6

https://doi.org/10.1038/s43247-024-01920-8

# A decade of declines in toothed whale densities following the Deepwater Horizon oil spill

Check for updates

Kaitlin E. Frasier <sup>®</sup><sup>1</sup> ⊠, Macey A. Kadifa<sup>1</sup>, Alba Solsona Berga <sup>®</sup><sup>1</sup>, John A. Hildebrand <sup>®</sup><sup>1</sup>, Sean M. Wiggins<sup>1</sup>, Lance P. Garrison<sup>2</sup>, Héloïse Frouin-Mouy <sup>®</sup><sup>2,3</sup>, Adolfo Gracia <sup>®</sup><sup>4</sup>, Arturo Serrano <sup>®</sup><sup>5</sup>, Lynne E. W. Hodge <sup>®</sup><sup>2,6</sup>, Carrie C. Wall<sup>7,8</sup>, Matthieu Le Hénaff <sup>®</sup><sup>9</sup> & Melissa S. Soldevilla<sup>2</sup>

Shortly after the *Deepwater Horizon* oil spill began in April 2010, a widely spaced passive acoustic monitoring array was deployed in the northeastern Gulf of Mexico to document the impacts of this unprecedentedly large and deep offshore oil spill on oceanic marine mammals. The array was subsequently maintained for over a decade. Here we document decadal density declines for seven of eight monitored species groups, including sperm whales (up to 31%), beaked whales (up to 83%), and small delphinids (up to 43%). Declines were observed both within and outside of the surface oil footprint. Though not conclusively linked to the oil spill, the broad spatial and temporal scale of these declines observed for disparate marine mammal species is consistent with *Deepwater Horizon* impacts. These declines have exceeded and outlasted post-spill damage assessment predictions, suggesting that the offshore ecosystem impacts of *Deepwater Horizon* may have been larger than previously thought.

Offshore oil spills have the potential to cause significant damage to marine ecosystems, and the 2010 Deepwater Horizon (DWH) oil spill in the Gulf of Mexico (GoMx) is the largest and deepest offshore humancaused release of hydrocarbons on record<sup>1</sup>. This spill occurred at a depth of over 1600 m, and an unprecedented step was taken to inject dispersants into the flow of oil from the wellhead<sup>2</sup>. This approach appears to have contributed extensively to the formation of a deep subsurface plume of dispersed oil particles and high methane concentrations at the 1000-1200 m depth horizon<sup>3</sup> in conjunction with the estimated 149,000 km<sup>2</sup> surface oil slick<sup>4</sup>. Over a decade later, the impacts of the spill and response on offshore oceanic GoMx ecosystems remain poorly understood<sup>5-7</sup>, due to limited pre-spill baseline measurements, challenges of offshore and deep ocean observation<sup>8</sup>, ongoing chronic impacts in the region<sup>9,10</sup>, and increasing effects of climate change on GoMx oceanography<sup>11-13</sup>. Passive acoustic monitoring using low-power autonomous systems is an effective way to observe the activity of acousticallyavailable apex predators in offshore deep-water locations and to quantify long-term changes.

The GoMx is inhabited by at least 18 species of odontocetes, or toothed whales, which rely on sounds for navigation, foraging, and communication<sup>14</sup>. Acoustic occurrence (presence or absence of detections) and densities (estimated numbers of individuals per unit area) of these species can be viewed as one indicator of ecosystem health in offshore regions<sup>15</sup>. However, these long-lived species are slow to reproduce, highly mobile, and sometimes migratory; therefore, long time series are required to determine statistically robust trends<sup>16–18</sup>.

We maintained five long-term passive acoustic monitoring stations in the GoMx from 2010 to 2020 (Fig. 1, Supplementary Table 1). Three deep sites were selected, including: Mississippi Canyon (MC) approximately 15 km northeast of the DWH wellhead within the 2010 surface oil footprint; Green Canyon (GC) a deep canyon 305 km west of the wellhead, and Dry Tortugas (DT) positioned on a sharp slope 520 km to the south of the wellhead, both outside of the surface footprint. The GC site was located within the estimated footprint of the deep plume<sup>19</sup>. Two shallow shelf sites monitored primarily shallow water species including: Main Pass (MP) located 56 km north of the wellhead, within the surface footprint, and De

<sup>1</sup>Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA, USA. <sup>2</sup>NOAA NMFS Southeast Fisheries Science Center, Miami, FL, USA. <sup>3</sup>Cooperative Institute for Marine and Atmospheric Studies, University of Miami, Miami, FL, USA. <sup>4</sup>Instituto de Ciencias del Mar y Limnología, UNAM, México City, México. <sup>5</sup>Universidad Veracruzana, Veracruz, México. <sup>6</sup>University Corporation for Atmospheric Research's (UCAR) Cooperative Programs for the Advancement of Earth System Science (CPAESS), Boulder, CO, USA. <sup>7</sup>University of Colorado Boulder, Boulder, CO, USA. <sup>8</sup>NOAA National Centers for Environmental Information, Boulder, CO, USA. <sup>9</sup>NOAA Atlantic Oceanographic and Meteorological Laboratory (AOML), Miami, FL, USA.



**Fig. 1** | **Site location map.** Orange circles indicate five sites monitored acoustically for toothed whale occurrence from 2010–2020. Site MC occupied two slightly different locations (1) and (2). The black triangle indicates the location of the *Deepwater Horizon* wellhead, with cumulative surface oil footprint<sup>73</sup> in red. Bathymetric contour depths are given in meters. Bathymetry obtained from GEBCO<sup>74</sup>.

Soto Canyon (DC) located 230 km to the east of the wellhead, just outside of the footprint.

We estimated weekly densities from acoustic detections<sup>20-24</sup> of eight of the most prevalent categories of odontocetes in the GoMx including sperm whales (Physeter macrocephalus), three species of beaked whales (Ziphius cavirostris, Mesoplodon densirostris, and Mesoplodon europeaus), pygmy and dwarf sperm whales (Kogia spp.), Risso's dolphins (Grampus griseus), and two additional delphinid categories: (1) A high frequency (HF, peak energy 20-80 kHz, inter-click interval near 0.06 s) delphinid category presumed to include bottlenose dolphins (Tursiops truncatus) and five species in the genus Stenella, and (2) A low frequency (LF, peak energy 10-40 kHz, inter-click interval near 0.15 s) delphinid category associated with the larger delphinids with lower frequency echolocation including short-finned pilot whales (Globicephala macrorhynchus)<sup>25</sup>, melon-headed whales (Peponocephala electra)<sup>26</sup>, false killer whales (Pseudorca crassidens)<sup>27</sup>, rough-toothed dolphins (Steno bredanensis)<sup>28</sup>, and killer whales (Orcinus orca)<sup>29</sup>. Groupbased density estimation methods were used to convert acoustic presence in short five-minute time widows and average group sizes into local density estimates using modeled site-specific detection probabilities<sup>20-24,30</sup>. Weekly densities were estimated by accounting for the total time present, group size, detection probability, area, vocal probability, recording effort, and false positive and false negative rates<sup>30</sup>. Uncertainty for each of these parameters was incorporated using the Delta method<sup>31</sup> (see Methods for details).

Density trends over the decade following DWH, evaluated for each category at each site, strongly suggest long-term density declines for seven of the eight categories. A bootstrap process was used to estimate trend uncertainty, and decadal trends are described as declining only if the interquartile range (IQR) of the distribution resulting from the bootstrap process is entirely negative. Similarly, a trend is considered to be positive if the IQR is entirely positive, and neutral if it spans both positive and negative values. These trends cannot be conclusively linked to the DWH spill due to minimal pre-spill baselines and limited direct evidence of acute impacts on offshore species following the event; however, they may be indicative of chronic impacts<sup>32</sup>. Vessel-based oceanic marine mammal visual surveys conducted in 2003, 2004, and 2009 have provided pre-spill species-level total abundance estimates for US waters of the GoMx<sup>33</sup>. However, while these regional snapshots are valuable, the spatial and temporal resolution of these broad surveys was not sufficient (nor designed) to detect acute changes, or to compare reliably with the point sample data used in this study.

