OCS Study
BOEM 2023-041

Power Analysis for Optimal Design of a Passive Acoustic Monitoring Network for US East Coast Offshore Wind

U.S. Department of the Interior
Bureau of Ocean Energy Management
Sterling, VA

BOEM
Bureau of Ocean Energy Management

OCS Study
BOEM 2023-041

Power Analysis for Optimal Design of a Passive Acoustic Monitoring Network for US East Coast Offshore Wind

July 2023

Authors:
Magda Chudzinska, Len Thomas

Prepared under [Purchase Order 140M0122P0011]
By
Centre for Research into Ecological and Environmental Modelling (CREEM)
University of St Andrews
The Observatory
Buchanan Gardens
St. Andrews
KY16 9LZ
UK

U.S. Department of the Interior
Bureau of Ocean Energy Management
Sterling, VA

BOEM
Bureau of Ocean Energy Management

DISCLAIMER

Study concept, oversight, and funding were provided by the U.S. Department of the Interior, Bureau of Ocean Energy Management (BOEM), Environmental Studies Program, Washington, DC, under Contract Number 140M0122P0011. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the opinions or policies of BOEM, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

REPORT AVAILABILITY

Download a PDF file of this report at https://espis.boem.gov/Final%20Reports/BOEM_2023-041.pdf. To search other studies completed by BOEM's Environmental Studies Program, visit https://www.boem.gov/environment/environmental-studies/environmental-studies-information/

CITATION

Chudzinska M, Thomas L. 2023 (CREEM, St Andrews, UK). Power analysis for optimal design of a passive acoustic monitoring network in the Atlantic Ocean. Sterling (VA): U.S. Department of the Interior, Bureau of Ocean Energy Management. 85 p. Report No.: OCS Study BOEM 2023-041.

ACKNOWLEDGMENTS

This work was partly supported by the Coastal States Stewardship Foundation (CSSF) on behalf of the Regional Wildlife Science Collaborative for Offshore Wind (RWSC), hosted by the Northeast Regional Ocean Council (NROC) and the Mid-Atlantic Regional Council on the Ocean (MARCO). We thank members of the Marine Mammal Subcommittee of the RWSC and BOEM for providing feedback on our planned analyses. We thank Jason Roberts of Duke University for providing the marine mammal density model outputs we used to simulate baseline animal density, Megan Ryder of SMRU Consulting for extracting Automatic Identification System (AIS) data, Jason Wood from SMRU Consulting for providing feedback. We also would like to thank Saana Isojunno and Cornelia Oedekoven from CREEM for discussions on the analytical methods.

Executive Summary

The goal of this project was to provide analyses and recommendations to support the optimal design of a baleen whale passive acoustic monitoring (PAM) network along the U.S. East Coast. We used computer simulation to perform a statistical power analysis and determine the power to detect biologically realistic changes in whale distribution and behavior associated with construction and operation of wind farms using three candidate designs. The three designs were: 1) a recommended regional PAM monitoring design provided by Van Parijs et al. (2021, doi:10.3389/fmars.2021.760840) consisting of a 20 x 20 km grid of sensors ("small PAM grid") located around wind energy areas (WEAs) and a 40 x 40 km grid ("large PAM grid") between the WEAs; 2) a modified design with a 10 x 10 km grid replacing the small PAM grid; and 3) a second modified design with a linear array of PAM stations in a T-configuration (i.e., with three "arms") centered on each WEA and denser sensor placement towards the center of the T, again replacing the small PAM grid. In all cases, each PAM station—which was assumed to consist of a bottom-mounted archival recorder and subsequent acoustic processing to derive counts of detected vocalizations per unit time—was assumed to be done independently at each station (so, for example, the possibility of localizing calls was not considered).

Statistical power is the probability of obtaining a statistically significant result in a hypothesis test, given that some specified change exists. We set a target power of 80% and a nominal false-positive rate (i.e., probability of detecting a change if none exists) of 5%. We chose four study species: fin, sei, minke, and North Atlantic right whale (NARW). We focused on three example WEAs, approximately evenly distributed from north to south: Vineyard Wind 1 Area (VYWA), Empire Offshore Wind Area (EOWA), and Maryland Offshore US Wind Area (MAWA). Power analyses were also undertaken for Coastal Virginia Offshore Wind Area (VOWA) and are reported in a companion report. We undertook local analyses, focusing just on the area around each WEA, and also regional analyses with the latter just looking at operational scenarios. For local analyses, we determined power using just the sensors within 20 km of each WEA ("small monitoring area") and within a larger circle around each WEA out as far as the continental shelf ("large monitoring area").

We simulated acoustic detection rate data by simulating from spatially and temporally referenced whale density surfaces provided by Duke University at the spatial scale of 5 x 5 km and temporal scale of a month, applying the hypothesized changes due to wind farm construction or operation, and converting the simulated whale numbers to acoustic detection numbers using assumed average animal vocalization rates and effective detection ranges. We applied a statistical test to the simulated data, looking to see whether the relationship between acoustic detection rate and distance from wind farm changed between baseline data (collected over 1 year) and data collected during construction (1 year, only construction months) or operation (1 year or 5 years). We repeated the simulation 500 times for each hypothesis, species, and WEA (for local analyses) and calculated power as the proportion of simulations that yielded a statistically significant result.

We specified eight hypotheses on how whales of the study species may respond to construction and operation of wind farms. These could be broadly grouped into four categories:

1. Construction and operation of wind farms had no effect on whale distribution and acoustic behavior (H1 and H8). In H1, there was no systematic change in whale density or distribution over time; in H8 there was a region-wide decline, but unrelated to wind farms.
2. Construction of wind farms caused a change in acoustic behavior (cue production rate) (H2). The change was to be strongest close to the construction location and declined with increasing range.
3. Construction of wind farms caused temporary displacement of whales (H3–H5). Again, the effect was strongest close to the construction location; the hypotheses differ in whether the direction of displacement is independent of habitat preference (H3), dependent on habitat preference (H4), or

dependent on habitat preference but tempered by other construction-related activities (such as shipping) (H5).
4. Operation of wind farms caused long-term change in whale distribution (either attraction, H6, or displacement, H7).

For hypotheses involving response to construction, we used two assumed dose-response functions, taken from a separate expert elicitation on the possible response of NARW to pile driving combined with a simple sound propagation model. In one dose-response function (DR1) corresponding with the assumption that whales are less sensitive, there was a sharp decline in probability of response within the first 2 km from the source and reaching 1% at 18 km; in the other function (DR2) corresponding with the assumption that whales are more sensitive, the probability of response declined gradually, reaching 1% at 30 km.

Hypothesis H1 specifies no effect of wind farms and hence any significant result in the power analysis under this hypothesis is a false positive. We found that false-positive rates were higher than the nominal level of 5%, ranging from 5–24%, but that false-positive rate was lower when larger monitoring areas were used, and was lower under the T-design than the other designs.

Hypotheses H2–H5 involved effects of construction, using DR1 and DR2. We found that power under the Van Parijs et al. design was generally low (i.e., below the target of 80%) for all species and these hypotheses. This was largely because the effect size, i.e., the proportion of animals monitored that responded, was small. Power was higher under the more sensitive dose-response function (DR2), but the difference was not great. Power was higher if only the sensors within 20 km of each WEA were used (the "small monitoring area"), because this is focused on the area where the effect is strongest. (Note that sensors from the small monitoring area can be from the 20 x 20 km or 40 x 40 km grids.) Power was particularly low for fin whales because they can be detected over large distances, and so even sensors close to the sound source detect a mixture of responding and non-responding whales. Minke whales had the highest power, likely partly because they were only detectable over short ranges but despite this had a higher acoustic encounter rate than other species (except fin whale). Sei whales and NARW typically had power values between fin and minke whales. Power to detect displacement responses (H3–5) was slightly higher than an acoustic-only response (H2).

Power was also low under the Van Parijs et al. design to detect changes due to operation (H6 and H7) at the level of an individual wind farm; power was higher at the regional level (i.e., combining all wind farms), and was above the 80% threshold for minke whales but not any other species for this level. Increasing the number of years of operational monitoring from 1 to 5 years increased power, but power still did not exceed 80% for the other three species in regional analysis.

The alternative monitoring designs both resulted in substantially higher power to detect effects of construction and operation, because sampling effort was concentrated closer to the wind farm footprint where change was greatest. Power was higher for the T-design than the 10 x 10 km grid design, with high power (> 80%) to detect change for almost all hypotheses related to operation (H6–H7) and construction (H2–H5) for all WEAs for minke whales, and for sei and NARW at the WEA where the two species were most abundant (VYWA).

Power to detect a region-wide decline (H8) was high for most species and WEAs under all designs; this trend shows the potential for PAM as a population monitoring tool (although this would require additional information about spatial and temporal patterns of vocalization rate and acoustic detectability). However, we did not explicitly examine power to detect long-term trends.

Our results are strongly contingent on the assumptions made, and we suggest future studies that could be undertaken to improve the reliability and scope of the analysis. These studies include looking at finer

temporal scales (which would require example acoustic data on which to base simulations) and examining alternative analysis methods.

Based on our results, we recommend replacing the 20 x 20 km small monitoring grid of sensors around WEAs with an alternative array that concentrates sensors where a response is expected and distributes sensors relatively evenly across the WEAs that are to be used as study sites. Of the designs we tested, the T-design appears better than a 10 x 10 km grid of sensors, but other linear designs with closer sensor spacing nearer the center of each wind farm are possible.

There is in additional need for sensors at distances from the WEAs where no response is expected, and this role could be fulfilled by the use of the entire 40 x 40 km grid. Monitoring over a larger area reduced the false-positive rate.

To maximize the sample size of acoustic sensors, we recommend pooling resources across stakeholders who are deploying sensors. Power will also likely be higher if analyses of construction effect were pooled across WEAs, and this is one of the additional investigations we recommend.

For species like fin whales with large acoustic detection distances, consideration should be given to localizing calls and undertaking effects analysis using the localizations.

One method to improve power is to accept a higher false-positive detection rate (i.e., “alpha-level” for the statistical tests). We used a nominal false-positive rate of 5% and a target power of 80%, but these values are conventions and consideration could be given to using other values.

Contents

List of Figures

List of Tables

List of Abbreviations and Acronyms

Short Form	Long Form
AIS	Automatic Identification System
BOEM	Bureau of Ocean Energy Management
BRS	Behavioral Response Study
CREEM	Centre for Research into Ecological and Environmental Modelling
DR	dose-response
EDA	effective detection area
EOWA	Empire Offshore Wind Area
EDR	effective detection radius
GAM	generalized additive model
MAWA	Maryland Offshore US Wind Area
MGEL	Marine Geospatial Ecology Laboratory
NARW	North Atlantic right whale
PAM	passive acoustic monitoring
PG	phase-gradient
RWSC	Regional Wildlife Science Collaborative for Offshore Wind
SPL	sound pressure level
U.S., US	United States
VOWA	Coastal Virginia Offshore Wind Area
VYWA	Vineyard Wind 1 Area
WEA	Wind Energy Area

1 Introduction

1.1 Aims of the Project

The overall aim of the project is to provide the Regional Wildlife Science Collaborative for Offshore Wind (RWSC) Marine Mammal Subcommittee with analyses and recommendations to support the optimal design of a baleen whale passive acoustic monitoring (PAM) network along the U.S. East Coast. A recommended regional PAM monitoring design for the Northeast U.S. offshore wind energy areas (WEAs) was provided by Van Parijs et al. (2021) (see Figure SI-2 in that paper, referred to here as the "Van Parijs et al. design"). The design is based on a network of PAM stations at two spatial resolutions: a 20 x 20 km grid around the WEAs (referred to here as the "small PAM grid"), and a 40 x 40 km grid between the WEAs (referred to as "large PAM grid"). Each PAM station comprises a single hydrophone bottom-mounted archival recorder. Here we present results, based on computer simulation, of the statistical power of this proposed design to detect biologically realistic changes in simulated whale distribution and behavior associated with construction and operation of wind farms within the WEAs. We additionally evaluate two alternative designs for each WEA: (1) a 10 x 10 km small PAM grid, and (2) a linear array of PAM stations in a T-configuration (i.e., with three "arms") centered on each development area ("T-design").

We focus on the area between northern North Carolina and Cape Cod (referred to as the "regional study site" or "study area"). Although development of offshore wind is planned along the entire U.S. East Coast, the data on baseline density of baleen whales in a format necessary for this analysis were only available for the North Carolina to Cape Cod area. We chose as local study sites three WEAs, approximately evenly distributed along the East Coast: Vineyard Wind 1 Area (VYWA), Empire Offshore Wind Area (EOWA), Maryland Offshore US Wind Area (MAWA). Additionally, power analysis was undertaken for Coastal Virginia Offshore Wind Area (VOWA) and results of that analysis are presented in Chudzinska et al. (2023). Figure 1A shows these sites and the Van Parijs et al. design, and Table 1 gives the schedule (at time of writing of this report) of construction for these wind farms.

In a set of *local analyses*, we estimate power to detect hypothesized changes in acoustic detections associated with wind farm construction and operation separately for each WEA. We also undertake *regional analyses* (Figure 1B) to estimate power assuming all 27 currently proposed wind farms are operating simultaneously (NROC 2009). In all cases, the spatial resolution of our computer simulation is 5 x 5 km and the temporal resolution is one month—this resolution is based on the resolution of the input density surfaces we used. We assume that the acoustic data from each PAM station is processed independently to produce a count of detected vocalizations of the species of interest, and this forms the input data for analysis. We do not consider in our power analysis any additional inferences that could be drawn from a joint acoustic analysis of the sensor array—for example, localization—but we do return to this in the Section 4 (Discussion).

Statistical power is the probability of obtaining a statistically significant result in a hypothesis test, given that some specified change exists (Steidl and Thomas 2001). We follow convention (Cohen 1988) in setting a target power of 80% or greater. It is proportional to sample size, nominal false-positive rate ("alpha-level"), and effect size (i.e., level of change that exists) and inversely proportional to the variance in the quantity being tested. The nominal false-positive rate, also called the significance criterion or α-level, is the *p*-value of the significance test below which the test is deemed to be statistically significant. If the effect size is zero, the statistical test should yield a statistically significant result (i.e., a false positive) $\alpha \times 100\%$ of the time, although in practice the true false-positive rate can be higher in real-world

situations, such as when assumptions of the test are violated. We again follow convention in setting an α-level of 0.05, i.e., a nominal false-positive rate of 5%.

