellison, D. P. Nowacek, and P. L. Tyack. 2019. "Marine Mammal Noise Exposure Criteria: Updated Scientific Recommendations for Residual Hearing Effects." *Aquatic Mammals* 45(2): 125–232.

Southall, B. L., D. Moretti, B. Abraham, J. Calambokidis, S. L. DeRuiter, and P. L. Tyack. 2012. "Marine Mammal Behavioral Response Studies in Southern California: Advances in Technology and Experimental Methods." *Marine Technology Society Journal* 46(4): 48–59.

Southall, B. L., D. P. Nowacek, A. E. Bowles, V. Senigaglia, L. Bejder, and P. L. Tyack. 2021. "Marine Mammal Noise Exposure Criteria: Assessing the Severity of Marine Mammal Behavioral Responses to Human Noise." *Aquatic Mammals* 47(5): 421–464.

Southall, B. L., D. P. Nowacek, P. J. Miller, and P. L. Tyack. 2016. "Experimental Field Studies to Measure Behavioral Responses of Cetaceans to Sonar." *Endangered Species Research* 31: 293–315.

Southall, B. L., R. S. Schick, W. R. Cioffi, S. L. DeRuiter, H. J. Foley, C. M. Harris, A. E. Harshbarger, et al. 2025. "Atlanticbrs/zc_sim_source_resp: Dataset and Code: Behavioral Responses of Goose-Beaked Whales (*Ziphius cavirostris*) to Simulated Military Sonar (v1.0-Accepted)." Zenodo. https://doi.org/10.5281/zenodo.17640059.

Stanistreet, J. E., D. P. Nowacek, S. Baumann-Pickering, J. T. Bell, D. M. Cholewiak, J. A. Hildebrand, L. E. W. Hodge, H. B. Moors-Murphy, S. M. Van Parijs, and A. J. Read. 2017. "Using Passive Acoustic Monitoring to Document the Distribution of Beaked Whale Species in the Western North Atlantic Ocean." *Canadian Journal of Fisheries and Aquatic Sciences* 74(12): 2098–2109.

Stimpert, A. K., S. L. DeRuiter, B. L. Southall, D. J. Moretti, E. A. Falcone, J. A. Goldbogen, A. Friedlaender, G. S. Schorr, and J. Calambokidis. 2014. "Acoustic and Foraging Behavior of a Baird's Beaked Whale, *Berardius bairdii*, Exposed to Simulated Sonar." *Scientific Reports* 4(1): 7031.

Tyack, P. L., W. M. Zimmer, D. Moretti, B. L. Southall, D. E. Claridge, J. W. Durban, C. W. Clark, et al. 2011. "Beaked Whales Respond to Simulated and Actual Navy Sonar." *PLoS One* 6(3): e17009.

Van Parijs, S., A. DeAngelis, and D. Cholewiak. 2024. *Analysis of Acoustic Ecology of North Atlantic Shelf Break Cetaceans and Effects of Anthropogenic Noise Impacts*. Technical Report to U.S. Navy Marine Species Monitoring Program. https://www.navymarinespeciesmonitoring.us/reading-room/project-profiles/acoustic-ecology-northwest-atlantic-shelf-break-cetaceans-and-effects-anthropogenic-noise-impacts/.

Wellard, R., K. Lightbody, L. Fouda, M. Blewitt, D. Riggs, and C. Erbe. 2016. "Killer Whale (*Orcinus orca*) Predation on Beaked Whales (*Mesoplodon* spp.) in the Bremer Sub-Basin, Western Australia." *PLoS One* 11(12): e0166670.

Wensveen, P. J., S. Isojunno, R. R. Hansen, A. M. von Benda-Beckmann, L. Kleivane, S. van IJsselmuide, F.-P. A. Lam, et al. 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.

Wiggins, S. M., and J. A. Hildebrand. 2016. "Long-Term Monitoring of Cetaceans Using Autonomous Acoustic Recording Packages." In *Listening in the Ocean* 35–59. New York, NY: Springer New York.

Wood, S. N. 2003. "Thin Plate Regression Splines." *Journal of the Royal Statistical Society, Series B: Statistical Methodology* 65(1): 95–114.

Wood, S. N. 2017. *Generalized Additive Models: An Introduction with R*. Boca Raton, FL: Chapman and Hall/CRC.

Wood, S. N., N. Pya, and B. Säfken. 2016. "Smoothing Parameter and Model Selection for General Smooth Models." *Journal of the American Statistical Association* 111(516): 1548–63.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Southall, B. L., R. S. Schick, W. R. Cioffi, S. L. DeRuiter, H. J. Foley, C. M. Harris, A. E. Harshbarger, et al. 2026. "Behavioral Responses of Goose-Beaked Whales (*Ziphius Cavirostris*) to Simulated Military Sonar." *Ecosphere* 17(1): e70501. https://doi.org/10.1002/ecs2.70501