# Results

Echolocation signals (clicks) associated with one or more of the eight categories of odontocetes were identified on 94.1% of days in the 37.0 instrument-years of cumulative recordings, and 937 million detections were analyzed across the five monitoring sites (Supplementary Table 1). Sperm whales were the most commonly detected class at deep sites, followed by the HF delphinid category presumed to consist primarily of four species in the genus Stenella including pantropical, spinner, Clymene, and striped dolphins (S. attenuata, longirostris, clymene, and coeruleoalba respectively) and bottlenose dolphins (Tursiops truncatus), the largest offshore delphinid stocks in the GofMx according to visual surveys<sup>33</sup>. At the shallow sites, the HF delphinid category primarily consists of Atlantic spotted dolphins (S. frontalis) and bottlenose dolphins<sup>33</sup>. HF delphinids were the most common category detected at the two shallow sites, where deep-diving species, including sperm whales and beaked whales, were rarely detected. Rarer species included Blainville's beaked whales (Mesoplodon densirostris) and the LF delphinid category.

# Local densities and trends

Weekly mean sperm whale densities (number of animals per 1000 km<sup>2</sup>, abbreviated hereafter as ind. Per 1000 km<sup>2</sup>) were highest at site MC, nearest the wellhead (4.0 ind. per 1000 km<sup>2</sup>), and GC (1.3 ind. per 1000 km<sup>2</sup>; Supplementary Table 2), where they were present year-round. At these locations, sperm whale densities gradually declined by 24% and 31%, respectively, over the 10-year period (Table 1, Fig. 2). Densities were lowest and most variable at the southernmost site DT (0.6 ind. per 1000 km<sup>2</sup>); however, in contrast with the northern sites, local weekly mean sperm whale densities more than doubled over the 10-year period (172% change). Cuvier's and Gervais' beaked whales occurred in high densities at the DT site, with weekly means of 17.4 ind. per 1000 km<sup>2</sup> and 5.6 ind. per 1000 km<sup>2</sup>, respectively, and much lower densities at the northern Gulf sites (0.9-1.9 ind. per 1000 km<sup>2</sup>) (Supplementary Table 2). Densities of Gervais' beaked whales declined by between 68% and 83% across the deep sites over a 10year period (Table 1, Fig. 2). Cuvier's beaked whale densities also declined by 75% at site DT and to a lesser degree at sites GC and MC (49% and 17% reductions, respectively). Occurrence of Blainville's beaked whales was very low across all sites throughout the monitoring period, with estimated mean weekly densities below 0.4 ind. per 1000 km<sup>2</sup> and no discernible trends (Fig. 2, Supplementary Table 2).

*Kogia* spp. were primarily detected at the two northern sites, with mean densities of 8.7 and 5.4 ind. per 1000 km<sup>2</sup> at sites MC and GC, respectively (Supplementary Table 2). Density declines of 38% and 37% were estimated at these two sites over the 10-year period, respectively. Of note, the *Kogia* spp. density time series exhibited a period of elevated density at site MC in 2014 and 2015 that coincided with a period of low density at site GC (Fig. 2). Estimated densities were low but stable at the southernmost site DT (1.5 ind. per 1000 km<sup>2</sup>).

Estimated HF delphinid densities were high at all sites, with weekly means ranging from a low of 571 ind. per 1000 km<sup>2</sup> at site GC, to highs of 953-1067 ind. per 1000 km<sup>2</sup> at sites MC and DC (Supplementary Table 2). Density declines between 21% and 43% were estimated over the 10-year monitoring period (Table 1, Fig. 3) at all sites except DC, where mean densities increased by an estimated 21%. Risso's dolphin densities declined at two of the four sites where the species was encountered, with declines between 17% and 22% at GC and MC respectively, with neutral trends at the other sites. The LF delphinid category was the only group that showed density increases across multiple sites during the monitoring period. Estimated LF delphinid densities at site MC remained constant over this period, while increases were observed at the three other sites. A strong increase in this class was observed at site DC, where estimated weekly mean densities increased more than three-fold over the 10-year period, though overall densities remained low (4.8 ind. per 1000 km<sup>2</sup>) relative to HF delphinids. Discrimination of individual species within the HF and LF delphinid categories will require further classifier development, currently underway. Subsequent analyses may reveal delphinid species-level trends that differ

#### Table 1 | Site and species-specific trends in marine mammal group densities over the 2010 to 2020 period

	Site	Est. 2010 density (ind. per 1000 km²) Mean (CV)	Est. 2020 density (ind. per 1000 km²) Mean (CV)	Annual Change Median [IQR]	Decadal Percent Change Median [IQR]	Change Type (D/I/N)
Sperm whale	MC	4.6 (0.04)	3.5 (0.07)	-0.11 [-0.14, -0.08]	-24.6 [-29.7, -18.4]	D
	GC	1.6 (0.06)	1.1 (0.06)	-0.05 [-0.06, -0.04]	-31.1 [-36.2, -28.1]	D
	DT	0.3 (0.15)	0.8 (0.06)	0.05 [0.05, 0.06]	172.1 [140.9, 208.7]	I
Cuvier's BW	MC	1.5 (0.06)	1.2 (0.11)	-0.03 [-0.04, -0.01]	-17.3 [-27.7, -6.7]	D
	GC	1.2 (0.09)	0.6 (0.08)	-0.06 [-0.07, -0.05]	-49.0 [-51.9, -44.5]	D
	DT	29.9 (0.04)	7.2 (0.10)	-2.27 [-2.36, -2.15]	-75.4 [-77.2, -73.7]	D
Gervais' BW	MC	2.6 (0.08)	0.4 (0.26)	-0.22 [-0.23, -0.20]	-83.4 [-85.9, -79.2]	D
	GC	3.1 (0.07)	1.0 (0.12)	-0.21 [-0.22, -0.20]	-68.7 [-71.7, -66.7]	D
	DT	10.3 (0.05)	1.8 (0.13)	-0.85 [-0.89, -0.81]	-82.9 [-84.8, -81.1]	D
Blainville's BW	MC	0.2 (0.03)	0.2 (0.03)	0.00 [-0.00, -0.00]	0.0 [-0.0, 0.0]	Ν
	GC	0.5 (0.07)	0.4 (0.07)	-0.01 [-0.01, -0.00]	-18.8 [-25.7, -9.7]	D
	DT	0.2 (0.00)	0.2 (0.00)	0.00 [-0.00, -0.00]	0.0 [-0.0, 0.0]	N
<i>Kogia</i> spp.	MC	11.1 (0.08)	6.8 (0.10)	-0.42 [-0.50, -0.34]	-38.4 [-43.6, -30.8]	D
	GC	6.8 (0.07)	4.3 (0.09)	-0.25 [-0.29, -0.22]	-37.8 [-41.3, -32.7]	D
	DT	1.6 (0.13)	1.5 (0.10)	0.00 [-0.03, -0.00]	0.0 [-17.4, 0.0]	Ν
Risso's dolphin	MC	34.3 (0.06)	26.7 (0.06)	-0.75 [-0.97, -0.53]	-21.9 [-28.1, -15.3]	D
	GC	14.0 (0.07)	11.5 (0.06)	-0.24 [-0.34, -0.15]	-17.6 [-23.3, -11.3]	D
	DT	75.6 (0.05)	71.2 (0.05)	-0.37 [-0.94, 0.09]	-4.9 [-12.2, 1.3]	Ν
	DC	23.4 (0.06)	24.7 (0.06)	0.14 [-0.04, 0.30]	6.0 [-1.9, 13.8]	Ν
Delphinid HF	MC	1186.9 (0.05)	764.3 (0.06)	-40.90 [-47.07, -36.32]	-34.7 [-39.0, -31.4]	D
	GC	679.6 (0.05)	484.2 (0.04)	–19.56 [–22.71, –16.82]	-28.6 [-32.6, -25.1]	D
	DT	821.2 (0.06)	654.0 (0.08)	–16.85 [–22.08, –11.97]	-21.1 [-26.1, -15.2]	D
	DC	952.3 (0.05)	1160.8 (0.06)	20.43 [13.95, 27.88]	21.1 [14.5, 29.9]	I
	MP	970.8 (0.07)	550.4 (0.06)	-41.45 [-47.56, -36.32]	-43.3 [-47.1, -39.2]	D
Delphinid LF	MC	20.0 (0.09)	20.9 (0.08)	0.11 [-0.12, 0.31]	5.7 [-5.7, 16.1]	Ν
	GC	22.4 (0.08)	26.6 (0.06)	0.41 [0.23, 0.61]	18.7 [9.7, 28.5]	I
	DT	11.9 (0.09)	16.5 (0.08)	0.44 [0.35, 0.59]	38.0 [28.4, 53.1]	I
	DC	1.8 (0.30)	7.2 (0.08)	0.55 [0.47, 0.60]	304.8 [222.2, 386.5]	1

Mean densities are from May during the first and last year of the monitoring period. Annual and decadal rates of change are based on a seasonally-detrended linear Theil-Sen fit. Percent change over the 10year monitoring period is reported as the difference between the value of the linear fit between May 2010 and 2020, relative to value of the linear fit in May 2010. Means, CVs, medians, and inter-quartile ranges (IQR) are computed from a bootstrap process in which each fit was computed 100 times from independent random samples of 50% of weekly density estimates. Density units are ind. per 1000 km<sup>2</sup>. The "Change Type" column summarizes the trend as decreasing (D) when the IQRs of the annual and decadal rates of change are entirely negative, increasing (I) when entirely positive, or neutral (N) when the IQRs include zero.

from those of these combined categories, and may help explain the increased LF delphinid densities. Natural periodic fluctuations for these species and regions are not known, therefore the most robust cases are those in which the year-by-year trends are steady and inter-annual variability is low, as seen for the more common species (sperm whales, Cuvier's and Gervais' beaked whales, and HF delphinids).