The remainder of this document is laid out as follows. In th Introduction, we provide a brief overview of the spatial ecology of baleen whale species along the U.S. East Coast and a summary of the observed effects of disturbance related to offshore wind farm development. We use this information to develop hypotheses for what changes might occur, including also possible changes not related to offshore wind farms. These hypotheses of change provide the basis for our simulation study, focusing on the acoustic signal generated by the change. We give an overview of issues to be considered in using PAMs to detect such changes. In the Methods section, we provide an overview of the power analysis and give details on each step of the analysis. We then present Results and a Discussion of the results, the caveats accompanying them, and our recommendations.

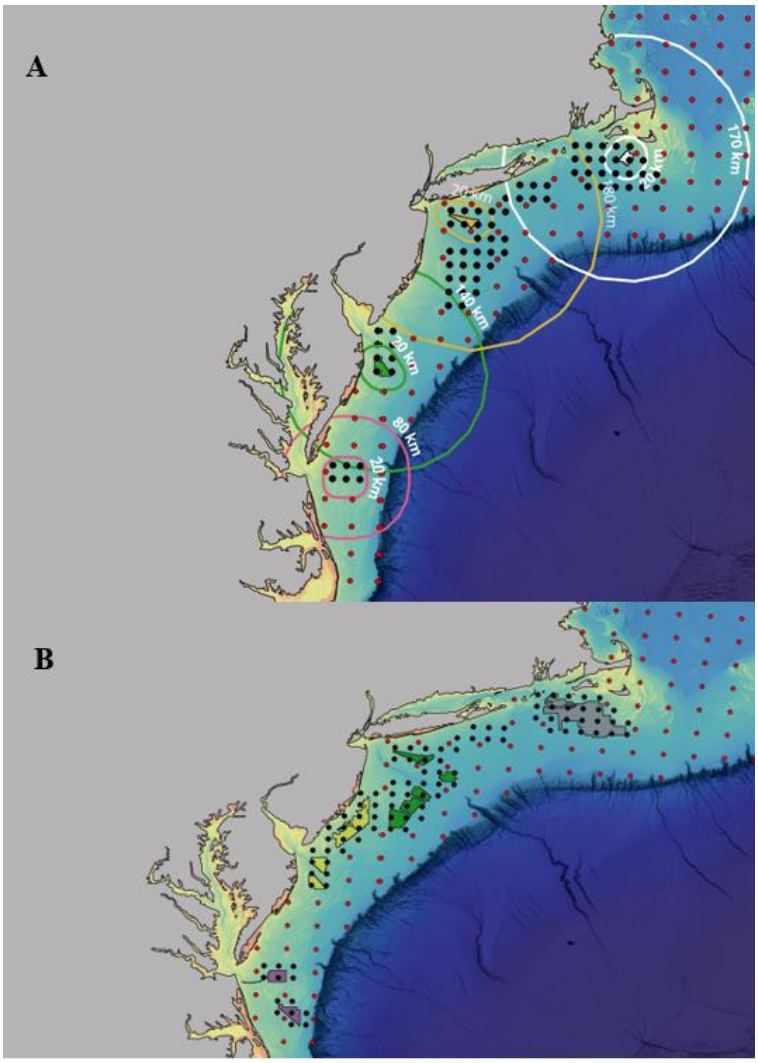


Figure 1. Wind energy areas between North Carolina and Cod
A) The three wind energy areas studied (filled polygons), with monitoring areas around them (lines) and two PAM grids (filled dots). The wind energy areas are, from north to south: Vineyard Wind 1 Area (VYWA, shown in white), Empire Offshore Wind Area (EOWA, shown in yellow), Maryland Offshore US Wind Area (MAWA, shown in green). Also shown, further south, is the Coastal Virginia Offshore Wind Area (VOWA, shown in pink) for which a power analysis was presented in a separate report. Each wind energy area has two sizes of monitoring areas: 20 km buffer and larger buffer spanning over the continental shelf. The two PAM grids (Van Parijs et al. 2021) have a 40 km grid

spacing (red dots) and a 20 km spacing (black dots, only shown around the chosen wind energy area sites). B) All wind energy areas that were used for regional analysis, with all PAM sensors from the Van Parijs et al. design that were considered: 40 km grid spacing (red dots) and a 20 km spacing (black dots). Wind farms considered in the regional analysis were grouped into four clusters (shown in matching colors) based on the close proximity of the individual wind farms.

1.2 Baleen Whales Along the East Coast of the U.S.

Six baleen whale species are found in the western North Atlantic along the U.S. East Coast: minke (*Balaenoptera acutorostrata*), North Atlantic right (NARW; *Eubalaena glacialis*), humpback (*Megaptera novaeangliae*), sei (*Balaenoptera borealis*), fin (*Balaenoptera physalus*), and blue (*Balaenoptera musculus*) whales. These are mainly migratory species moving annually between winter breeding grounds and summer foraging areas.

Minke whales are the most abundant baleen whale species in the study area (between northern North Carolina and Cape Cod) (Roberts and Yack 2022a, Roberts et al. 2022). Minke whales breed south of the study area in the winter, and their main summer foraging ground is north of the study area, but they are present at the study area between April and October (Risch et al. 2013, Risch et al. 2014).

NARW is one of the least abundant species in the study area. Although their main foraging grounds are north of the study area, NARW have been observed in the study area between January and June (Whitt et al. 2013, Davis et al. 2017, Quintana-Rizzo et al. 2021) more frequently since 2010 after a large-scale shift in the occurrence of this species was observed (Meyer-Gutbrod et al. 2018). Within the study area, waters off southern New England are the main foraging ground for NARW. They calve off South Carolina, Georgia, and northeastern Florida between late November and early March (Kenney et al. 2020).

Like minke whales, humpback whales migrate between their breeding grounds south of the study area in the winter and foraging ground north of the study area in the summer, but some individuals are present year-round in the study area (Katona and Beard 1990, Vu et al. 2012, Davis et al. 2020), especially juveniles (Barco et al. 2002, Aschettino et al. 2020, Stepanuk et al. 2021). Humpback whales return from their breeding ground to the study area from June and leave the area to breed around November (Katona and Beard 1990), but their timing of migration is subject to large interannual variation (Salisbury et al. 2018, Davis et al. 2020).

Migratory movements of sei whales are the least understood of the six baleen whale species, but they are believed to move northward from southern New England in June and July to eastern Canada, with a return southward in September and October (Mitchell 1975, CETAP 1982). Their range is therefore mainly north of the study area. Recent data collected by acoustic monitoring revealed the presence of sei whales at the northern part of the regional study site (Southern New England and New York Bight) between March and July (Davis et al. 2020), although they are also observed more widely closer to the shelf break (PACM 2023).

Fin whales are the second most abundant species in the study area (Roberts et al. 2022). This species does not undergo as large a seasonal migration as humpback whales, but their foraging grounds overlap. Fin whales are observed all year within the study area, mainly in the New York Bight, with lowest densities between May and July (Morano et al. 2012, Davis et al. 2020, Roberts et al. 2022).

Blue whales are mainly observed north of the study area and in off-shelf waters but can be present in the northern part of the study area (New York Bight) in winter months (Muirhead et al. 2018, Davis et al. 2020, Zoidis et al. 2021).

In terms of conservation status, on the International Union for Conservation of Nature Red List, fin whales are listed as vulnerable, sei and blue whales as endangered, and NARW as critically endangered.

There is a large variation in reported measured or modeled PAM detection ranges (Section 2.5) and acoustic behavior (cue rate) of the six baleen whale species. Fin whales have the largest reported detection range (> 120 km (Salisbury et al. 2018, Estabrook et al. 2021)) and sei, minke, and North Atlantic right whales the smallest, although there is large variation for these species reported between WEAs (Salisbury et al. 2018). As the detection range for each species may vary over space, time, and oceanographic properties, we use estimations of detection range closest to or within the study area; we also require more information than just mean or maximum detection range, and use only reports that provide at least three percentiles for detection (e.g., ranges at which 5, 50, and 95% of calls are estimated to be detected, see details in Methods section) (Salisbury et al. 2018, Estabrook et al. 2021). Studies estimating detection ranges from other areas or not providing percentiles (e.g. Gervaise et al. 2019, Kowarski et al. 2020, Palmer et al. 2022) were therefore not used here. There is also a large variation in cue production rate (rates at which individuals produce acoustic signals) between species (Fregosi et al. 2022). For cue production rate, fin whales have been estimated to produce on average 45 cues/h (Stimpert et al. 2015), and minke whales and NARW only ~ 6 cues/h (Parks et al. 2011, Martin et al. 2013); however, information for fin, sei and minke whales come from the Pacific region. Apart from NARW, no information on call rates from the East Coast is available for minke, sei, fin, and blue whales.

After consulting with BOEM and RWSC, four baleen species were chosen for analysis: North Atlantic right, fin, minke, and sei whales. Fin and minke whales are the most abundant species and are at the two extremes in terms of detection ranges and cue production rates; fin whales are also present year-round at some of the chosen WEAs. NARW and sei whales are less abundant species, but NARW are of highest conservation concern.

1.3 Potential Effects of Offshore Wind Development on Behavior of Cetaceans

The majority of studies describing the effects of construction and operation of wind farms on cetaceans come from Europe, where small odontocetes (dolphins and porpoise mainly) have been the main focus. Studies have reported displacement effects at ranges up to 10–26 km from the footprint of offshore wind farms during construction (Dähne et al. 2013, Brandt et al. 2016, Dähne et al. 2017, Brandt et al. 2018, Graham et al. 2019, Benhemma-Le Gall et al. 2021, Graham et al. 2023). Together, these studies indicate that the distance and duration of avoidance may be related to habitat quality, received noise level (which in turn is a function of source level and sound propagation conditions), hearing characteristics of the studied species, distance to the noise source, duration of exposure, level and type of mitigation, and presence of other noise sources like construction vessels. There is variation among studies in the reported duration of effect, from hours (Dähne et al. 2017) and days (Brandt et al. 2018) to years (Teilmann and Carstensen 2012), which also suggest that operation of offshore wind farms may affect cetacean behavior, including distribution. These large differences are hypothesized to be mainly driven by the habitat quality and noise characteristics at the area during construction and operation.

The acoustic frequency range used by baleen whales to communicate, and presumed hearing sensitivity, overlaps more with frequencies produced by pile driving (Pyć et al. (2018), see Fig. 3 in Van Parijs et al. (2021)) than in the case of odontocetes. Little is known about the effect of pile-driving noise on baleen whales, and we are not aware of any empirical data on behavioral response to pile driving in this group; the effects studied for odontocetes may not be directly transferable to baleen whale species. However, some informed estimates can be made based on understanding of their hearing, response to other low-frequency noise sources (e.g., seismic or vessel), and models or auditory masking (Hatch et al. 2012, Erbe et al. 2016).

We suggest eight hypotheses, listed in Table 2A and 2B, on how simulated whales of the study species may respond to construction and operation of wind farms, and how this may affect acoustic detections. We also outline what additional data (in addition to PAM) or analyses maybe required to distinguish between the hypotheses, although this is not the focus of the current report. Broadly, the hypotheses can be grouped into four categories, as follows.

1. Construction and operation of wind farms has no effect on whale distribution and acoustic behavior (H1 and H8). In H1, there is no systematic change in whale density or distribution over time; in H8, there is a region-wide decline, but it is unrelated to wind farms.
2. Construction of wind farms causes a change in acoustic behavior (cue production rate) (H2). The change will be strongest close to the construction location and decline with increasing range.
3. Construction of wind farms causes temporary displacement of whales (H3–H5). Again, the effect will be strongest close to the construction location; the hypotheses differ as to whether the direction of displacement is dependent on habitat preference or other construction-related activities (such as vessel traffic).
4. Operation of wind farms causes long-term change in whale distribution (H6–H7).

In cases where we hypothesize that construction of wind farms results in behavioral response (either displacement or change in cue rate), to calculate the number of animals responding, we multiply the number of animals at each distance from the source (here derived from animal density maps such as Roberts et al. (2022)) by the probability that each animals will respond, obtained from an assumed dose-response function (Tyack and Thomas 2019). If the dose-response function uses received noise level rather than distance as the dose metric, then the range-specific received level must be estimated (e.g., using a sound propagation model). To our knowledge, no empirical dose-response function has been derived for a baleen whale species' response to construction of offshore wind farms. However, an interim function for NARW was obtained using expert elicitation in December 2022 by a team from the University of St Andrews as part of a BOEM-funded project "Assessing Population Effects of Offshore Wind Development on North Atlantic Right Whales." A distribution of functions was elicited, representing scientific uncertainty in the dose-response function; we use the first and third quartiles from this distribution (less and more sensitive, respectively, see Methods for details) and assume these apply to all species.

We note that baleen whale dose-response functions have been developed for other noise sources (e.g. Sivle et al. 2015, Dunlop et al. 2017, Dunlop et al. 2018, Dunlop et al. 2020), and indeed this information was part of the input to the expert elicitation. However, in many cases the data have been summarized in the form of a received level at which probability of response is a given value, typically 50%. For example, 50% probability of inducing behavioral responses may be expected at a received sound pressure level (SPL) of 140–160 dB re 1 µPa based on the responses of migrating grey whales to airguns (Malme et al. 1984, Wood et al. 2012, see also summary in Pyć et al. 2018). The use of binary 50% thresholds to estimate zones of impact has been criticized by Tyack and Thomas (2019), who showed that it can lead to a large underestimation of the number of animals responding. They advocate the use of a dose-response function or, if a single threshold value is preferred, then it should be based on the concept of an effective response range (see paper for details). Nevertheless, 50% thresholds may provide a helpful frame of reference. Based on a literature review and sound propagation modeling, Pyć et al. (2018) predicted 50% probability of response for baleen whale species in the Vineyard Wind 1 Area WEA to be at 2–7 km from turbine installation, depending on exposure scenario, foundation type and attenuation level, and assuming a step function with certain response at a frequency-weighted SPL of 140 dB (see Tables 10 and 13 in Pyć et al. 2018). Depending on the scenario (number of pilings per day and attenuation level), no more than 1 to 40 individuals (depending on the baleen whales species) would be estimated to experience sound levels above the threshold criteria, which is no more than 4% of the population of a given species (Pyć et al. 2018). The recommended distance for monitoring and mitigation during Vineyard Wind 1 construction for NARW was 10 km (Pyć et al. 2018).