# Discussion

Long-term declining trends in marine mammal density at the monitored locations were observed over the 10 years following the 2010 DWH oil spill, for seven of the eight identified species categories. In most cases, local mean densities declined by between 13% and 81% from 2010 to 2020, and these trends were most evident at locations where initial densities were high. The largest declines were observed for beaked whales, particularly Gervais' beaked whales, which were initially consistently present at all three deep sites, with densities declining by over 70% across all sites during the monitoring period. In two cases (sperm whale and LF delphinids), sites where certain species were initially rare experienced relatively large increases in local densities of those species. This included the southern DT site, where

sperm whale densities increased, although densities there at the end of the monitoring period were still considerably lower than those observed at the northern sites, and lower than the more common species. A sperm whale demographic study over the 2010–2017 period suggests a greater increase in the occurrence of large, presumed transient males southern DT site than of matriarchal groups<sup>34</sup>. Density increases in the LF delphinid category require additional investigation, but explanations may include increased occupancy due to reduced competition or changes in prey composition<sup>35</sup>.

Due to a lack of exposure and health information for offshore GoMx cetaceans, efforts to quantify the magnitude of the offshore injury caused by the DWH oil spill combined models extrapolated from detailed studies of nearshore bottlenose dolphin populations with predictions of offshore cetacean survival and recovery based on expert elicitations<sup>5,36-39</sup>. Models used in the damage assessment process<sup>39</sup> estimated maximum changes in population sizes ranging from -3% to -23% for the full range of offshore marine mammal species. The primary predictive factor for impact was taken to be the proportion of each population estimated to have been located within the surface oil slick, defined using a cumulative oiling index. Where impacts were predicted, somewhat precipitous declines were expected over



Fig. 2 | Density trends of deep diving odontocetes at deep monitoring locations from 2010 to 2020. Black dots represent acoustically-derived weekly density estimates. Error bars represent standard deviation. Median annual density change

estimates ( $\Delta$ ) and associated interquartile ranges estimated using a bootstrap are noted in text, with seasonally-detrended Thiel-Sen fits of the time series denoted as red dotted lines. Pink bars indicate data gaps.

the first few years following the DWH spill, with gradual recovery beginning within 5–15 years. In a subsequent revision of these models<sup>5</sup> maximum predicted declines were revised to 1.3 to 8.9%, with the largest impacts estimated for spinner dolphins, with population recovery beginning 5 years after the DWH spill, and an estimated population time to recovery of 11.2 years. In contrast with model predictions, observed time series of estimated weekly marine mammal densities at five northeastern GoMx passive acoustic monitoring locations show no evidence of rapid declines associated with acute impacts (e.g., cessation of foraging or death due to direct oil exposure)<sup>40</sup> or habitat abandonment for odontocete species immediately following the 2010 DWH oil spill. Sperm whale density estimates at site MC were lower from May to July 2010 during the oil spill response phase, possibly indicating avoidance of the area, but increased to their highest observed levels by the fall of 2010. No other species' occurrence appears to

have been suppressed at site MC during the oil spill response period in this study. Long-term density declines measured in this 10-year study exceed model-predicted changes at these monitoring locations for critical species, including sperm whales (model-predicted decline 6.4%)<sup>40</sup>, beaked whales (3.8%), and *Stenella* spp. (1.3-8.9.%), and do not suggest recovery trends for affected species to date.

An unusual mortality event that began before the DWH oil spill in early 2010 was exacerbated by the spill, leading to 1,141 reported marine mammal strandings associated with the event from 2010–2014, of which only 5% were oceanic species<sup>9</sup>. Historical analyses suggest pooled carcass recovery rates of 0.4% for the major offshore species in this region<sup>41</sup>, therefore, strandings likely undercount total deaths under the unusual mortality event. Mortality associated with that event may be related to the density declines observed in this study. Movement may also partially





and associated interquartile ranges are noted in text, with seasonallydetrended Thiel-Sen fits of the time series denoted as red lines. Pink bars indicate data gaps.

explain the observed declines and occasional increases: Oceanic visual surveys conducted in 2017–2018 found more sperm whales in the western Gulf than observed prior to the spill, though shorter-term oceanic or seasonal drivers could be the cause<sup>42,43</sup>. Advancing our understanding of the natural trends of oceanic marine mammal populations in this region would likely improve our ability to interpret the observed trends in this study. Additional factors related to human activities, such as underwater radiated noise, ship strikes, oil and gas extraction, seismic surveys, fisheries, and climate change, also warrant thorough investigation. While these factors are beyond the scope of the present analysis, we aim to address them in a future study.

The trends observed in the northeastern GoMx do not necessarily apply across the broader GoMx region. Due to the limited spatial coverage of this dataset, we cannot distinguish between large-scale population distribution shifts and mortality as possible explanations for the observed local density changes. Our initial study design was based on the idea that sites located outside of the oil footprint could be used as controls for comparison against patterns observed at sites within the footprint. In fact, declines in density at the exposed locations, sites MC and MP, generally appeared to be paired with declines at other sites. In some cases, such as for Cuvier's beaked whale, rates of density decline at sites outside of the footprint with high initial densities, were larger than declines measured at sites within it. This likely reflects the fact that trends are more apparent when encounter rates are high<sup>44,45</sup>. Paired declines across these widely spaced sensors may also reflect the high interconnectivity of mobile marine mammal stocks in the region. It is notable that density increases, where observed, occurred outside of the oil footprint. An expanded passive acoustic study is underway to determine whether some taxa may have shifted their distributions west or south<sup>46</sup>.

These time series clearly show high temporal variability in local marine mammal occurrence at seasonal and interannual scales, illustrating the need for extended, consistent, and comparable data collection when aiming to resolve trends using this approach. Oceanographic patchiness, seasonal occupancy patterns, animal movement, and limited acoustic detection ranges influence short-term variability in these time series. Moored passive acoustic recording systems are a robust option for autonomous monitoring over an extended period. Minimization of temporal data gaps, careful calibration of individual recorders, and attention to data quality are critical for long-term trend estimation and inter-site comparisons.

Density estimation from passive acoustics is still a relatively new field, and it is likely that future studies will improve our ability to convert from acoustic detections to absolute numbers of individuals more accurately. The trends described here are computed directly from acoustic occurrence using scalar multipliers; therefore, these trends are expected to be robust to revisions such as improved group size estimates, or updated mean detection probabilities, provided these factors are stable on average over the monitoring period. For example, if the acoustic detection probability was revised for a particular species or site, the magnitudes of estimated weekly densities would scale linearly following Eq. 3, and the slope of the trend line would not change. The group-based density estimation strategy used was selected because it does not require classification of every click (challenging in multispecies scenarios) and facilitates the use of more robust classification of mean click features in short time windows (5 minutes in this study). Groupbased methods also have low sensitivity to potential changes in click rates, a large source of uncertainty in click-based density estimation methods<sup>30</sup>. However, they assume that the mean group sizes of each species are stable over the long term. Average group size estimates for a subset of species including sperm whales, Stenella species, bottlenose dolphins, pygmy killer whales, short-finned pilot whales, and melon-headed whales were 30 to 75 percent lower during 2017 and 2018 vessel-based visual surveys<sup>47</sup> than during 2003-2009 visual surveys<sup>48</sup>. Group sizes estimates are highly variable between surveys, in part due to limited sample sizes, and artifacts may occur due to modifications of survey methodologies; therefore, group size trends are not included in this study. A reduction in group size over the monitoring period of this study would increase the magnitudes of estimated density reduction rates for these species. The assumption that mean group sizes have remained constant despite declining densities, and the election of a density estimation method based on presence/absence in short time windows, rather than of absolute click counts that are likely to decline when fewer animals are present, are conservative choices that err on the side of underestimating declines. For these reasons, we believe that our estimates are robust and conservative for scientific purposes. However, group sizes may have declined in some cases<sup>48-50</sup>, therefore our estimates may be overlyconservative from a precautionary perspective for the management of these long-lived species.

Killer whale population monitoring over 25 years following the *Exxon Valdez* oil spill offers a rare opportunity for comparison<sup>16</sup>. Two matrilineal pods of distinct ecotypes were observed, one resident and one transient. Deaths occurred within both groups following the oil spill (impacts of acute exposure); however, the resident, fish-eating pod gradually began to recover, while the transient pinniped-eating pod has failed to reproduce and is expected to become extirpated<sup>51</sup>. Dietary differences leading to differences in chronic exposure or differential changes in prey availability may explain these diverging outcomes<sup>16</sup>. In contrast with the DWH event, the Exxon Valdez event was a smaller, localized coastal spill in shallow water. GoMx oceanic marine mammal stocks are many times larger than the pods occupying the Prince William Sound, and species differ widely in their prey preferences, which range from benthic invertebrates to fish, deep-sea squid, and marine mammals<sup>52,53</sup>. However, similar mechanisms could apply<sup>6</sup>, and numerous potential impacts and pathways of exposure have been described for nearshore GoMx dolphin stocks<sup>38</sup>. Sublethal health effects of heavy oil exposure were observed for bottlenose dolphin stocks in Barataria Bay, Louisiana following the DWH spill<sup>54</sup>. Subsequent modeling efforts<sup>55</sup> predict a 45% population decline, and a 35-year recovery period required for the population size to return to 95% of the pre-spill baseline.

The degree to which this sparse network of monitoring stations represents larger-scale trends in marine mammal density in the northern Gulf is an ongoing subject of investigation. Current research in the GoMx has expanded this network from 5 to 14 stations including randomly selected long-term and short-term stations, covering the northern and southern Gulf for a period of five years. Future analysis of these historic time series relative to levels and variability observed across more locations will help clarify the degree to which temporal trends are correlated across sites.

These local trends are concerning. If these rates of decline occur across a majority of the Gulf and continue, many of the species in this study could become rare at these monitored locations within the next 30 years. Ongoing environmental impacts from natural resource extraction, shipping, and climate change in the region complicate recovery prospects. A number of iconic marine mammal species are known to be critically endangered, and recognition of their endangered state has often come late (e.g. North Atlantic right whales, Southern Resident killer whales, Rice's whales, and vaquita56after many years of decline, despite the fact that these are often nearer-shore dwellers that are more tractable to monitor than the offshore species considered in this study. Late recognition of low population sizes is largely due to the high uncertainties and low sensitivity of trends produced with traditional survey methods, particularly for rare species, and the lack of timely recognition can ultimately minimize the effectiveness of protective measures once they are enacted. Additional challenges in observing offshore, deep-diving species make these limitations even more concerning. In this study, high temporal resolution passive acoustic monitoring provides an early warning that many protected marine mammal stocks in the GoMx may be on a troubling trajectory, and restoration efforts are needed now to support population health and encourage recovery through protective measures.

Measurement of change in offshore ecosystems has long been considered an intractable problem. In this study, long-duration autonomous passive acoustic recordings were used to measure decadal changes in oceanic marine mammal densities in a region of concern hundreds of kilometers offshore, following an unprecedented deep-water environmental disaster. The findings indicate widespread density declines that vastly exceed expert predictions, which had been based on observations of coastal populations and assumptions that mortality would be linked to direct contaminant exposure. Gradual declines revealed by this observational study instead suggest ongoing, long-term chronic impacts on marine mammal populations in the northeastern GoMx, which may not be suitably approximated by the estimation of relatively short-term exposure to an acute event. The DWH oil spill occurred in the context of decades of heavy exploitation of offshore GoMx ecosystems, and the effect of that event cannot be disentangled from those of the many other chronic stressors in the region. Marine mammal density trends indicate that some of these systems are in trouble, and consistent long-term observational passive acoustic datasets can provide a much-needed window into the status of these otherwise inaccessible regions.

# Methods

# **Data collection**

Passive acoustic recordings were collected using High-frequency Acoustic Recording Packages (HARPs)<sup>60</sup> at five monitoring stations between May 2010 and March 2020 (Supplementary Table 1). The total recording duration across the five sites amounted to 37 instrument-years of continuous data sampled at 200 kHz. Representative hydrophones were calibrated at the Navy's Transducer Evaluation Center (TRANS-DEC). Hydrophones (sensitivity -150 dB re:V  $\mu$ Pa<sup>-1</sup>) were buoyed approximately 20 m above the seafloor. Site locations were selected based on their position relative to the DWH surface oil footprint and were named according to the lease block in which they were located. Sites include three deep sites: Mississippi Canyon (MC) located approximately 15 km northeast of the DWH wellhead within the 2010 surface oil footprint; Green Canyon (GC) a deep canyon 305 km west of the wellhead, and Dry Tortugas (DT) positioned on a sharp slope 520 km to the south of the wellhead, both outside of the footprint. Two shallow shelf sites monitored primarily shallow water species including: Main Pass (MP) located 56 km north of the wellhead, within the surface footprint, and De Soto Canyon (DC) located 230 km to the east of the wellhead, just outside of the footprint. Due to a transcription error, the MC site occupied a slightly shallower position at MC(2), 13 km north of the original site from April 2014 to May 2017.

#### **Detection and classification**

Echolocation clicks and other short duration (<1 ms), impulsive signals with received amplitudes  $\geq$  125 dB peak-to-peak re:1 µPa were detected using a generic energy detector. All acoustic recordings were high-pass filtered below 5 kHz to exclude low and mid-frequency signals. The generic detector identified impulsive events with durations between 0.03 and 1.2 milliseconds, and a waveform energy envelope differencing metric greater than -0.5. The waveform energy envelope differencing metric computes the Hilbert transform of the waveform and then subtracts the mean amplitude of the second half of the waveform from the first half. This check takes advantage of the fact that odontocete clicks have most of their energy at the beginning of the signal in a decaying pattern (delphinids, sperm whales, and Kogia spp.) or balanced between the first and second halves in a pulse shape (beaked whales), rather than a strong increase in amplitude at the end. Both the duration and waveform envelope metric criteria are selected to be highly permissive, with the goal of detecting anything click-like, to minimize false negatives at the detection stage. This, combined with a relatively high received level threshold, allows the required density estimation assumption that all clicks above a certain threshold are available for classification to be met, and ensures they are accounted for in the false positive and false negative rate estimates.

A semi-supervised clustering process was used to generate training classes for 12 different signal categories based on spectral shapes and interclick interval distributions<sup>61,62</sup>. These features were compared by computing the correlation distance between pairs of clicks, and then multiplying the two distances to produce a combined similarity score for each pair. Unsupervised clustering was used to group similar signal types, and those types were further combined by an expert analyst across sites and then subsampled to produce training, test and validation sets (see Refs. 52,62 for detailed workflow). Eight odontocete classes including sperm whale (Physeter macrocephalus), Cuvier's beaked whale (Ziphius cavirostris), Gervais' beaked whale (Mesoplodon europaeus), Blainville's beaked whale (Mesoplodon densirostris), dwarf and/or pygmy sperm whale (Kogia spp.), Risso's dolphin (Grampus griseus), and two generalized delphinid classes, high frequency (HF, peak energy 20-80 kHz, inter-click interval near 0.06 s) and low frequency (LF, 10-40 kHz, inter-click interval near 0.15 s), were defined. Four non-target classes (snapping shrimp, boats, echosounders/ sonar, and noise) were included in the training set, and signals classified as any of these non-target classes were excluded from further analyses. The HF delphinid class likely represents offshore bottlenose dolphins (Tursiops truncatus) and delphinids in the genus Stenella. Pantropical spotted dolphins (S. attenuata) have historically been the most common offshore delphinid species in the GoMx region, with related Atlantic spotted dolphins (S, frontalis), spinner dolphins (S. longirostris), Clymene dolphins (S. clymene), and striped dolphins (S. coeruleoalba) also present<sup>62</sup>. All delphinid echolocation clicks at the shallowest site (MP), where bottlenose and Atlantic spotted dolphins are expected to be the primary species<sup>52,63</sup>, were consistent with the acoustic features of the HF delphinid category. Therefore, we expect that bottlenose dolphins are included in the HF delphinid class. The LF delphinid class likely represents a mix of the larger delphinid and lower frequency species, with longer inter-click intervals, including short-finned pilot whale<sup>25</sup>, melon-headed whale<sup>28</sup>, rough-toothed dolphin<sup>28</sup>, false killer whale<sup>28</sup>, and killer whale<sup>29</sup>.