For hypotheses where the response is displacement only, once the number of displaced animals is estimated, we will also consider the location to which the displaced animals move. Direct evidence of this is (to our knowledge) lacking, so we considered a range of hypotheses. We set an upper limit on the displacement distance and within that distributed responding animals uniformly ("symmetrical displacement", H3), according to the underlying habitat model (H4), or according to the habitat model but also accounting for patterns of vessel traffic (H5).Table 2

Construction of wind farms is not limited only to foundation installation: a range of activities happen before and after this installation, such as cable or turbine installation. In this project, we focus only on periods when foundations are installed (referred to as 'construction' throughout the text, Table 1), as we are not aware of any data on the effect of other activities on baleen whales. Piles can be installed in a variety of ways: impact hammering, vibrating piling, or drilling; a mix of the methods can be used for a given wind farm. The presented study focuses on impact pile driving as this is the loudest of the methods. We also consider hypotheses based around long-term operation.

Table 1. Assumed timing of construction for the studied WEAs and corresponding months of baseline monitoring

WEA	Construction Timing	Baseline Monitoring	Size of the Large Monitoring Area (km)
MAWA	April 2025–September 2026 April 2027–September 2027	April–September	140
EOWA	April 2025–October 2025 April 2026–December 2026	April–December	180
VYWA	May–October 2023	May–October	170

1.4 Challenges of Using PAM to Detect Changes in Behavior and Distribution of Baleen Whales and the Effect of These Challenges on Quantifying Power

Responses to anthropogenic disturbance, including noise, have been quantified using Behavioral Response Studies (BRSs). Such studies either involve controlled exposure experiments, where animals are exposed to a controlled level of a potential stressor, or in opportunistic contexts where exposure and concurrent activities are monitored in a strategic manner (see Harris et al. 2018 for a review on the context of disturbance from naval sonar). A variety of animal observation techniques have been employed, some focused on measuring the response of individual animals (e.g., animal-borne tags, ship-based or aerial focal follows, acoustic tracking arrays) and others on changes in occurrence or density of animals at the population level (e.g., visual surveys, PAM). Often, the approaches are complementary—for example, Tyack (2011) used results from both a controlled exposure experiment on a small number of individuals and a large-scale opportunistic population-level PAM study to infer an avoidance response by Blainville's beaked whales to naval sonar exercises.

We here focus on the use of a fixed array of archival PAM sensors to perform population-level inference on behavioral response. We note that other types of PAM systems exist (e.g., glider-mounted, towed or buoyed; real-time vs archival) and can be used for mitigation and monitoring in the context of wind energy development—see for example the review by Van Parijs et al. (2021). Archival data is post-processed once the recordings are recovered, and detections of vocalizations from the target species are identified. If each sensor is analyzed independently then the resulting data are a count of detections per unit of monitoring time. If sensors are close enough together that the same vocalizations can be detected

on multiple sensors, it may in addition be possible to localize animals making the sounds, allowing a finer level of inference about any possible behavioral change. Further, under some circumstances it may be possible to track individual vocalizing animals, allowing for individual level responses to be studied (see e.g. Durbach et al. 2021).

Returning to the independent sensor analysis, changes in detection rate associated with wind farm construction or operation could be caused by several factors: changes in detectability of vocalizations (e.g., via masking of the sounds), change in the frequency of false-positive detections, change in animal vocalization behavior, or changes in local animal distribution. Additional studies may be required to exclude the first two factors, although any masking should only occur while anthropogenic sounds are being made, and biological changes may be more long term. Distinguishing between changes in vocalization behavior and changes in animal distribution also requires additional data collection—but both constitute a behavioral response that are of interest to detect. However, detecting such a change on a single sensor concurrent with wind farm construction or operation is not enough to infer the wind farm has caused the change; any change may be part of some larger-scale process that is independent of the wind farm activities. Therefore, it is crucial to monitor at a range of distances from the wind farm site; a change in detection rates close to wind farm activities that does not occur at further distances is much stronger evidence that the change is caused by the activities. Detecting an interaction between change in detection rate and distance from wind farm, while accounting for other factors such as monthly changes in detection rate in the baseline data, is the core of the statistical test used in the power analysis.

The density of the four baleen species chosen for this analysis is generally low. In particular, for NARW, construction of some of the wind farms is scheduled to happen in the months when they are at seasonally low density (Pyć et al. 2018). As a result, the number of whales responding to construction (or operation) of wind farms is also going to be low. For a given statistical test and chosen α-level, statistical power increases with increasing effect size (i.e., true magnitude of change) and sample size, and decreases with increasing variability. For the statistical test evaluated here, effect size is related to the proportion of whales responding within the area monitored by the PAM devices, rather than the absolute number responding. Hence the low number responding will not necessarily cause low power if the proportion responding is consistent. However, low numbers may lead to high variability over time and space, which may result in low power. In addition, in some cases there may be no whales at all in the area, giving an effect size of zero.

Variability in baseline density in space and time will tend to decrease power unless it is dealt with as part of the analysis. The highest density of the studied species generally is along the continental shelf and in the areas around the Cape Cod peninsula (see for example Roberts et al. 2016). There is, therefore, a baseline gradient in whale distribution for wind farms that are far from continental shelf (MAWA) or from density hot spots, like VYWA being east from Nantucket Shoals. This gradient is accounted for in the analysis method used (which models acoustic encounter rate under baseline conditions as a function of distance from each wind farm) and so should not cause either decreased power or elevated false-positive rate. Variability in time is partly accounted for by including month as a factor in the model.

Another issue that could potentially affect power is that some species (fin whale in particular) can be detected at a considerable distance from the acoustic sensors. In this case, if response to wind farms only occurs relatively close to the wind farm footprint, then sensors placed near the wind farm boundary will detect a mix of responding and non-responding animals. This will dilute the measured effect size and reduce power. We return to this in the Discussion.

As well as quantifying statistical power given a specified change, it is important to quantify the false-positive rate—i.e., the probability of detecting a statistically significant effect when none exists. We used a nominal α-level of 0.05, which should result in a 5% false-positive rate. However, for complex statistical tests like those used here, the false-positive rate can be different from the nominal value. In

addition, it may be that random variation in animal density over time produces a change in distribution with respect to distance from a wind farm and so triggers a positive significance test. To evaluate false-positive rate, we include a hypothesis of no change (H1) in our test suite.

Table 2A. Summary of potential drivers, acoustic effect of these drivers, data additional to PAM data required to study the driver and additional reading supporting the driver related to the construction and operation of offshore wind farms

#	Drivers and Hypothesis	Effect Observed in PAM Data	Additional Data or Analysis Required to Distinguish from Other Hypotheses	Additional Reading	Tested Scenarios and Methods
1	Construction and operation activities have no effect on baleen whale distribution or behavior.	There are no changes in acoustic encounter rates with distance from wind farm between construction and operation versus the same area before construction (referred to hereafter as “over time”).	Visual surveys and tagging to confirm no changes in behavior, such as foraging or group behavior. Studies could additionally undertake analysis of stress hormones to understand whether there is a physiological response, even if no behavioral response is observed.	Bailey et al. (2010)	Simulation: no redistribution of whales or changes in cue rate will be generated for all the wind farms. Analysis: examine the relationship between acoustic encounter rate and distance to the wind farm under baseline and under construction and operation. Under this hypothesis, the relationship should not differ. The proportion of times a significant difference is found is an estimate of the false-positive rate, i.e., the probability of detecting a change if the change is not present. Analysis to be undertaken on a per wind farm basis, and also regionally.

#	Drivers and Hypothesis	Effect Observed in PAM Data	Additional Data or Analysis Required to Distinguish from Other Hypotheses	Additional Reading	Tested Scenarios and Methods
2	Construction activities have no effect on baleen whale distribution but have an effect on their behavior related to cue production.	We would observe a change in acoustic encounter rate over time with distance from wind farm (compared with the baseline pattern), where the change is driven by change in cue production or cue detection rate (see next column) and not by changes in distribution. As pile driving mainly occurs during the day (Heaney et al. 2020), an increase in cue rate may be observed at night as a function of time since last pile-driving event (although diurnal patterns in baseline would need to be considered). A decrease in encounter rate of acoustic detections could also result from decrease in cue detection rate due to masking (Erbe et al. 2016, Cholewiak et al. 2018). Comparing data collected during actual piling with in-between piling is needed to distinguish between these two options.	Additional data, e.g., a visual survey or tagging study, would be needed to confirm that there is no displacement.	Benhemma-Le Gall et al. (2021)	Simulation: the number of responding whales will be calculated based on a dose-response function. We then simulate 1) 100% decrease in cue rate of the responding whales (referred to as "H2_100"); 2) 50% decrease in the cue rate of the responding whales simulating either masking or changes in cue production occurring only during the day or a partial response (referred to as "H2_50"). No whale redistribution is assumed for this hypothesis. Simulation to be done for all wind farms. Analysis: examine the relationship between acoustic encounter rate and distance to the wind farm under baseline and under construction. The proportion of times a significant difference is found is an estimate of power.

#	Drivers and Hypothesis	Effect Observed in PAM Data	Additional Data or Analysis Required to Distinguish from Other Hypotheses	Additional Reading	Tested Scenarios and Methods
3	Construction activities cause displacement of whales away from the construction locations, with displacement occurring equally in all directions. There is no change in cue production of individual whales.	We would observe a decrease in acoustic encounter rate over time with distance from wind farm, compared to the baseline pattern.	Visual survey or tagging study to confirm that this is animal displacement and not a change in cue detection/production rate. Animals may increase their foraging effort when displaced from the wind farm to compensate for the lost foraging time when moving away from the site (Benhemma-Le Gall et al. 2021). In such a case we may expect an increase in foraging activity further away from the construction which could be confirmed by tagging. Note that this hypothesis is not considered very likely to be correct, as whales can be displaced to areas where they were previously not known to occur.	Kraus et al. (2019) Pyć et al. (2018) Benhemma-Le Gall et al. (2021) Sivle et al. (2016)	Simulation: the number of responding whales will be calculated based on dose-response functions. We then simulate displacement of responding whales from the footprint of the wind farms with equal probability of displacement in all directions. Simulation to be done for all wind farms. Analysis: as for H2.
4	Construction activities cause displacement of whales away from activity locations preferentially towards higher density areas outside wind farm.	We would observe a decrease in acoustic encounter rate over time with distance from wind farm (compared to baseline pattern) with a corresponding increase being proportional to the observed baseline densities in the areas around the wind farm.	Same as H3.	Davis et al. (2020) Salisbury et al. (2018) Roberts et al. (2016) BOEM and NOAA (2022) Rolland et al. (2016) Ellison et al. (2012)	Simulation: the number of responding whales calculated as for H3. Displacement locations will be proportional to baseline density. Simulation to be done for all wind farms. Analysis: as for H2. Note this means we will test for displacement as a function of distance. In this report, we do not additionally seek to distinguish between H3 and H4, but this could be done by fitting an acoustic encounter density surface to the data and looking for an increase in detections proportional to baseline density (or proportional to baseline animal density as estimated by e.g., Roberts et al. (2016)).

#	Drivers and Hypothesis	Effect Observed in PAM Data	Additional Data or Analysis Required to Distinguish from Other Hypotheses	Additional Reading	Tested Scenarios and Methods
5	Construction activities cause displacement of whales away from activity locations, preferentially towards higher density areas outside of the wind farm, but this preference is lessened by additional anthropogenic activities, such as shipping, that are associated with construction but take place away from the piling locations.	We would observe a decrease in acoustic encounter rate over time with distance from wind farm, compared to baseline pattern. The corresponding increase in density at nearby sites would be a function of baseline density and additional anthropogenic activity.	Same as H3, plus data on other anthropogenic activity: vessel traffic, etc.	www.northeastoceandata.org globalfishingwatch.org marinecadastre.gov/accessais/	Simulation: the number of responding whales calculated as for H3. Displacement locations will be proportional to baseline density, modified according to an anthropogenic pressure map. This map will be generated based on AIS data scaled from 0 to 1 with 0 meaning high anthropogenic pressure. Simulation to be done for all wind farms. Analysis: as for H2. Note, as for H4, this means we do not seek to distinguish between H3, H4, or this hypothesis.
6	Operating wind farms attract whales due to formation of “artificial reefs.”	We would observe an increase in acoustic encounter rate in the wind farm footprint during operation in comparison to that area before construction. The effect may differ for planktivorous and piscivorous whales.	Monitoring of prey composition and distribution within and outside wind farms.	Fernandez-Betelu et al. (2022) Claisse et al. (2014)	Simulation: wind farms in the known foraging grounds of whales (VYWA, EOWA) attract all focal species at the time when these species are present at these areas. We simulate this by increasing densities for these species by a factor of two by two within the wind farm footprint. For regional analysis, we assumed that all WEAs attract whales. Analysis: as for H2, except that here the focus is on a difference between baseline and operation (not construction). Analysis to be undertaken on a per wind farm basis and regionally.

#	Drivers and Hypothesis	Effect Observed in PAM Data	Additional Data or Analysis Required to Distinguish from Other Hypotheses	Additional Reading	Tested Scenarios and Methods
7	Displacement and alteration of whale behavior during construction leads to long-term displacement of whales. Such long-term displacement may be related to the noise of the operational turbines, anthropogenic activity related to maintenance or changes in prey distribution/behavior at the wind farm.	The displacement would not only occur during construction but also several months after construction stops and operation starts.	Aerial surveys to confirm displacement and distinguish from changes in cue production rate. Monitoring of prey composition and distribution within and outside wind farm could determine if prey changes are driving effect.	Though operational turbine noise levels are generally low and close to ambient, larger turbines may lead to behavioral response in low-frequency specialists, such as baleen whales, within ~1.4 km of turbines (Teilmann and Carstensen 2012, Stöber and Thomsen 2021).	Simulation: to calculate the number of responding whales, density within the wind farm footprint was set to zero; displaced whales were redistributed according to baseline density (as for H4). Displacement took place in all months. Analysis: as for H6. Analysis to be undertaken on a per wind farm basis, and also regionally.
8	Displacement or decline of whales is not linked to activating related to construction or operation of wind farms but it is a large-scale phenomenon related to global changes in environment.	The decrease or displacement of whales would occur over the entire study region.	If PAM is to be used to detect a long-term decline in density or abundance within the region, then additional data collection is needed to determine if detectability and/or cue rate changes over time. Alternative population monitoring such as visual surveys may be used.	Global shift or decline of whales has been observed along the U.S. East Coast, especially for NARW and changes in distribution of their primary food (Ramp et al. 2015, Davis et al. 2017, Meyer-Gutbrod et al. 2018)	Simulation: a region-wide decrease in acoustic encounter rate was simulated that was equal in magnitude to the average decline under H3 within 20 km of a wind farm during construction. Analysis: unlike previous hypotheses, the focus here is on whether the effect of year is statistically significant (and whether the interaction between distance and phase is not).

Table 3B. Summary of tested scenario for each analysis

Hypothe-sis	Phase	Shortened Description of the Simulated Effect	Local: large VP grid + small VP grid	Local: small VP grid only	Local: large VP grid + 10 x 10 km grid	Local: 10 x 10 km grid only	Local: large VP grid + T-design	Local: T-design grid only	Region: large VP grid + small VP grid	Region: small VP grid only	Region: large VP grid + 10 x 10 km grid	Region: 10 x 10 km grid only	Region: large VP grid + T-design	Region: T-design grid only
H1	Construction	No effect – false-positive rate	yes	yes	yes	yes	yes	yes	no	no	no	no	no	no
H1	Operation	No effect – false-positive rate	yes	yes	yes	yes	yes	yes	yes	yes	no	no	yes	yes
H2	Construction	Decrease in detected cues but no displacement	yes	yes	yes	yes	yes	yes	no	no	no	no	no	no
H3	Construction	Displacement of whales occurring equally in all directions away from WEAs	yes	yes	yes	yes	yes	yes	no	no	no	no	no	no
H4	Construction	Displacement of whales towards higher density areas away from WEAs	yes	yes	yes	yes	yes	yes	no	no	no	no	no	no
H5	Construction	Displacement of whales towards higher density areas away from WEAs and away from vessel traffic	yes	yes	yes	yes	yes	yes	no	no	no	no	no	no
H6	Operation	WEAs attract whales	yes	yes	yes	yes	yes	yes	yes	yes	no	no	yes	yes
H7	Operation	Displacement of whales towards higher density areas away from WEAs	yes	yes	yes	yes	yes	yes	yes	yes	no	no	yes	yes
H8	Global	Large-scale, global decline of whales not related to activities at WEAs	yes	yes	yes	yes	yes	yes	no	no	no	no	no	no

Notes: 'VP' refers to Van Parijs et al. design. * Results for VOWA are presented in a separate report (Chudzinska et al. 2023).

2 Methods

2.1 Overview

For the selected study species (fin, minke, and sei whales and NARW), the power analysis was conducted in four steps.

1. Local, at the level of each of the three selected wind farms (VYWA, EOWA and MAWA), based on the PAM design of Van Parijs et al. and examining hypotheses H1–H8.
2. Local, as with Step 1, but using alternative PAM design(s) if Step 1 showed low power to detect change.
3. Regional, assuming all wind farms are constructed and operational, to determine power to detect possible changes due to operation (H6 & H7).
4. Regional, as with Step 3, but using alternative PAM designs if Step 3 showed low power to detect change.

An overview of these steps is given below. All analyses were undertaken using the R statistical software (R Core Team 2022). A summary table listed all tested scenarios is given in Table 2B.

2.1.1 Local Analyses

For Step 1, we treated each of the three wind farms separately. To evaluate hypothesized changes within and around each wind farm, we did not use the entire regional PAM network. Rather, we sought to understand how the power to detect change is affected by the size of the monitoring area around each wind farm and by the PAM density. We defined two sizes of monitoring area: 1) the wind farm footprint plus a 20 km buffer (the "small monitoring area"), and 2) the wind farm footprint plus a buffer large enough to cover the continental shelf ("large monitoring area"). This setup resulted in a gradient of large monitoring area size, with the largest in the north at VYWA (buffer 170 km) and the smallest in the south at MAWA (buffer 80 km). Large monitoring area sizes are given in Table 1 and both large and small monitoring areas are shown in Figure 1A. To examine PAM density, we calculated power using 1) both the large (40 x 40 km spacing) PAM grid and the small (20 x 20 km spacing) PAM grid suggested by Van Parijs et al. and 2) just the small PAM grid. If power is high and false-positive detection rate is low with just the small PAM grid, then the monitoring scheme will be considerably less expensive. The grids are shown in Figure 1A.

The use of two monitoring areas and two PAM densities gives four combinations for each species, wind farm and hypothesis (see Table 2B for summary). However, we did not evaluate the combination of small monitoring area and both PAM grids, because very few large PAM grid sensors fell within the small monitoring areas, so results would be almost identical to the combination of small monitoring area and small PAM grid. We therefore evaluated the following combinations:

- large monitoring area and both PAM grids
- large monitoring area and small PAM grid only
- small monitoring area and small PAM grid only

At the local level, we evaluated all eight hypotheses. For each species and hypothesis, the power analysis proceeded through a series of sub-steps, shown in Figure 2. We first simulated a set of 500 random replicate density surfaces on a 5 x 5 km grid for relevant months and years from habitat-based density models (Sub-step 1). For hypotheses involving construction (H2–H5), we defined the sound source location for each month as a subset of the WEA (Sub-step 2). We then calculated the number of responding simulated whales during construction or operation according to the hypothesis (Sub-step 3)

and, for hypotheses that involved displacement (H3–H8), we generated the appropriate spatial change in simulated whale density (Sub-step 4). We next transformed the simulated whale densities per grid cell into a number of vocalizations ("cues") per month (Sub-step 5) and, for the given PAM design, to the number of cues detected on each PAM sensor (Sub-step 6). These detection numbers formed the input data for a statistical analysis to determine whether the hypothesized effect was detected or not. Analysis involves comparing one year of baseline data (just the months of construction for the construction-related hypotheses) with the same months of data from a construction or operation period. This process was repeated 500 times, once for each replicate simulated dataset, and the proportion of statistically significant results was used as the estimate of power (or false-positive rate for H1 since this hypothesis is that there is no change related to construction or operation) (Sub-step 7). Note that while Sub-step 1 involves stochasticity, all other sub-steps are deterministic.

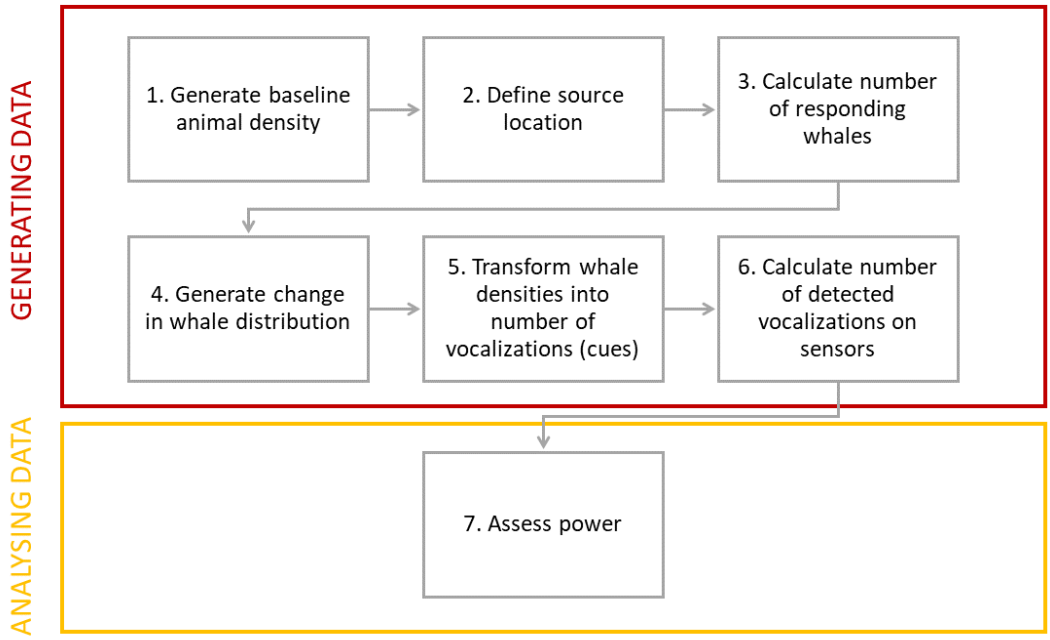


Figure 2. Flow chart of seven sub-steps of the analysis

For construction-related hypotheses, we assumed 1 year of baseline monitoring and 1 year of construction, just in the months where construction activities were planned in each WEA (Table 1). For operation-related hypotheses (H6 and H7), we did two sets of runs: one assuming 1 year of baseline and 1 year of operation, and another assuming 1 year of baseline and 5 years of operation.

After completing Step 1 (i.e., power analysis based on the Van Parijs et al. design) we found that power to detect hypotheses relating to construction and operation was not above the target level of 80% for any scenario. Therefore, in Step 2, we created two alternative PAM designs to replace the 20 x 20 km small grid of Van Parijs et al. in the vicinity of each wind farm: I) a 10 x 10 km grid and II) a linear array of PAM stations in a T-configuration. We repeated the same sub-steps to evaluate power.

2.1.2 Regional Analyses

The aim of Step 3 was to determine the power of the regional monitoring grids under the scenario that all 27 of the planned wind farms had already been constructed and were now operational. We estimated power for the operation-related hypotheses (H6 and H7) under the assumption of data from 1) both the

large (40 x 40 km spacing) PAM grid and the small (20 x 20 km spacing) PAM grid suggested by Van Parijs et al. or 2) just the small PAM grid. We assumed 1 full year of baseline data and either 1 or 5 years of monitoring of the operational wind farms. We used the same sub-steps as outlined above to evaluate power. We also evaluated the false-positive rate using H1, assuming 1 year of baseline and 1 year of monitoring during operation.

We found that power was less than the target of 80% for most cases, and so for these cases in Step 4 we evaluated the power of the 10 x 10 km grid and the T-design, using the same method as for Step 3.

In the following sub-sections, we give details on each of the sub-steps shown in Figure 2.

2.2 Generating Baseline Animal Density

Baseline animal density surfaces were generated using habitat-based marine mammal density models for the U.S. Atlantic (Roberts et al. 2016, Roberts et al. 2023) produced by the Duke University in collaboration with Marine Geospatial Ecology Laboratory (MGEL) of Duke University produced habitat-based marine mammal density models for the U.S. Atlantic. These models were based on analysis of aerial and shipboard visual line transect survey data from multiple sources—see Roberts et al. (2016) and the above web site for details. For each species, we used the most recent version of the models (see Table 4) available at the time we undertook our analysis; for minke whale there was a separate model for summer (April–October) and winter (November–March).

The models predict density of the studied species with one-month temporal resolution and 5 x 5 km spatial resolution. Two types of models were produced by MGEL, depending on the species (and, for minke whale, the season): for fin whale and minke whale (winter) the models used climatological covariates whose value changed by month but were averaged across years, while for fin whale, NARW, and minke whale (summer) the models used contemporaneous dynamic covariates whose value changed by month and year. In the latter case, we generated density surfaces using the most recent 5 years in the model (see Table 4). The spatial coverage of the MGEL models varied between species (Figure 3A): models for fin and minke whales in summer extended past the study area borders, but for NARW did not extend north of Cape Cod, while those for sei whales and minke whales in winter did not extend much further southwest than the southernmost wind farm (VOWA). In these cases, our power analyses excluded areas for which there were no density surfaces—but these were a small portion of the overall areas analyzed.

The density models include two sources of uncertainty: 1) uncertainty in the predicted average density surface given covariate values, represented by variances and covariances on model parameter values, and 2) variability in the number of animals present in a given grid cell on any day, represented by a statistical distribution on predicted per-cell numbers (the MGEL models used a Tweedie distribution). To incorporate the first source of uncertainty, we used parametric bootstrap resampling (sampling model parameters from a multivariate normal distribution) to generate 500 realizations of the density surfaces for each species and month. For models that use contemporaneous dynamic covariates, we sampled 100 realizations from each year, making 500 in total. To incorporate the second source of uncertainty, for each realization and 5 x 5 grid cell, we generated a random value from the Tweedie distribution with mean equal to the parametric bootstrap value and scale equal to the value from the MGEL model divided by 30. The reason to divide by 30 was to account for the fact that we want an average density per grid cell per month while the MGEL models use day as the sampling unit.

The above procedure occasionally generated unrealistically high-density values, and we hence implemented an outlier removal procedure. We removed densities greater than an outlier threshold defined as $Q3 + 1.5 * IQR$ where *Q3* is the third quartile and *IQR* the interquartile range of the densities.

The removed predictions were then replaced by new predictions generated by bootstrapping and generating values from Tweedie distribution, so the number of realization maps was always equal to 500. For all four species, no more than 20 realizations per species were above the defined threshold.

An example generated density map is shown in Figure 3B, and further examples for each species, wind farm and month are given in Appendix B.