A deep neural network trained from these classes was used to automatically attribute identity classification labels to echolocation click signals in successive five-minute time windows based on mean spectrum, waveform and inter-click interval features<sup>61</sup> (Supplementary Fig. 1). To reduce misclassification rates, the snapping shrimp category was only allowed for the shelf sites, because snapping shrimp are not acoustically present at the deep sites. If snapping shrimp was identified as the most probable class for a deep site bin by the classifier, this label was rejected and the second most probable class label, according to the Softmax probabilities output by the neural network, was applied to that bin. This approach was used as an alternative to training separate networks for each site. An analyst manually classified raw detections in 3% of all detection-positive hours from each site, using a graphical user interface<sup>64</sup> (Supplementary Table 3). These manual classification labels were compared with the automatic labels to estimate bin-level species-specific and site-specific false positive rates ( $\hat{c}_p$ ) and false negative rates ( $\hat{c}_p$ ) (Supplementary Table 3), computed as:

$$\hat{c}_{\rm p} = FP/(FP + TN) \tag{1}$$

$$\hat{c}_{n} = FN/(FN + TP)$$
(2)

where FP and FN are the number of false positive and negative bins, respectively, and TP and TN represent the number of true positive and true negative bins, respectively. False negatives in this study are those due to misclassification. False positives from cases in which entire bins containing clicks have been missed by the detector are assumed to be rare, due to the very permissive design of the generic detector<sup>61</sup>.

Classification errors varied by species class (Supplementary Fig. 2, Supplementary Table 3). For sperm whales, false positive and negative rates ranged from 1-7% at deep sites due to confusion with ship noise in cases where characteristic modal inter-click intervals were not discernible in the inter-click interval distributions (Supplementary Table 3). Sperm whale false positives tended to occur at the beginnings and ends of vessel transits, while false negatives tended to occur during dense bouts of clicking with many overlapping click trains. Across species classes, false negatives also occurred when multiple classes were present simultaneously and one (or more) was missed, such as bins containing thousands of clicks of a common class, which can prevent small numbers (tens or less) of clicks of a minority class from forming a cluster for classification. This is likely the cause of elevated false negative rates for Kogia spp.: their very high-frequency clicks (>90 kHz) attenuate strongly, tend to be recorded in small numbers, and are occasionally overshadowed by large clicking events associated with other species. Highfrequency attenuation with distance also affects the received spectral content of signals from other species; however, spectral peaks and troughs in the lower frequency ends of the spectra, as well as inter-click intervals, are the diagnostic features for classification purposes. By preserving amplitude information in the classification process, the neural network is able to learn that high-frequency amplitude generally declines with received signal amplitude, minimizing misclassifications due to variable amounts of high-frequency attenuation. Overall, beaked whale species and Kogia spp. tended to have low false positive rates (<1%), while false negative rates were higher (3-10%) due to low overall detection rates for these species. Sperm whale, beaked whale, and Kogia spp. densities were not computed for the shallow sites DC and MP, but error rates are given for completeness and false negative rates can be high due to small sample sizes.

Classifier confusion between the aggregated delphinid categories in some cases led to somewhat elevated false positive and negative rates compared to the deep diver classes (Supplementary Table 3). Bins with small numbers of low amplitude delphinid clicks tend to be more difficult to classify. The HF delphinid category includes multiple subtypes, including one with banding patterns sometimes confused with Risso's dolphin. This type is found primarily at site DC leading to higher error rates for the Risso's dolphin, HF and LF delphinid classes at that location. Site DC and MP also have snapping shrimp activity, which can be difficult for both automatic and manual labeling when events have few detections. The manual classification was more likely to label such cases as delphinids based on weak modal interclick intervals, elevating false positive rates for HF delphinids. Reported error metrics are for the entire detection and classification pipeline.

#### **Density estimation**

Group-based density estimation methods were used to convert acoustic presence into local density estimates<sup>20–24,30</sup>. If a species is acoustically detected in a short five-minute time window, group-based methods assume that a group of average size is present within a defined area around the monitoring station. Window duration is short to allow the simplifying assumption that no animals enter or exit the monitored area during the window. Detection radii are determined by the maximum acoustic detection range for each species (frequency, source level and orientation-dependent). Species-specific group detection probabilities within that range are computed using Monte Carlo simulations to incorporate acoustic and behavioral variability and uncertainty. Mean group sizes and variance are typically derived from visual surveys, and acoustic cue rates, as well as subsurface depths and orientations, are estimated from tag and acoustic tracking data.

Weekly marine mammal densities per site  $(\hat{D}_{kt})$  were estimated for each species following Marques et al. <sup>30</sup> using a group-counting method,

$$\hat{D}_{kt} = \frac{n_{kt}\,\hat{s}\,(1-\hat{c}_{pk})(1+\hat{c}_{nk})}{\pi\,w^2\,\hat{P}_k\,\hat{P}_v\,T_{kt}} \tag{3}$$

in which  $n_{kt}$  is the mean number of positive five-minute windows per day in week t at site k,  $\hat{s}$  is the mean group size (from visual surveys),  $\hat{c}_{pk}$  is the estimated site-specific false positive rate,  $\hat{c}_{nk}$  is the estimated site-specific false negative rate,  $\hat{P}_k$  is the estimated probability of detecting a group within radius w of the recording station,  $\hat{P}_v$  is the probability that at least one animal in a group is vocalizing in a five-minute period, and  $T_{kt}$  is the mean number of time bins analyzed in week t at site k (used to account for monitoring effort). Coefficients of variation (CVs) were estimated using the delta method<sup>31</sup> as

$$CV(\hat{D}_{kt}) = \left(\hat{D}_{kt}^{2} * (CV(\hat{P}_{v})^{2} + CV(\hat{P}_{k})^{2} + CV(\hat{s})^{2})\right)^{1/2}$$
(4)

to incorporate uncertainty in vocalization probabilities, detection probabilities, and group size estimates.

For each marine mammal species class and site combination, the probability  $(\hat{P}_k)$  of detecting a group of animals in a five-minute time window within the effective detection radius was estimated using a Monte Carlo simulation. These simulations took into account acoustic and behavioral parameters including minimum click amplitude detection thresholds, echolocation clicks peak frequency content and distance-dependent attenuation, animal dive depths, ascent and descent rates, and vocalization rates during different dive phases (Supplementary Table 4). To incorporate uncertainty and variability, ranges for the means and standard deviations of each of these parameters were selected. For each site and species, a simulation was iterated 500 times, each time randomly selecting a different mean and standard deviation for each parameter from the defined ranges. Then, within each iteration, 100,000 animat models were simulated, with properties pulled from randomly-generated parameter distributions defined by the selected means and standard deviations. Species and site-specific mean detection probabilities and associated uncertainties were computed across the 500 model iterations. Previously-published behavioral parameter estimates for beaked whales<sup>24</sup>, sperm whales<sup>65</sup>, delphinids<sup>23</sup>, and Kogia<sup>20</sup> were used for simulation of group behavior. However, in their original published forms, each of these studies used a different minimum received-level threshold when estimating detection probabilities. Therefore, all simulations were reevaluated for this project with a consistent minimum received level (125 dB<sub>nn</sub>) for all species. This received level threshold was selected to minimize the influence of variable background noise levels on detection counts. For simplicity, the Monte Carlo simulation procedure uses a static received level threshold to determine whether a simulated click is detectable, rather than

mimicking variable noise conditions in which detector performance can be difficult to predict. Iterative manual review of detector performance across a majority of the deployments indicated that the selected threshold minimized noise masking in the high-pass filtered data, such that clicks could still be reliably detected in the presence of ships and elevated wind conditions. A higher threshold would further minimize potential noise masking but would begin to greatly reduce the number of clicks detected for small delphinids with lower source levels and highly directional beaked whale signals, likely reducing classification accuracy (characteristic inter-click intervals distributions can become degraded due to small sample size). Independent simulations were run to estimate detectability for the two MC locations (1 and 2) to account for potential differences in detectability due to sensor depth (980 m vs. 800 m). Climatological mean sound speed profiles (full water column, 1 m resolution) for the months of January and July were extracted from the Global Digital Elevation Map v3.066 and used to compute frequency-dependent propagation loss volumes at each site at the typical echolocation click peak frequency for each species. In some cases, mean delphinid detection probability estimates differed slightly but significantly between summer and winter models due to differences in the depth and temperature of the surface mixed layer (Supplementary Table 5). A twosample t-test was used to test for equal means with the null hypothesis rejected at the 5% significance level. Where significantly different, the summer detection probability was used for density estimation from May to October, and the winter detection probability was used for the remaining months. Deep divers, including sperm whales, beaked whales, and Kogia spp., were primarily detected at the deep sites. Densities were not computed for these species for shallow sites, despite occasional occurrence, due to a lack of information on shallow water behaviors.