Table 4. Density surface distribution models used for each of the studied species

Species	Model Version	Response Variable Distribution and its Parameters	Prediction Years	Reference for the Most Recent Model Description
Fin whale	v12	Tweedie (p = 1.14, scale = 7.20)	-	Roberts et al. (2022)
NARW	v12	Tweedie (p = 1.23, scale = 13.45)	2016–2020	Roberts and Yack (2022b)
Minke whale	v10	Summer: Tweedie (p = 1.13, scale = 6.97)	2015–2019	Roberts and Yack (2022a)
Minke whale	v10	Winter: Tweedie (p = 1.06, scale = 5.96)	-	Roberts and Yack (2022a)
Sei whale	v10	Tweedie (p = 1.22, scale = 14.36)	2016–2020	Roberts and Yack (2022c)

Notes: "-" indicates the model did not contain year-referenced covariates

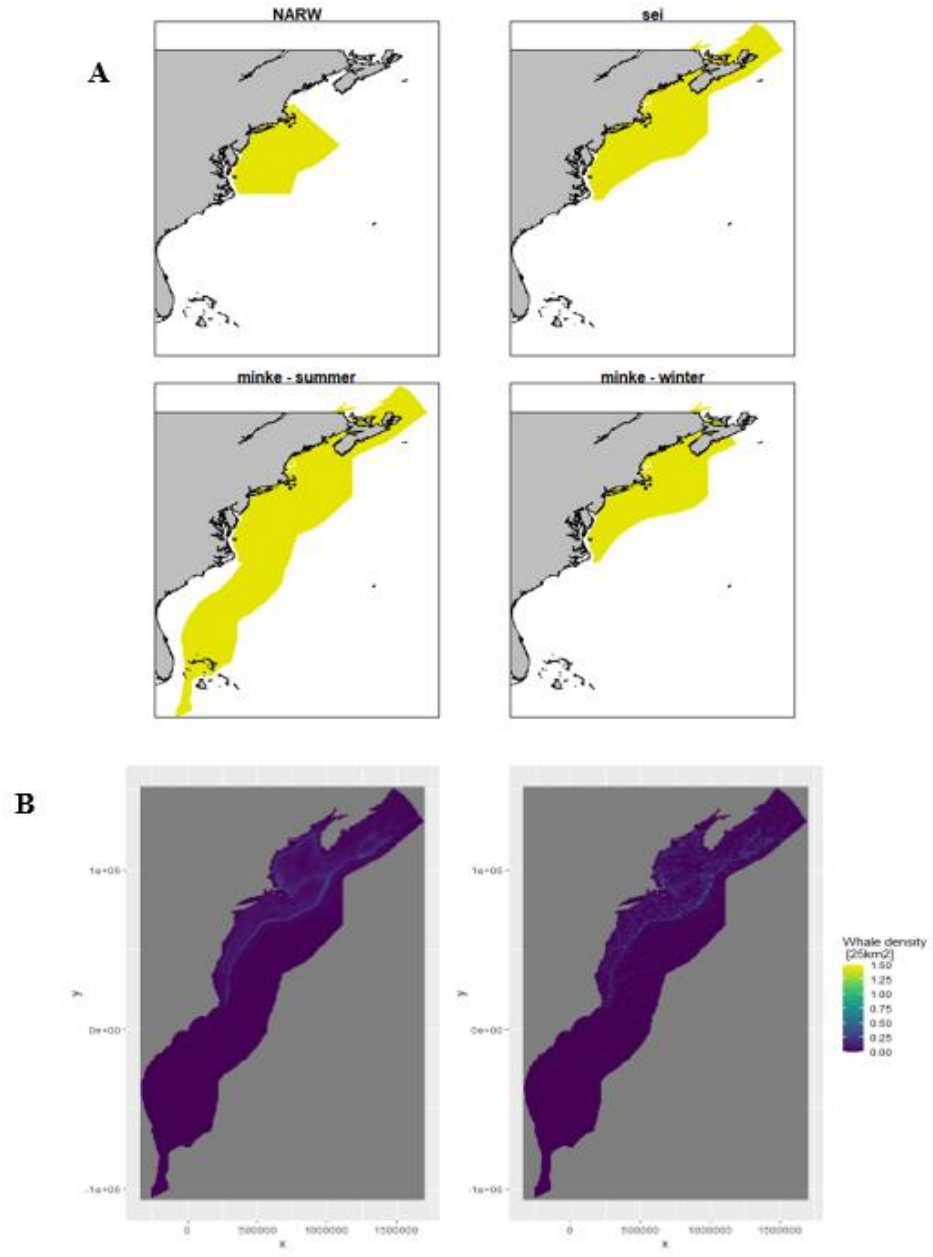


Figure 3. Spatial extent of the density surface models
A: model extent for NARW, sei whale and minke whale. The red square show location of the southernmost wind farm (VOWA) for reference. B: Density of fin whales in January an example of one out of 500 realization maps generated by parametric bootstrap (left) and then in addition accounting for daily variability (right).

2.3 Defining Source Location

This sub-step is relevant only for hypotheses involving wind farm construction (H2–H5, Table 2). As the detailed location and order of construction of each turbine was not available for each wind farm, we assumed that at each month of the construction, piling took place within one section of the wind farm at a time. To define sound source location, the wind farm was, therefore, divided into as many approximately similarly sized polygons as months of piling (Table 1, see for example Figure 4). The temporal ordering of the sections was chosen at random.

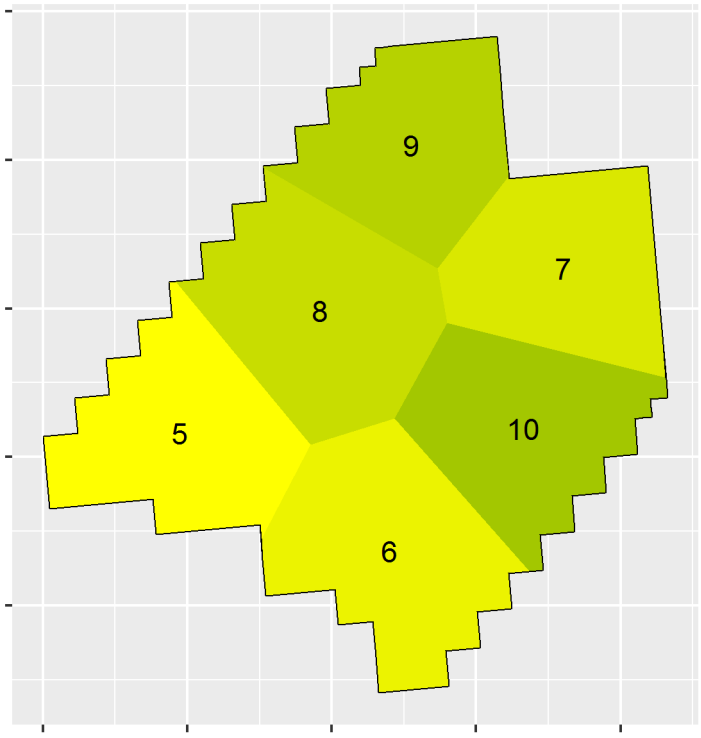


Figure 4. Division of the VYWA footprint based on location of piling
The piling takes place over 6 months and the numbers in each polygon indicate the piling month (e.g., 5 is May, 6 is June, etc.).

2.4 Calculating Number of Responding Simulated Whales

The number of responding simulated whales for the hypotheses related to construction (H2–H6, Table 2) was calculated based on dose-response functions and uncertainty around them established for NARW during an expert elicitation conducted in December 2022 (Booth et al. 2023). The functions elicited related to the received level of sound from pile driving at which a foraging NARW would switch from a foraging to non-foraging state for at least the rest of the pile driving on that day. We chose two functions, one below the elicited mean dose-response function equal to 1st quartile (i.e., less sensitive) and one above the elicited mean equal to 3rd quartile (i.e., more sensitive). We refer to these as DR1 and DR2 respectively. The same two functions were used for the entire calculations. The dose-response functions established in expert elicitation process were based on daily responses. Due to lack of information on cumulative response, we assumed that dose-response by month was the same as that based on daily responses. This would be the case, for example, if whales return to the baseline distribution at the end of each day after displacement.

We assumed impulsive source level of 200 dB re 1 µPa, including 10 dB broadband sound attenuation (following Pyć et al. (2018)) and transmission loss (TL) of

$$TL = 15\log_{10}(R \times 1{,}000) + \alpha R$$

where R is distance (km) from the source and α attenuation coefficient (dB km^{-1}). We chose α = 1.2, which was the mean value for areas of ~40 m depth as presented in Heaney et al. (2020) (see Table 3-1 of

that document). We also assumed that at the distance where probability of response <= 1%, this probability was 0.

These dose-response functions resulted in a very sharp decline in probability of response within the first 2 km from the source and reaching 1% at 18 km for DR1 (green line, Figure 5) and a more gradual decline with probability of response of 1% at 30 km for DR2 (orange line, Figure 5).

To calculate the number of responding simulated whales for each 5 x 5 km density grid cell, the number of animals in that grid cell was multiplied by the probability of response from the dose-response function. Distances were taken from the center of each density grid cell to the closest edge of the sound source polygon. For grid cells with centers inside the sound source polygon, all simulated whales were assumed to respond (see Discussion for effect of this assumption). An example showing probability of response under DR1 and DR2 for five months of piling at VYWA is shown in Figure 6

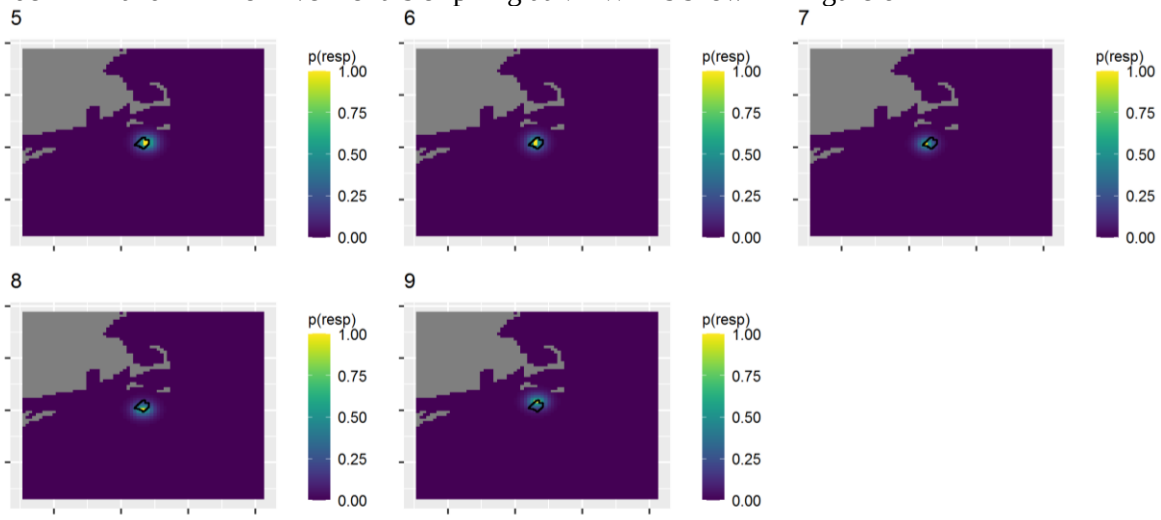


Figure 6.

For hypotheses related to operation (H6 and H7), we assumed that all simulated whales within the footprint of the wind farm respond. In practice that meant altering the density of 5 x 5 km density grid cells whose center fell within the wind farm footprint, and not altering density for cells with center outside the footprint. An example for VYWA is shown in Figure 7. One exception was for regional analyses under H7, which specifies that activities associated with wind farm operation causes displacement of simulated whales. Because some wind farms are directly adjacent to one another, or close by, we hypothesized that the response may occur over a cluster of wind farms rather than each individual one. We therefore created a set of four clusters (Figure 8) and assumed all animals within these clusters responded.

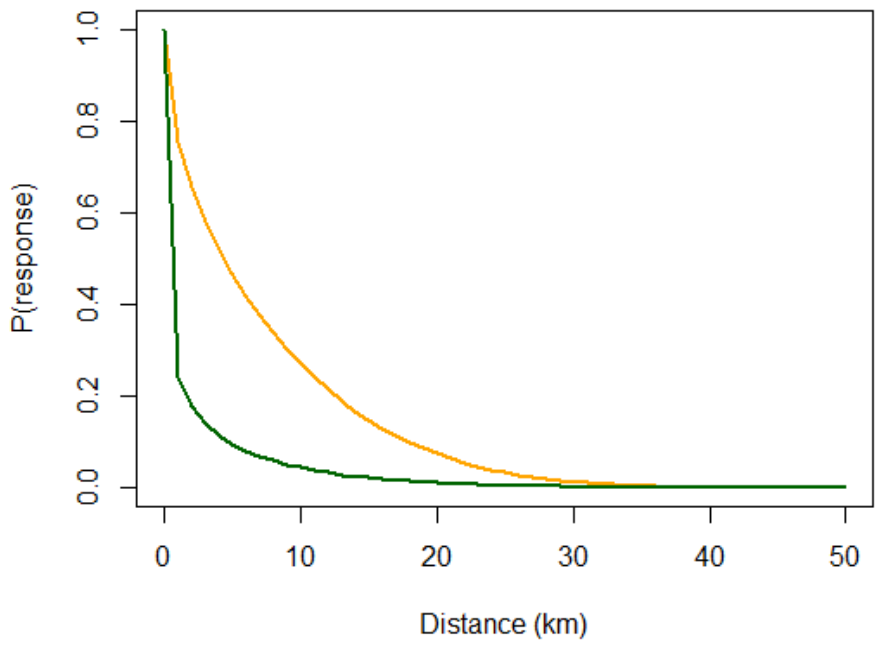
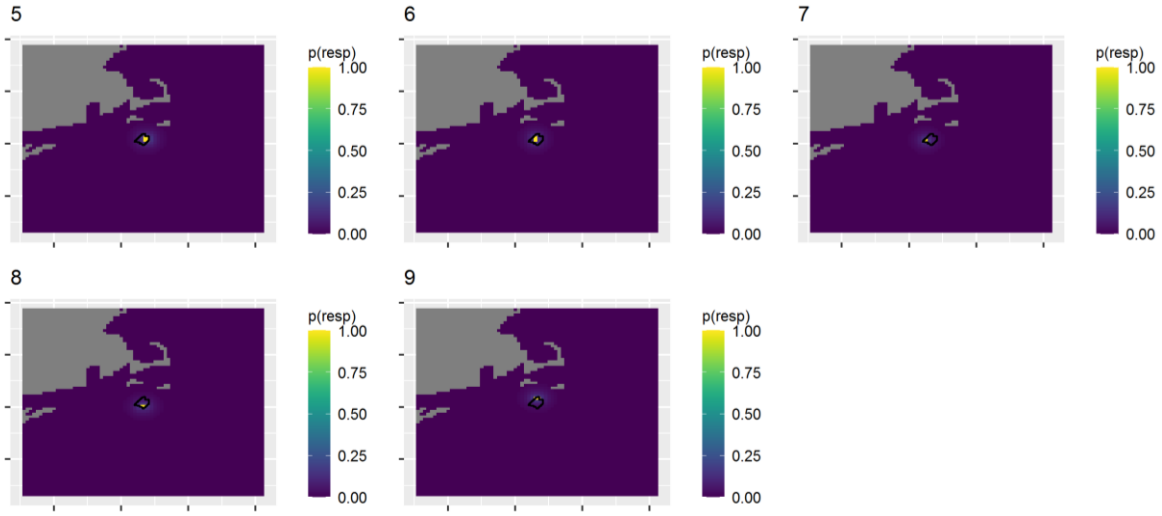


Figure 5. Probability of response with distance from the source for the two chosen dose-response functions assuming 200 dB source level
Green line is DR1, orange line is DR2.



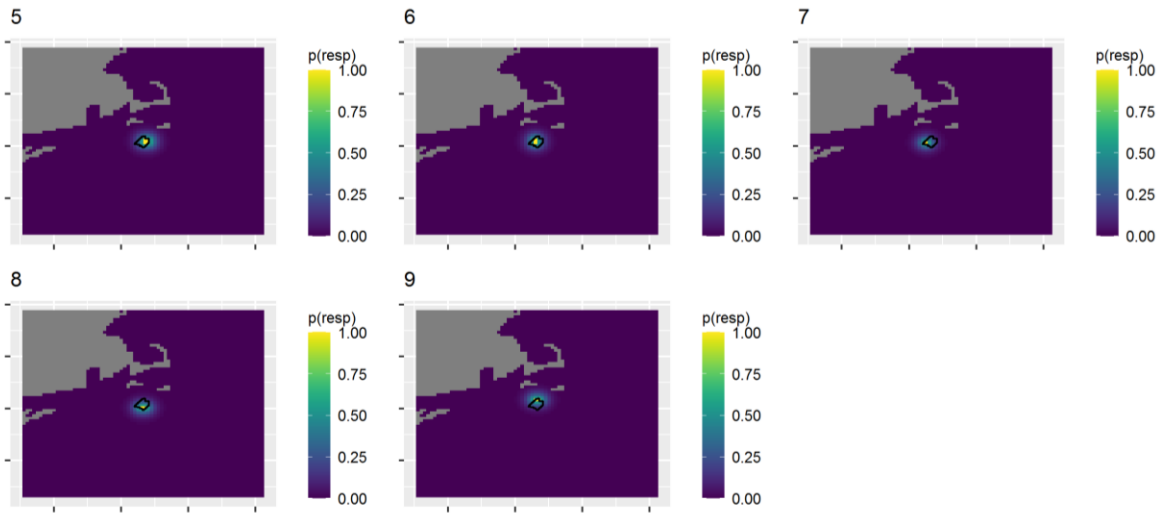


Figure 6. Probability of response with distance from the windfarm
First five panels show the probability of response (p(resp)) for each of the piling months denoted by the title of each panel for based on DR1 for VYWA. The lower five panels were based on DR2. The difference between months was the location of the sound source.

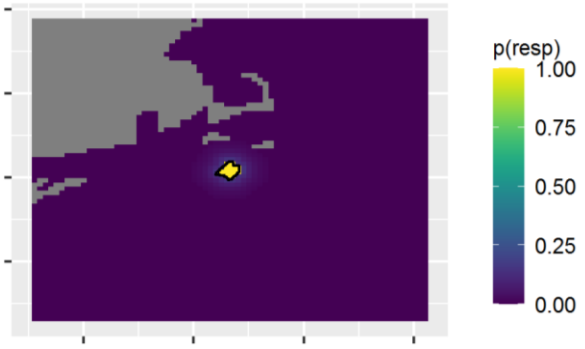


Figure 7. The probability of response (p(resp)) for one month of operation of VYWA
Probability of response was assumed to be 1 inside the wind farm footprint and 0 outside.

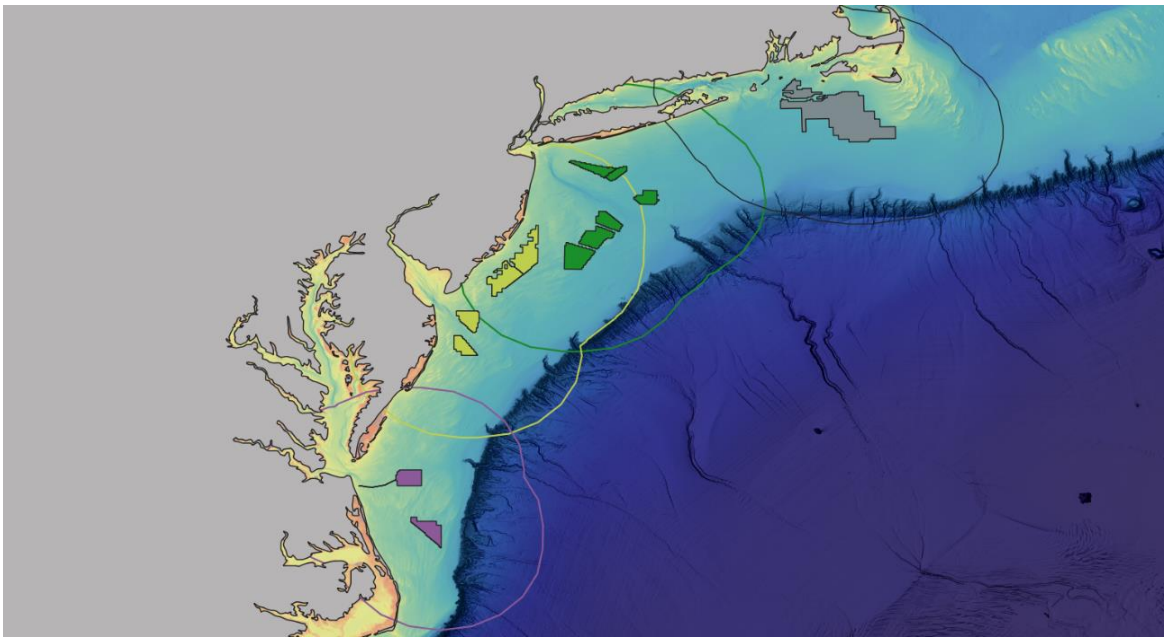


Figure 8. Wind farms considered in the regional analysis were grouped into four clusters (shown in matching colors) based on the close proximity of the individual wind farms
Under hypothesis H7, all animals within these clusters were assumed to be displaced. The lines depict 100 km buffers around each cluster within which displaced simulated whales were redistributed.

2.5 Generating Change in Simulated Whale Distribution and Density

Hypotheses H3–H5 relate to redistribution due to construction, while H6 and H7 to redistribution due to operation. The locations of redistributed whales differed between hypotheses (Table 2). For each of the hypotheses, change was generated separately for each construction/operation month. We therefore assumed that changes in behavior of whales in one month was independent from changes in the previous or following months. Animals were assumed to redistribute within 100 km of the wind farm (but only within habitat covered by the density models, e.g., not onto land).

Under H3 (displacement due to construction, equally in all directions), responding animals were redistributed uniformly within the 100 km radius, excluding the wind farm footprint. Under H4 (displacement due to construction, with preference for higher density locations), animals were redistributed in proportion to the underlying simulated density surface, so areas of high density received proportionally more responding animals than those with low density, under the assumption that they will be more attractive to displaced animals. This is reasonable because, in the density surface models we used, high density areas are interpreted as areas of high habitat suitability.

Hypothesis H5 is that higher density locations receive more displaced animals, but that this is tempered by increased anthropogenic activities outside of the wind farm, such as construction-related vessel traffic. As an approximation of this, we assumed that current vessel traffic is an indication of the locations of possible future increased traffic. We calculated mean monthly AIS per 5 x 5 km grid cell for each monitoring month based on mean of four days for each month (1, 10, 15, and 20th day of each month) from 2019 (source BOEM (2023)). We did not differentiate between traffic related to construction of windfarms and the remaining types. We then log transformed the number of vessels per grid cell and scaled to values between 0 and 1. We redistributed animals in proportion to the underlying simulated density surface multiplied by the scaled metric of vessel traffic.

Under H6 (operating wind farms attract simulated whales), we simply doubled the density in grid cells within the wind farm footprint. Under H7 (operating wind farms displace simulated whales), we allocated displaced simulated whales according to the underlying simulated density surface (as with H4), so any simulated whales that would have been located within the windfarm moved outside of it. For H7 under the regional analyses, we used the wind farm cluster footprints, rather than the footprints of individual wind farms, as detailed in the previous subsection.

Hypothesis H8 involves a global decline in simulated whale density, not related to wind farm construction or operation. To keep the size of the decline comparable to that in the vicinity of the wind farms under other hypotheses, we divided density by the ratio of the number of simulated whales remaining within 20 km of each wind farm under H3 (symmetric displacement) DR2 divided by the number under H1 (no response). By including H8 in the analysis, it may be possible to understand whether a 20 km monitoring size allows for a distinction between a larger-scale, global decline and decline in detected cues due to activities at the WEA.

2.6 Transforming Simulated Whale Densities into Number of Vocalizations (Cues)

The above sub-steps generated animal density per grid cell and month under baseline conditions and under each of the hypotheses of change. For each species, these were transformed into the number of vocalizations (cues) produced per grid cell and month by multiplying by an assumed cue production rate (Table 4), which was (for lack of better information) assumed to be constant over space and time. Note that information on cue production rate for fin, sei, and minke whales come from the Pacific region as such information are not available, to our knowledge, for the study site. Information on cue production rate for minke whales are based on ‘boing’ calls, which are not observed at this study site.

Hypothesis H2 involved a decrease in cue production rate for responding animals. Two alternatives were used: 1) 100% decrease in cue rate of responding simulated whales (referred to as “H2_100”) and 2) 50% decrease in cue rate of responding simulated whales (“H2_50”). The second alternative could be seen as a partial decrease in cue rate or a lessened ability to detect cues on the hydrophones during piling activities either due to masking or due to the fact that whales may stop calling when piling occurs (see Table 2A). (Detectability is dealt with in the next sub-step, but the two explanations were not analyzed separately because their effect on the data at the temporal level of a month would be similar.)

Table 5. Individual cue production rate, type of cues and effective detection range for the four whale species

Species	Cue Rate (cues/hour/individual), *c*	Type of Cues	Reference	Effective Detection Range (km), *v*
Minke whale	6.04	'boing' calls	Martin et al. (2013)	All wind farms: 8.6[3]
Fin whale	45.08	20 Hz pulse	Stimpert et al. (2015)	MAWA: 99.9[5] EOWA, VYWA: 93.8
Sei whale	10	30–87 Hz upsweeps and downsweeps combined	Baumgartner and Fratantoni (2008)[1] Calderan et al. (2014)[2]	All wind farms: 21.1[4]
NARW	6.2	Upcalls, variable tonal calls, gunshot sounds, and exhalations combined	Parks et al. (2011)	MAWA: 24.9 EOWA, VYWA: 9.2[5]

[1] No individual cue production rate is given in this study. We used the minimum, but no zero values, from Fig. 6.
[2] Back calculated based on call duration and inter-call intervals (Table 1 in the cited study).
[3] Percentiles of detection range for minke whale were not estimated in Estabrook et al. (2021) and we used values estimated by Salisbury et al. (2018) for all wind farms.
[4] Percentiles detection range for sei whale was not estimated in Salisbury et al. (2018) and we used values estimated by Estabrook et al. (2021) for all wind farms.
[5] Value used in regional analysis for all the wind farms.

2.7 Calculating Number of Detected Vocalizations on Sensors

In this sub-step the cue densities were used, in conjunction with an assumption about cue detectability, to determine the number of detections per sensor. We first describe how sensor locations were determined—i.e., the PAM designs—and then how number of detections was determined given a sensor location.

2.7.1 PAM Designs

We calculated the number of detected cues for three PAM designs: Van Parijs et al. and two alternative designs. The two alternative PAM designs have sensors placed closer to the footprints of each wind farm, where the effect of simulated whale distribution and behavior was expected to be largest (Figure 5). The first design, referred to as "10 x 10 km grid," is an equally spaced grid of PAM sensors spaced 10 km from each other (Figure 9A). The original design for small PAM grid by Van Parijs et al. used 20 km spacing. The second design, referred to as "T-design," consists of PAM sensors in three rows with each 'arm' of T situated across the expected gradient of simulated whale density (Figure 9B). Although such gradient difference between the four species both spatially and between months, there is a general trend for these species to have higher densities along the continental shelf. One 'arm' for the T-design is therefore, always pointing towards the continental shelf. The distance between sensors in the T-design is 2, 3, 4, 5, 6 and 8 km with sensors closer to each other in the center of the footprints of each wind farm. For both alternative designs, the numbers and locations of large PAM grid remains the same (Figure 9). The number of PAM sensors suggested by Van Parijs et al. in small PAM grid varied for the studied wind farms between 8 and 24 sensors (Table 6, Figure 1). The number of sensors in two alternative designs is more similar between sites: 19 for the T-design and 21–25 for the 10 x 10 km grid (Table 6).