Vocalization probabilities were taken from the literature for all nondelphinid species (Supplementary Table 4). The best practice for density estimation is to use parameter estimates derived from populations as similar to the target population and target behavioral states as possible. In this case, estimates were from northern GoMx populations, with the exception of Blainville's beaked whale, which is relatively rare in the GoMx, and for which best available estimates were from individuals in the Bahamas<sup>67</sup>. Current studies in the GoMx are underway to produce and refine local estimates for vocalization probabilities and other parameters to reduce uncertainty and potential bias associated with using measurements from other regions. For delphinids, which exhibit patterns of primarily nocturnal echolocation in this region, the probability of vocalization was computed by taking into account day and night differences at each site, to account for seasonal differences in daylight duration (Supplementary Table 6). First, all-time bins were assigned to daylight or night based on sunrise and sunset times. A bootstrap approach was used to repeatedly select 25% of days and to compute the proportions of positive daylight bins  $(\hat{D}_{p})$  and positive night bins  $(\hat{N}_{p})$  relative to all positive bins in each bootstrap sample. Next, the proportion of daylight hours  $(P_{day})$  in each week was computed for the entire time series. Assuming that groups of delphinids click constantly while foraging at night, an adjusted probability of vocalization  $(\hat{P}_{v})$  was computed as

$$\hat{P}_{\rm v} = (P_{\rm day}\,\hat{D}_{\rm p}) + ((1 - P_{\rm day})\hat{N}_{\rm p}) \tag{5}$$

Group sizes for the multi-species HF delphinid and LF delphinid classes were computed as the weighted average of the included species' group size estimates, scaled by their relative population sizes as estimated by NOAA visual surveys in 2017 and 2018<sup>48</sup> (Supplementary Table 7). For example, if a class includes two species A and B with respective estimated group sizes  $\hat{s}_{A}$  and  $\hat{s}_{B}$ , then the combined group size  $\hat{s}_{AB}$  is estimated as

$$\hat{s}_{AB} = \frac{n_A \hat{s}_A}{n_A + n_B} + \frac{n_B \hat{s}_B}{n_A + n_B}$$
(6)

where  $n_A$  and  $n_B$  are the estimated numbers of animals of species A and B, respectively.

**Delphinid Vocalization Probabilities**. The diel vocalization probability analysis showed that deep divers generally did not have a diel pattern in vocalization rate, while delphinids were far less likely to be detected during the day (Supplementary Table 6). This trend for delphinids was particularly strong at the deep sites where daytime click-positive bins represented only 4% to 22% of total click-positive bins (CVs = 0.00–0.02) for the three delphinid classes (Risso's, HF and LF). At the shelf sites, the pattern was weaker, with 20 - 36% of click-positive bins (CVs = 0.02–0.05) occurring during daylight hours on average. This difference is likely due to differences in species composition and foraging strategies<sup>68</sup>.

#### Long-term trend analysis

Long-term trends were estimated at each site for each species from weekly density estimates which were de-seasoned using a monthly seasonal pattern decomposition procedure<sup>69</sup>. This procedure consists of detrending the data, regressing the detrended time series against a set of monthly indicator variables, subtracting the seasonal component, and adding the trend back in. De-seasoned time series were fit using a Thiel-Sen regression<sup>70,71</sup>, a robust linear estimation process that is insensitive to outliers. The Theil-Sen algorithm computes a slope between each pair of points in a time series, using the true time difference, and then estimates the overall slope as the mean across all pairs, yielding a robust fit of the data, insensitive to outliers. Means and associated CVs, and interquartile ranges were estimated using a bootstrap approach<sup>72</sup>, selecting 50% of points and computing the associated slope 100 times. Trends are reported as change in number of animals per 1000 km<sup>2</sup> per year. The values of the linear bootstrapped Thiel Sen fit of the de-seasoned data in May 2010 and May 2020 were used to approximate mean local density at the beginning and end of the monitoring period time series, with percent change computed as the difference between those two points, divided by the estimated mean starting density in May 2010. For summarization and interpretation, trends are described as declining if the interquartile range of the slope spans only negative values, increasing if the interquartile range spans only positive values, and neutral if the range includes zero.

#### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

# Data availability

The dataset can be accessed at: https://doi.org/10.5061/dryad.9zw3r22n4.

#### Code availability

The code can be accessed at: https://github.com/MarineBioAcousticsRC/ 2010-2021\_timeseries\_analysis.

Received: 23 May 2024; Accepted: 20 November 2024; Published online: 21 December 2024

#### References

- 1. Joye, S. B. Deepwater Horizon, 5 years on. *Science* **349**, 592–593 (2015).
- Kujawinski, E. et al. Fate of dispersants associated with the Deepwater Horizon oil spill. *Environ. Sci. Technol.* 45, 1298–1306 (2011).
- Testa, J. M., Eric Adams, E., North, E. W. & He, R. Modeling the influence of deep water application of dispersants on the surface expression of oil: A sensitivity study. *J. Geophys. Res.: Oceans* **121**, 5995–6008 (2016).
- 4. Berenshtein, I. et al. Invisible oil beyond the Deepwater Horizon satellite footprint. *Sci. Adv.* **6**, eaaw8863 (2020).
- Marques, T. A. et al. Population consequences of the Deepwater Horizon oil spill on pelagic cetaceans. *Mar. Ecol. Prog. Ser.* **714**, 1–14 (2023).
- Sutton, T. T. et al. The open-ocean gulf of Mexico after Deepwater Horizon: synthesis of a decade of research. *Front. Mar. Sci.* 9, 753391 (2022).

- DWH NRDA Trustees. Deepwater Horizon oil spill: Final programmatic damage assessment and restoration plan and final programmatic environmental impact statement. noaa.gov/restorationplanning/gulfplan, Accessed 10 Feb 2017. (2016).
- Levin, L. A. et al. Global observing needs in the deep ocean. Front. Mar. Sci. 6, 241 (2019).
- Litz, J. A. et al. Review of historical unusual mortality events (UMEs) in the Gulf of Mexico (1990-2009): providing context for the multi-year northern Gulf of Mexico cetacean UME declared in 2010. *Dis. Aquat. Org.* **112**, 161–175 (2014).
- Hildebrand, J. A. Anthropogenic and natural sources of ambient noise in the ocean. *Mar. Ecol. Prog. Ser.* 395, 5–20 (2009).
- Wang, Z., Boyer, T., Reagan, J. & Hogan, P. Upper-Oceanic Warming in the Gulf of Mexico between 1950 and 2020. *J. Clim.* **36**, 2721–2734 (2023).
- Liu, Y., Lee, S. K., Muhling, B. A., Lamkin, J. T. & Enfield, D. B. Significant reduction of the Loop Current in the 21st century and its impact on the Gulf of Mexico. *J. Geophys. Res.: Oceans* **117** (2012).
- 13. Ochoa, J. et al. Deep-Water Warming in the Gulf of Mexico from 2003 to 2019. *J. Phys. Oceanogr.* **51**, 1021–1035 (2021).
- Mullin, K. & Hansen, L. Marine mammals of the northern Gulf of Mexico. *The Gulf of Mexico large marine ecosystem: assessment, sustainability, and management.* Blackwell, New York, 269-277 (1999).
- 15. Hazen, E. L. et al. Marine top predators as climate and ecosystem sentinels. *Front. Ecol. Environ.* **17**, 565–574 (2019).
- Matkin, C. O., Saulifis, E. L., Ellis, G. M., Olesiuk, P. & Rice, S. D. Ongoing population-level impacts on killer whales *Orcinus orca* following the Exxon Valdez oil spill in Prince William Sound, Alaska. *Mar. Ecol. Prog. Ser.* **356**, 269–281 (2008).
- 17. Nelms, S. E. et al. Marine mammal conservation: over the horizon. *Endanger. Species Res.* **44**, 291–325 (2021).
- White, E. R., Schakner, Z., Bellamy, A. & Srinivasan, M. Detecting population trends for US marine mammals. *Conserv. Sci. Pract.* 4, e611 (2022).
- Kessler, J. D. et al. A persistent oxygen anomaly reveals the fate of spilled methane in the deep Gulf of Mexico. *Science* **331**, 312–315 (2011).
- Hildebrand, J. A. et al. Assessing seasonality and density from passive acoustic monitoring of signals presumed to be from pygmy and dwarf sperm whales in the Gulf of Mexico. *Front. Mar. Sci.* 6, 66 (2019).
- Marques, T. A. et al. Estimating animal population density using passive acoustics. *Biol. Rev.* 88, 287–309 (2013).
- Thomas, L. et al. Distance software: design and analysis of distance sampling surveys for estimating population size. *J. Appl. Ecol.* 47, 5–14 (2010).
- Frasier, K. E. et al. Delphinid echolocation click detection probability on near-seafloor sensors. J. Acoust. Soc. Am. 140, 1918–1930 (2016).
- Hildebrand, J. et al. Passive acoustic monitoring of beaked whale densities in the Gulf of Mexico during and after the Deepwater Horizon oil spill. *Nature. Sci. Rep.* 5, 16343 (2015).
- Cohen, R. E. et al. North Atlantic odontocete echolocation click types using machine learning and spatiotemporal correlates. *PloS one* **17.3**, e0264988 (2022).
- Baumann-Pickering, S., Wiggins, S. M., Hildebrand, J. A., Roch, M. A. & Schnitzler, H. U. Discriminating features of echolocation clicks of melon-headed whales (*Peponocephala electra*), bottlenose dolphins (*Tursiops truncatus*), and Gray's spinner dolphins (*Stenella longirostris*) *Iongirostris*). J. Acoust. Soc. Am. **128**, 2212–2224 (2010).
- Baumann-Pickering, S. et al. False killer whale and short-finned pilot whale acoustic identification. *Endanger. Species Res.* 28, 97 (2015).
- Ziegenhorn, M. A. et al. Discriminating and classifying odontocete echolocation clicks in the Hawaiian Islands using machine learning methods. *PLOS One* **17**, e0266424 (2022).

- Leu, A. A., Hildebrand, J. A., Rice, A., Baumann-Pickering, S. & Frasier, K. E. Echolocation click discrimination for three killer whale ecotypes in the Northeastern Pacific. *J. Acoust.Soc. Am.* **151**, 3197–3206 (2022).
- Marques, T., Thomas, L., Ward, J., DiMarzio, N. & Tyack, P. Estimating cetacean population density using fixed passive acoustic sensors: An example with Blainville's beaked whales. *J. Acoust. Soc. Am.* **125**, 1982–1994 (2009).
- Powell, L. A. Approximating variance of demographic parameters using the delta method: a reference for avian biologists. *Condor* 109, 949–954 (2007).
- 32. Fair, P. A. & Becker, P. R. Review of stress in marine mammals. *J. Aquat. Ecosyst. Stress Recovery* **7**, 335–354 (2000).
- Garrison, L. P. & Aichinger Dias, L. Distribution and abundance of cetaceans in the northern Gulf of Mexico. NOAA Technical Memorandum NMFS-SEFSC-747, 40 p. (2020).
- Solsona Berga, A. et al. Accounting for sperm whale population structure in density estimation using passive acoustic monitoring. *Mar. Ecol. Prog. Ser.* 746, 121–140 (2024).
- Murawski, S. A. et al. Vulnerability and resilience of living marine resources to the Deepwater Horizon oil spill: an overview. *Front. Mar. Sci.* 10, 1202250 (2023).
- Schwacke, L. H. et al. Health of common bottlenose dolphins (*Tursiops truncatus*) in Barataria Bay, Louisiana, following the Deepwater Horizon oil spill. *Environ. Sci. Technol.* 48, 93–103 (2013).
- Schwacke, L. H. et al. Quantifying injury to common bottlenose dolphins from the *Deepwater Horizon* oil spill using an age-, sex-and class-structured population model. *Endanger. Species Res.* 33, 265–279 (2017).
- Takeshita, R. et al. The Deepwater Horizon oil spill marine mammal injury assessment. *Endanger. Species Res.* 33, 95–106 (2017).
- MMIQT, D. Models and analyses for the quantification of injury to Gulf of Mexico cetaceans from the *Deepwater Horizon* Oil Spill. 1–64 (2015).
- Geraci, J. R. In Sea Mammals and Oil: Confronting the Risks (eds Geraci, J. R. & St. Aubin, D. J.) Ch. 6, 167-192, (Academic Press, Inc., 1990).
- Williams, R. et al. Underestimating the damage: interpreting cetacean carcass recoveries in the context of the Deepwater Horizon/BP incident. *Conserv. Lett.* 4, 228–233 (2011).
- Morano, J. L. et al. Seasonal movements of Gulf of Mexico sperm whales following the *Deepwater Horizon* oil spill and the limitations of impact assessments. *Mar. Pollut. Bull.* **161**, 111627 (2020).
- Roberts, J. J. et al. Habitat-based cetacean density models for the US Atlantic and Gulf of Mexico. *Nat. Sci. Rep.* 6 (2016).
- Barnes, R. The problem of precision and trend detection posed by small elephant populations in West Africa. *Afr. J. Ecol.* **40**, 179–185 (2002).
- Taylor, B. L., Martinez, M., Gerrodette, T., Barlow, J. & Hrovat, Y. N. Lessons from monitoring trends in abundance of marine mammals. *Mar. Mammal. Sci.* 23, 157–175 (2007).
- Rafter, M. R. et al. 2010-2021 Long-term Investigations into Soundscapes, Trends, Ecosystems, and Noise in the Gulf of Mexico. Marine Physical Laboratory, Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA., MPL Technical Memorandum 662 (2022).
- Garrison, L. et al. Gulf of Mexico Marine Assessment Program for Protected Species (GOMMAPPS): Marine Mammals. Vol. 2, Appendix C: Gulf of Mexico marine mammal spatial density models. Obligation No.: M1217PG00013. Report No.: OCS Study BOEM 02023-00042. New Orleans LA.: US Department of the Interior, Bureau of Ocean Energy Management, 1264 p. (2023).
- 48. Garrison, L. P., Ortega-Ortiz, J. & Rappucci, G. Abundance of Marine Mammals in Waters of the U.S. Gulf of Mexico During the Summers of

2017 and 2018. National Marine Fisheries Service, Southeast Fisheries Science Center, (2020).