Table 6. Number of PAM sensors used in the three PAM designs for each wind farm

Wind Farm	Small PAM Grid in Van Parijs et al.	10 x 10 km Grid	T-design
VYWA	24	21	19
EOWA	36	24	19
MYWA	8	21	19

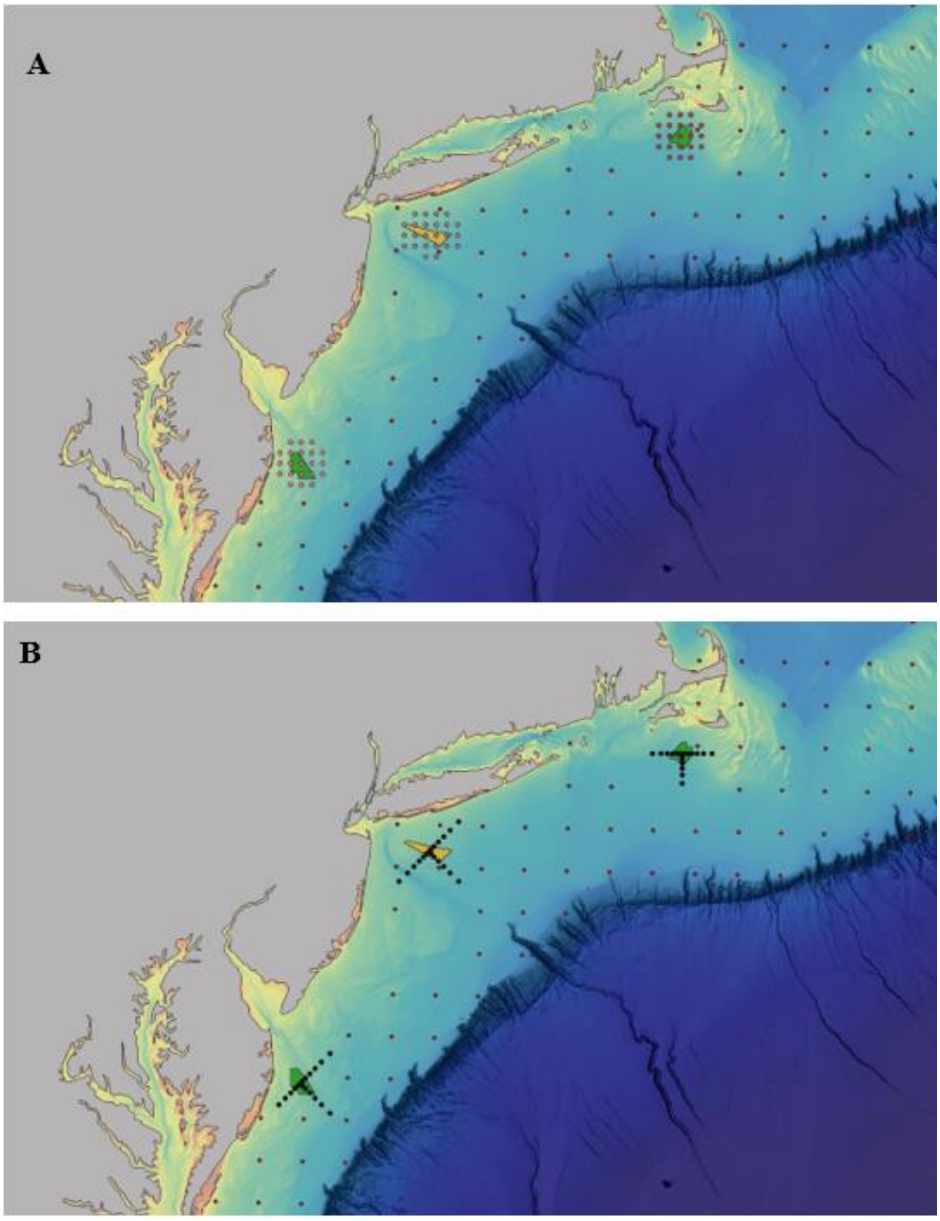


Figure 9. First alternative modification of small PAM grid suggested by Van Parijs et al. for the studied sites: A) 10 x 10 km grid, B) T-design
The large PAM grid is shown as red dots.

For the regional analysis, we tested only one alternative PAM design to Van Parijs et al.: T-design (Figure 10). The Van Parijs et al. design for the small PAM grid suggested a total of 98 sensors within the study site. For the regional alternative design, we redistributed the same number of sensors as the Van Parijs et al. design, but closer to the wind farms. The number of sensors within each cluster for this alternative design is proportional to baseline density of NARW at the 100 km buffer around the four clusters and the size of the clusters (**Error! Reference source not found.**). We placed 26, 40, and 19 sensors in the northern, two middle combined and southern cluster respectively. The distance between the sensors is between 5 and 15 km (Figure 10).

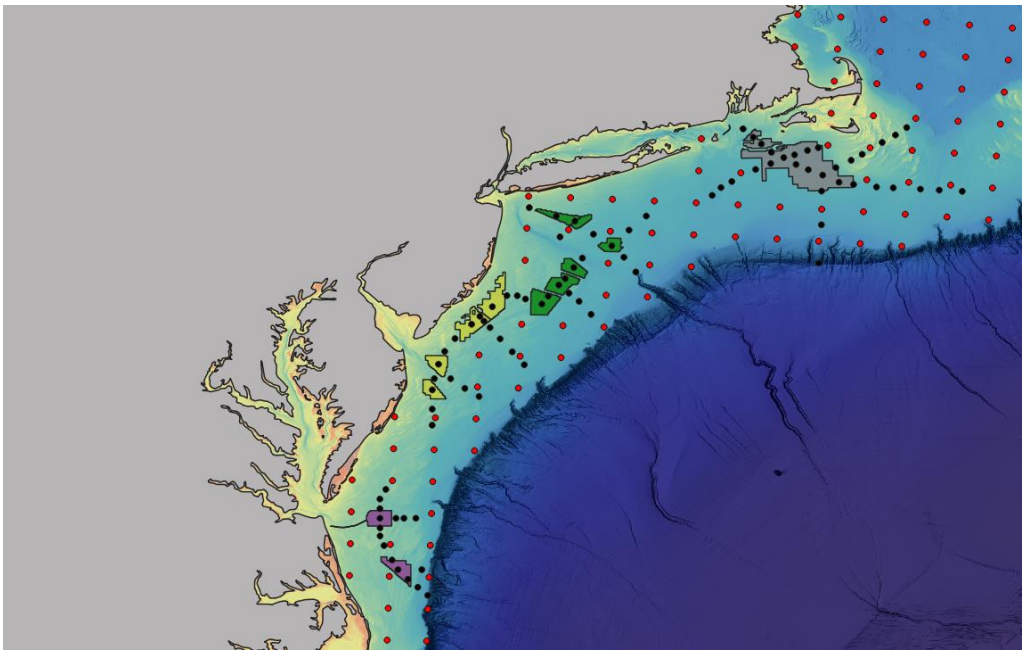


Figure 10. Alternative modification of small PAM grid suggested by Van Parijs et al. for the regional analysis (black dots)
The large PAM grid is shown as red dots.

2.7.2 Accounting for Detectability

To determine the number of cues detected on each sensor given the sensor location one must account for the detection range of the cues. In general, the probability of detecting vocalizations (cues) of a given species decreases with increasing distance between animal and sensor (although many other factors are also involved). We used a concept from the distance sampling literature (see, e.g., Buckland et al. (2001) (Section 3.1.3) and Marques et al. (2009) of the effective detection area (EDA)). The EDA is the circular area around a sensor within which as many vocalizations are missed as are detected outside it; hence the EDA can be thought of as a measure of the area monitored by a sensor. The radius of this circle is called the effective detection radius (EDR). EDA or EDR values have not been published for any of the four species for the entire regional study site, but empirical measurements of detectability have been taken at some sites and summary statistics published that allow them to be estimated. We used the reported detection ranges at which 5, 50, and 95 percent of vocalizations were estimated to have been detected by Salisbury et al. (2018) (see Table 5.2) for VOWA and MAWA and Estabrook et al. (2021) (mean of all stations presented in Table 29) for the remaining wind farms unless indicated differently in

Table 5. We fitted a three parameter detection function (a two-part mixture of half-normal functions, Miller and Thomas (2015), to these three summary statistics using a least-squares algorithm, and given the fitted detection function we estimated EDA as

$$\hat{v} = 2\pi \int_{r=0}^{w} r\hat{g}(r)dr$$

where *r* is range [km], $\hat{g}(r)$ is the estimated detection probability, and w is some suitably large truncation distance so that $\hat{g}(r) \approx 0$ at this range.

To calculate the number of detected vocalizations at a sensor given the EDA and the 5 x 5 km density grid, we summed the number of vocalizations produced in all grid cells within the EDA. For grid cells partially within the EDA, we pro-rated the number of vocalizations by the proportion of the grid cell within the EDA.

2.8 Assessing Power to Detect Response

We used analysis methods based on the concept of a phase-gradient (see Mackenzie et al. (2013), Methratta (2021) for an overview of the methods) design to detect effect of wind farm construction/operation. In essence, this process involves estimating the relationship between the number of detected cues and distance from wind farm under baseline conditions (which thereby takes account of any pre-existing gradients) and determining whether this relationship is different during construction or operation. In statistical terms this means determining whether there is a significant interaction in acoustic detection rate between distance from wind farm and phase (baseline, or construction and operation).

The specific analysis used was a generalized additive model (GAM), implemented using the *gam* function from the *mgcv* package (Wood 2017) in R. We used number of detected cues per each PAM station as the response, main effects for month, distance from wind farm, phase (baseline vs construction or operation) and (for the 5-year operational monitoring) year, and an interaction between distance and phase.[1] The terms involving distance were specified as smooths (thin plate regression splines), allowing for flexibility in the relationship between number of detected cues and distance. The response was modeled with an over dispersed Poisson error structure with log-link function; estimation was via restricted maximum likelihood (REML). We recorded whether the interaction between distance and phase was statistically significant at an α-level of 0.05 and also for H8 (global changes over time not related to wind farm construction), where the interaction was not significant, but the main effect of phase was significant.

Power for each species, hypothesis, dose-response curve, and monitoring area and grid was estimated as the proportion of the 500 simulations cases when interaction between phase and distance was significant. If this percentage >= 80% we refer to it as 'high power.' For the remaining cases we refer to as 'low power.'

[1] The interaction term was specified in *gam* as s(distance, by=phase), with phase being an ordered factor with baseline the first level. This meant that the main effect of distance corresponded to the effect of distance under baseline and the interaction term to the effect under construction/operation.

3 Results

The results are presented in Appendices C, D, and E.

3.1 Simulated Whale Density, Acoustic Encounter Rates and Number of Responding Animals

Minke whales had the highest average simulated density in the vicinity of the wind farms of all the study species, followed by fin whales, sei whales, and NARW (Table C-1). By contrast, fin whales had by far the highest average simulated acoustic encounter rate (2 orders of magnitude higher than the next highest, minke whales, Table C-1) because both their cue rate and effective detection range was the highest of the four species. The simulated acoustic encounter rate was lowest for NARW in two northern most WEAs: VYWA and EOWA; even though their density is relatively high in this area, this is the area where their EDR was lowest.

The proportion of simulated animals responding to construction under H2–H5 is driven by the dose-response function used, the size of monitoring area, and the pattern of animal density within this area. We used two dose-response functions—DR1 (more sensitive) and DR2 (less sensitive, Figure 5)—both of which had a low probability of response beyond 30 km. This gave a mean proportion of animals responding within the small monitoring area (effect size) (20 km around the wind farms) of 0.03 for DR1 and 0.09 for DR2, and within the large monitoring area (out to the continental shelf) of 0.002 for DR1 and 0.004 for DR2 (the results for large monitoring area are not presented). The proportion varied somewhat between wind farms (Table C-1). The simulated mean number of animals responding is a function of the proportion responding and the density of each species; in the majority of months and wind farms the number of responding simulated whales was <1 for sei whales, fin whales, and NARW, with a maximum of 8.9 for minke whales under the more sensitive dose-response function (DR2). Note that construction activities are not scheduled to occur in winter and early spring (December–March), which is the peak density of NARW.

For hypotheses involving wind farm operation (H6 and H7), only simulated animals within the wind farm footprint were assumed to respond, so proportion responding is a function of the size of the wind farm footprint, the size of monitoring area and the pattern of animal density within this area. The mean effect size (proportion of animals responding) for small monitoring area for H7 was 0.01 in local analysis. Assuming all 27 wind farms are operational at the same time (regional analysis), summing the effect across wind farms, the region-level mean number of animals responding was between 3.5 in November to 121.9 in May for minke whales, 4.3 in November and 24.9 in July for fin whales, 0.8 in August and 27 in February for NARW, and 0.3 in August to 9.9 in May for sei whales (Table C-2).

3.2 False-Positive Rates

The simulations under H1 include no effect of wind farm construction or operation on acoustic detections, and hence any case with a statistically significant interaction between distance from wind farm and phase (before or during construction or operation) is a false positive. Given a threshold for significance of 0.05, we expect a 5% false-positive rate, but since underlying density surfaces were sampled at random for each year (representing underlying variation in animal density over time), a higher false-positive rate is possible.

In the local analyses, the false-positive rate under H1 ranged from 5.2–23.8% (Tables D1, E1, and E2; see rows with H1 under "Hypothesis"). The rate was generally lower during construction (5.2–20%) than operation (6.2–23.8%). It was lower when all PAM sensors within the large monitoring area were used in the analysis, higher when only the small PAM grid was used within the large monitoring area, and highest when only the small PAM grid was used in the small monitoring area. There was no clear difference between the Van Parijs et al. design and the 10 x 10 km grid, but the false-positive rate was generally lower for the T-design (max 5.2–13.2%). There was no clear pattern among species and wind areas.

At the regional level (where only operation was assessed, not construction), false-positive rate was generally lower than for operation at the local level (compare rows with H1 under "Hypothesis" in Tables D2, E3 with those from Tables D1, E1, and E2), and ranged from 5.3–15.5%. There were no consistent differences between species or designs.

3.3 Statistical Power Under Van Parijs et al. Design

3.3.1 Local Analysis

There was low statistical power (< 80%) for all hypotheses related to construction and operation (H2–H7) regardless of the studied species, PAM grid use (small only or all), the chosen dose-response function, and the wind farm (Table D1).

Comparison of power between hypotheses showed that for hypotheses related to construction (H2–H5), power was lower for the hypothesis involving change in acoustic behavior (H2) vs those involving displacement (H3–H5). Of those involving displacement, estimated power was similar within species for all three hypotheses. There was high statistical power (> 80%) to detect a global decline (H8) for most species and wind farms regardless of the monitoring size and PAMs grid used. The only exception was for fin whales at the three WEAs for the large monitoring area and both PAM grids, where power was < 80%.

Comparison of power between species showed that highest power (but still below the 80% threshold) was estimated for minke whales and lowest for fin whales. There was no consistent pattern in power between WEAs within species.

Comparison of power between monitoring areas and PAM grids used showed that power was higher if only the small PAM grid was used in the smaller monitoring area, and lowest if both PAM grids were used in the large monitoring area. This seems like a paradoxical result—that using more sensors and monitoring a bigger area produces lower power—and we return to this in the Discussion.

Comparing the two dose-response functions, power was 1–5% higher for DR2 (the more sensitive dose-response function) than for DR1.

3.3.2 Regional Analysis

Only hypotheses related to operation (H6 and H7) were tested at the regional level. Power was high (> 80%) to detect change under both hypotheses for minke whales, but low (< 80%) for the other three species (with sei whales and NARW having intermediate power, and fin whales being very low) (Table D2). Increasing the number of years of operational monitoring for these species from 1 to 5 years increased power, but not enough to make power high (Table D2).

As with the local analyses, power to detect change was higher if only the small PAM grid was used.

3.4 Statistical Power Under Alternative PAM Designs Local Analysis

Power was substantially higher for the 10 x 10 km grid (Table E1) for all hypotheses related to construction and operation (H2–H7) relative to that from the Van Parijs et al. design, although power only exceeded the 80% threshold for minke whales for H4 and H5 at the WEA where this species is most abundant, VYWA. Comparison of power between monitoring areas and PAM grid used showed that power was higher by 2–7% only if the modified small PAM grid (either 10 x 10 km or T-design) was used in comparison to use of all grids (either 10 x 10 km or T-design together with large grid by Van Parijs et al.).