- Rappucci, G. & Garrison, L. SEFSC GoMMAPPS Vessel Survey 2018 Summer. Data downloaded from OBIS-SEAMAP (http://seamap.env. duke.edu/dataset/1984) on 2024-09-17 (2019).
- Rappucci, G. & Garrison, L. SEFSC GoMMAPPS Vessel Survey 2017 Summer. Data downloaded from OBIS-SEAMAP (http://seamap.env. duke.edu/dataset/1980) on 2024-09-17 (2019).
- Esler, D. et al. Timelines and mechanisms of wildlife population recovery following the *Exxon Valdez* oil spill. *Deep Sea Res. Part II: Top. Stud. Oceanogr.* **147**, 36–42 (2018).
- Davis, R. W. et al. Physical habitat of cetaceans along the continental slope in the north-central and western Gulf of Mexico. *Mar. Mammal. Sci.* 14, 490–507 (1998).
- Pitman, R. L., OSullivan, S. & Mase, B. Killer whales (*Orcinus orca*) attack a school of pantropical spotted dolphins (*Stenella attenuata*) in the Gulf of Mexico. *Aquat. Mamm.* 29, 321–324 (2003).
- Schwacke, L. H. et al. Health of common bottlenose dolphins (*Tursiops truncatus*) in Barataria Bay, Louisiana, following the Deepwater Horizon oil spill. *Environ. Sci. Technol.* 48, 93–103 (2014).
- Schwacke, L. H. et al. Modeling population effects of the Deepwater Horizon oil spill on a long-lived species. *Conserv. Biol.* 36, e13878 (2022).
- NMFS & NOAA. Endangered and Threatened Wildlife; 90-Day Finding on a Petition To List the Gulf of Mexico Bryde's Whale as Threatened or Endangered Under the Endangered Species Act. NOAA-NMFS-2014-0157 (2015).
- Pace, R. M. III, Corkeron, P. J. & Kraus, S. D. State–space mark–recapture estimates reveal a recent decline in abundance of North Atlantic right whales. *Ecol. Evol.* 7, 8730–8741 (2017).
- Rosel, P. E., Wilcox, L. A., Yamada, T. K. & Mullin, K. D. A new species of baleen whale (*Balaenoptera*) from the Gulf of Mexico, with a review of its geographic distribution. *Mar. Mammal. Sci.* 37, 577–610 (2021).
- Krahn, M. M. et al. Status review of southern resident killer whales (Orcinus orca) under the Endangered Species Act. U.S. Department of Commerce, NOAA Tech. Memo NMFS-NWFSC-54, 133 p. (2002).
- Wiggins, S. M. & Hildebrand, J. A. High-frequency Acoustic Recording Package (HARP) for broad-band, long-term marine mammal monitoring. *Underw. Technol. Workshop Sci. Use Submar. Cables Relat. Technol.* 2007, 551–557 (2007).
- Frasier, K. E. A machine learning pipeline for classification of cetacean echolocation clicks in large underwater acoustic datasets. *PLoS Comput. Biol.* **17**, e1009613 (2021).
- Frasier, K. E. et al. Automated classification of dolphin echolocation click types from the Gulf of Mexico. *PLoS Comput. Biol.* 13, e1005823 (2017).
- Fulling, G. L., Mullin, K. D. & Hubard, C. W. Abundance and distribution of cetaceans in outer continental shelf waters of the US Gulf of Mexico. *Fish. Bull.* **101**, 923–932 (2003).
- Solsona-Berga, A., Frasier, K. E., Baumann-Pickering, S., Wiggins, S. M. & Hildebrand, J. A. DetEdit: A graphical user interface for annotating and editing events detected in long-term acoustic monitoring data. *PLOS Comput. Biol.* **16**, e1007598 (2020).
- Solsona Berga, A. Advancement of methods for passive acoustic monitoring: a framework for the study of deep-diving cetaceans. Doctoral dissertation, Universitat Politècnica de Catalunya, (2019).
- 66. US Navy, Lockheed Martin. GDEM-V QAV Analysis and Delivery (N62306-08-F-7504). (2009).
- Warren, V. E. et al. Spatio-temporal variation in click production rates of beaked whales: Implications for passive acoustic density estimation. *J. Acoust. Soc. Am.* **141**, 1962–1974 (2017).
- Cohen, R. E., Frasier, K. E., Baumann-Pickering, S. & Hildebrand, J. A. Spatial and temporal separation of toothed whales in the western North Atlantic. *MEPS* **720**, 1–24 (2023).
- Cleveland, R. B., Cleveland, W. S., McRae, J. E. & Terpenning, I. STL: A seasonal-trend decomposition procedure based on loess. *J. Off. Stat.* 6, 3–33 (1990).

- Thiel, H. A rank-invariant method of linear and polynomial regression analysis, Part 3. Proc. Koninalijke Nederlandse Akademie van. Weinenschatpen A 53, 1397–1412 (1950).
- Sen, P. K. Estimates of the regression coefficient based on Kendall's tau. J. Am. Stat. Assoc. 63, 1379–1389 (1968).
- Wilcox, R. R. Some results on extensions and modifications of the Theil-Sen regression estimator. Br. J. Math. Stat. Psychol. 57, 265–280 (2004).
- Kobara, S. Deepwater Horizon Oil spill dataset compiled for Gulf of Mexico Ocean Observation (GCOOS). (2019). https://gisdata.gcoos. org/datasets/afafd2255f9d43bd8a5531de7e98c9a5/explore?layer=1 Accessed Aug. 15, 2023).
- GEBCO Compilation Group. GEBCO 2020 Grid. (2020). https://doi. org/10.5285/a29c5465-b138-234d-e053-6c86abc040b9. Accessed Oct. 12, 2021).

# Acknowledgements

Many people have participated in the collection and analysis of this dataset. The authors would like to thank HARP technicians Bruce J. Thayre, John P. Hurwitz, T. Christianson, B. Hurley, Chris Garsha, Joshua Jones, Ryan Griswold, and engineers K. Lenssen, K Rosier, S. Wagner. Data processors and analysts included Hannah R. Bassett, Erin O'Neil, Shelby Bloom and Diego Majewski. Vessels and key cruise leads who made this work possible include Tony Martinez, Jesse Wicker, Dr. Keith Mullin and the crews of R/V Gordon Gunter and R/V Pisces, and Tad Berkey and crew of the R/V Pelican. We also would like to thank Dr. Steven Murawski and Sheryl Gilbert for their long-term support of this effort. Funding for this study was part of the "Assessing long-term trends and processes driving variability in cetacean density throughout the Gulf of Mexico using passive acoustic monitoring and habitat modeling" project under federal funding opportunity Grant No. NOAA-NOS-NCCOS-2019-2005608 from the National Oceanic and Atmospheric Administration's RESTORE Science Program (ROR - https:// ror.org/0042xzm63) through the Gulf Coast Restoration Trust Fund to the NOAA Southeast Fisheries Science Center. Ship-time aboard the R/V Pelican for HARP deployment and servicing cruises in August 2020 and August 2021 were funded under the Office of Naval Research Task Force Ocean (Robert Headrick). Funding for historic HARP data leveraged by the LISTEN GoMex program were provided by the Deepwater Horizon Natural Resource Damage Assessment program, the Gulf of Mexico Research Initiative through C-IMAGE I, II and III at USF (Steve Murawski and Sheryl Gilbert), and NOAA NMFS' Office of Science and Technology (Mridula Srinivasan), NOAA's Southeast Regional Office (Laura Engleby). Acoustic data collection at four long-term sites (GC, MC, DC, DT) over the 2020-2021 period, marine mammal occurrence analyses in these recordings, and all ambient and anthropogenic noise analyses from historic and 2020-2021 recordings were funded by the Deepwater Horizon Open Ocean Trustee Implementation Group's "Reduce Impacts of Anthropogenic Noise on Cetaceans" project to restore natural resources injured by the 2010 Deepwater Horizon oil spill in the Gulf of Mexico. Funding for historic HARP data collection and analysis was also provided by the Natural Resource Damage Assessment partners (20105138), the US Marine Mammal Commission (20104755/E4061753), the Southeast Fisheries Science Center under the Cooperative Institute for Marine Ecosystems and Climate (NA10OAR4320156) with support through Interagency Agreement #M11PG00041 between the Bureau of Offshore Energy Management, Environmental Studies Program, and the National Marine Fisheries Service, Southeast Fisheries Science Center. A subset of the data used for this study are archived by the Gulf of Mexico Research Initiative at https://data.gulfresearchinitiative.org/data/R4.x267.180:0011 maintained by the Gulf Research Initiative Information and Data Cooperative. M. Le Hénaff received partial support by the NOAA Atlantic Oceanographic and Meteorological Laboratory (Miami, FL) and was supported in part under the auspices of the Cooperative Institute for Marine and Atmospheric Studies (CIMAS), a cooperative institute of the University of Miami and NOAA (agreement NA20OAR4320472). The scientific results and

conclusions, as well as any views or opinions expressed herein, are those of the authors and do not necessarily reflect those of NOAA or the Department of Commerce.

# Author contributions

K.E.F. led the study, conducted data analyses, and wrote the manuscript. M.A.R. conducted the detection and classification tasks and assisted with data analysis. A.S.B. contributed to sperm whale density estimation. J.A.H. initiated the monitoring program, led the development of density estimation processes, and advised on the project. S.M.W. developed recording systems, led data collection, and provided the system with calibrations. L.P.G. contributed to the exploration of links with visual survey and damage assessment efforts and provided guidance. H.F.M. assisted with project management, provided review, and assisted with beaked whale density estimation. A.G. contributed Gulf of Mexico oceanographic knowledge and provided manuscript reviews. A.S. contributed Gulf of Mexico ecological knowledge and provided manuscript reviews. L.E.W.H. assisted with data management organization and distribution and provided reviews. C.C.W. assisted with data management organization and distribution, and provided reviews. M.L.H. contributed to the Gulf of Mexico oceanographic and oil spill plume context knowledge and provided a review. M.S.S. was the funding lead who secured support to complete this analysis. She also contributed extensively through project direction, guidance, research process development, and many reviews.

# **Competing interests**

The authors declare no competing interests.

# **Additional information**

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43247-024-01920-8.

**Correspondence** and requests for materials should be addressed to Kaitlin E. Frasier.

**Peer review information** *Communications Earth & Environment* thanks Alexandra Constaratas, Francesco Caruso and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Alice Drinkwater. [A peer review file is available.]

# Reprints and permissions information is available at http://www.nature.com/reprints

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2024