The T-design resulted in high power to detect change for all hypotheses related to operation (H6–H7) and construction (H2–H5) for all WEAs for minke whales apart from H2–H5. The T-design resulted in high power for all hypothesis related to construction which assumed redistribution of whales (H3–H5) for sei whale and NARW at the WEA where these two species are most abundant: VYWA.

As was found for the Van Parijs et al. design, power was higher for DR2 than DR1 by 1–5%.

3.4.1 Regional Analysis

Power in regional analyses was just calculated for the T-design. This gave higher power than the Van Parijs et al. design, with high power for minke and sei whales under H6 after 1 year of operational monitoring, and high power for all species except fin whales under H7 after 1 year of operational monitoring (Table E-3 in Appendix E).

4 Discussion

4.1 Discussion of Results

The statistical power to detect biologically plausible changes in acoustic encounter rate due to construction and operation was generally low (i.e., < 80%) under the Van Parijs et al. design, regardless of the studied species and WEA. The main determinant of these results is the small effect size, which is a result of three factors.

First, the distance over which the response was assumed to occur is small. For hypotheses involving construction, where a dose-response function for response was assumed, probability of response was almost 0 by 35 km for both DRs and below 0.1 by 18 km for the less sensitive DR. If a more sensitive dose-response function were assumed, where the probability of response remained higher at greater distances, power to detect change would also be greater. For hypotheses related to operation of the wind farms, the effect size was even smaller than for hypotheses related to construction as assumed probability of response = 0 outside the wind farm footprints. This explains why power was higher for hypotheses related to construction than operation for the individual WEA (local analysis).

Second, under the Van Parijs et al. design, relatively few sensors are placed within or close to the wind farm footprint, in the area where probability of response is high. The design assumed a minimum 20 km spacing of the PAM sensors, which covers almost the entire dose-response function. The alternative designs both had considerably more sensors within and close to the wind farm footprints (and allocated sensors more evenly between WEAs) and the result was a substantial increase in power. Of the two, the T-design gave higher power because the smaller between-sensor spacing at the middle of the wind farms resulted in more sensors where the response was greatest.

Third, baseline density of all four species was low, in some WEAs equal or very close to zero for some months. This may have produced a zero-effect size in some simulation replicates. For the T-design, sensors were placed along gradients of highest baseline density, and this may have contributed somewhat to the higher power.

Comparing the species, two factors seemed to drive variation in power between species, regardless of hypothesis and PAM design. The major factor was variation in detection ranges, summarized in the EDR measure. Fin whales, in particular, had a large EDR and low power. As described earlier, the large detection distances meant that PAM sensors detected a mix of responding and non-responding whales, effectively "diluting" the effect size and resulting in low power. There are two potential solutions to this. The first is to increase (or impose) the received level (or signal to noise ratio) detection threshold for this species in the PAM detector, thereby decreasing the range over which detections are made. This would result in a decrease in the sample size of detections, but since detections per month were estimated to be over 100,000 in many months, this is unlikely to be problematic. A second solution is to use the fact that multiple sensors can detect the same vocalization for this species (given the sensor spacing) and attempt to localize vocalizations. This then would enable a more refined analysis of effect, which presumably would have higher power. Determining whether localization is or is not possible, and the optimal acoustical processing method, is outside the scope of this report, and we recommend consulting experts on this topic if this is determined to be a priority for the large EDR species.

A second factor driving variation in power between species appeared to be the sample size of detections. This size was affected by three factors: the generated density of each species, the species-specific cue production rate, and EDR. The most abundant species producing large number of cue and detected over large distances (large EDR) would have highest sample size of detection. Although highest sample size of detections was simulated for fin whales, the power to detect change was low for this species (as explained in the above paragraph) due to large EDR, high abundance of this species, and highest cue production rate out of all four species. Although the cue production rate for minke whale and NARW is comparable and so is EDR at EOWA and VYWA, the power for minke whale was higher as this is the more abundant species.

The construction schedule was designed to avoid peak occurrence of NARW; nevertheless, we obtained high power to detect displacement with the T-design at the wind farm where NARW are most abundant, VYWA. Power for other hypotheses and wind farms, although not above the 80% threshold, was not far off in some cases.

With the Van Parijs design, power was higher under the displacement hypotheses H3–H5 than the acoustic behavior changes hypothesis H2. This is because the displacement hypotheses not only decreased acoustic detections in the vicinity of the WEAs, but also increased it at further distances from the WEAs (up to 100 km) by redistributing the displaced animals. Note that H3 (symmetrical displacement) is unlikely to be realistic, as it places animals in locations where they are not recorded at all under baseline conditions. H5 was based on AIS data, which is likely not an accurate representation of how disturbance could increase around WEAs during construction as we did not distinguish between AIS of the vessels from regular shipping lanes and vessels related to the construction of the wind farms. In any case, the difference in power between H3, H4, and H5 was small, and no consistent pattern was detected.

In monitoring during operation, we found that power was somewhat higher with 5 years of operational monitoring than 1. However, both scenarios assumed 1 year of baseline (pre-construction) monitoring; it may be that, instead of 1 year baseline and 5 years operation monitoring, a more balanced design (e.g., 3 years baseline, 3 years operation) will have higher power.

One seeming paradox in the results is that power was somewhat higher when just the small PAM grid within the small monitoring area was considered and was lower when both the small and large PAM grids

were considered within the large monitoring area. This finding is explained by again considering effect size. The large monitoring area contains mostly grid cells that are far from the WEAs and hence where no effect takes place. Hence, the average effect size over the large monitoring area is smaller than that over the small monitoring area. Similarly, the large grid contains PAM sensors that are farther from the WEAs on average than the small grid, which is clustered around each WEA. Therefore, the average effect size when large and small sensors are combined is less than for the small grid alone.

A conclusion from the above might be that there is no need for the large monitoring grid. However, the results from the false-positive tests showed that false-positive rate is lower with a larger monitoring area. This is a good reason to monitor over a larger area. Another reason is that the design will then be more robust to a misspecification of the dose-response function should animals be displaced over larger distances than were used here. False-positive rates were found to be lower for the T-design than other designs.

Power was generally good to detect a global decline. Although this was only tested at the WEA level, power is likely to be higher still when WEAs are combined in a regional analysis. We did not examine the potential of the PAM designs to monitor long-term population changes—this could be done with a longer-term region-level simulation, where simulations are based on plausible scenarios for population change. However, based on this initial study with H8, it seems that this might be a powerful approach compared with the current standard of visual line transect surveys, where power is often low due to high estimated variance. However, visual line transects surveys produce estimates of absolute density and abundance (notwithstanding issues associated with trackline detection bias and availability bias), while the PAM network tested here is designed only to produce acoustic cue detection rate. Separate studies of detection ranges and cue production rates would be required to enable production of absolute density estimates from the PAM network (Marques et al. 2009, Marques et al. 2013)

4.2 Limitations of Analysis and Potential Future Studies

In undertaking the power analysis, we made several assumptions. Here, we review these assumptions and their potential effect on estimated power. We suggest future studies that might be undertaken to improve reliability of the estimated power values.

- The presented analysis is based on simulated data, where we estimated the number of potentially detected cues using animal density models derived from visual data and literature-reported cue production rates for each species. For all four studied species, these cue rates come from very different habitats than the studied site. Additionally, the cue production rate was assumed to be constant over time and space for each species. If cue production rates are higher than our assumed values, then power will be higher; on the other hand, variation in cue production over time and space will reduce power. Targeted studies of cue rates in the studied species and within the study area (for example studies of factors affecting the calling behavior of NARW) would be required to produce better-supported inputs.
- One method for checking the realism of the simulated acoustic data would be to compare baseline simulated acoustic encounter rates with those measured on PAM sensors previously deployed within the study area. At the time of this study, only data on acoustic presence-absence were available, so further acoustic processing may be required to obtain measured acoustic encounter rates.
- The simulations assumed a consistent effect of each hypothesized change in each grid cell, month, and WEA. For each simulation realization, after randomly sampling from the animal density surface, the other steps were deterministic. For example, it is infeasible that cue production rate is constant over space and time. This likely caused power to be over-estimated given that in the real world a consistent, deterministic effect cannot be expected.

- The assumption that all animals within the wind area footprint responded during operations is almost certainly an overestimate of response and therefore results in an overestimate of power. However, the assumption for operation that no animals respond outside the wind area footprint likely results in an underestimate of power. The dose-response functions used were assumed to apply to all species and may be too sensitive or not sensitive enough. A simple propagation model was used in converting from a received level-based dose-response function to a distance-based one, and a single assumption about sound source level was used. The direction of any resulting bias in estimated power is unknown.
- Assuming that the monthly variation in density between grid cells in the density surface had a scale parameter that was $1/30^{th}$ the daily visual survey-derived, estimates may over or underestimate the variability (and hence power). Truncating outliers in the generated density surface likely had a minimal effect on power.
- The use of EDA in lieu of a full detection function when determining acoustic detections is an approximation and had an unknown effect on power (although likely not large).
- We assumed that EDA was constant in space and time, or only varied in space between sets of WEAs. Variation in detection area, like other sources of variability, will decrease power unless it can be accounted for by collecting additional data alongside the acoustic encounter rates that will allow detectability to be estimated.
- We assume no change in response (sensitization or desensitization) with repeated exposure. Either of these would change the effect size and hence power.
- In the analysis, some models fitted to simulated data failed to fully converge. The effect of this is likely minor but could be investigated further. For some scenarios, especially the one which applied to small (20 km) monitoring size, number of PAM sensors used in the simulations was low (six in case on Van Parijs et al. design in VOWA and MAWA), which was the frequent reasons of models not converging. Increasing number of sensors, as suggested in the two alternative PAM designs, greatly reduced this problem.

In evaluating the power of hypotheses relating to construction we undertook only local analyses, analyzing each wind farm separately. Our experience from the regional analysis of operational scenarios was that combining wind farms into a regional analysis increased power. Therefore, we expect higher power to detect construction effects could be gained by combining monitoring at several wind farms simultaneously, and this would make a useful future study.

We analyzed data at the temporal scale of one month because this was the resolution available to us from the habitat-based density models. However, there is likely a strong signal of any construction effect in hourly or daily acoustic records (see, e.g., the diurnal pattern suggested under H2 in Table 2). Hence, power is likely greater at this temporal level to detect a construction effect. We suggest that further efforts be made to obtain acoustic data that could be used as the basis for a simulation study at the daily level.

In creating and testing alternative designs, we tried two alternatives, the 10 x 10 km grid and the T-design, but many others are possible. In particular, the T-design could be replaced by a cross (x or +), although the branch running towards the shoreline would likely need to be truncated and would likely enter very low whale density habitat. In the T-design, different inter-sensor spacings could be investigated, and the robustness of the results to misspecification of the dose-response function checked (i.e., checking power if a very different dose-response function is used than the one used to optimize the sensor spacing). If an updated dose-response function were available, the T-design could be adjusted to account for the new function. Additionally, the T-design allows for adjusting spacing between sensors without necessarily increasing the total number of sensors used in the design. Such flexibility is lessened in a regularly spaced grid—firstly because of the regular spacing, and secondly because the number of sensors required goes up with the square of the grid spacing (e.g., a 5 x 5 km grid requires four times as

many sensors as a 10 x 10 km grid to cover the same area). Overall, the optimal spacing of the sensors should be directly related to expected probability of response with distance from the wind farms.

The power analysis we undertook looked only at power to detect an effect. After an effect of wind farm construction or operation is detected then natural follow-up questions are: 1) what is the dose-response function, and 2) how many animals are affected? It is not certain that the design best suited to detecting an effect will also produce the most precise dose-response function or estimate of number of affected animals. For example, the most important part of a dose-response function for estimating number of affected animals is the part at larger ranges, because many more animals are exposed at these ranges than at smaller ranges (Tyack and Thomas 2019). Obtaining a precise dose-response function at larger ranges may require more sensors at these ranges than in the T-design (although there are also the sensors of the large grid) or alternative data collection, such as through tagging individuals. Optimizing the sensor array to estimate these quantities could be the subject of a future study.

The statistical model we fitted was a one-dimensional model in the sense that the only spatial metric was distance from the sound source. It is possible, and potentially more powerful given sufficient sensors, to estimate effects using a two-dimensional spatial model (i.e., modeling acoustic encounter rate in the horizontal plane). This is the basis for the MRSea package (Scott-Hayward et al. 2017), which is routinely used in Europe to undertake phase-gradient analyses. Even if not used as part of the power analysis, an MRSea-type analysis may be worth considering if the PAM monitoring grid is commissioned. It would, for example, give insight into where in space animals are redistributing to (assuming a redistribution scenario) rather than just making inferences that there is some redistribution with respect to distance from sound source.

We have not considered the power of the proposed grid for detecting long-term global trends in cetacean abundance. This issue would make a potentially useful future study. One factor that would need to be considered is that the sensor placement to detect wind farm effects may be non-random, and the resulting data may not be suitable for analysis using typical “design-based” methods but instead may be used as part of a spatial modeling exercise. Such models are already a standard part of the analysis of long-term wildlife monitoring schemes.

The Van Parijs. et al. design covers the area from the coast to the continental shelf (Figure 1). It is worth noting that all of the four chosen species frequently occur further off the shelf (e.g. Roberts et al. 2016), beyond the area covered by Van Parijs et al. and the alternative designs suggested in this report. Expending the design further offshore was not tested in this study.

4.3 Recommendations

- Based on the results of this power analysis, we recommend replacing the proposed 20 x 20 km small grid of sensors around WEAs, with an alternative array that concentrates sensors where a response is expected and distributes sensors relatively evenly across WEAs used as study sites.
- However, there is also a need for sensors at distances from the WEAs where no response is expected, and hence there is a role for the 40 x 40 km grid, at least out to the distances we tested (out to the continental shelf). Including both the local sensors and the 40 x 40 km grid appeared also to reduce the false-positive detection rate.
- Of the designs we tested, the T-design appears better than a dense grid of sensors in terms of the amount of statistical power generated for a fixed investment of sensors.
- To maximize the sample size of acoustic sensors, we recommend pooling resources across stakeholders who are deploying sensors.
- For species with large acoustic detection distances (like fin whales), consideration should be given to localizing calls and undertaking effects analysis using the localizations.

- The power analysis we have undertaken could be improved, and we have made some suggestions for future studies in the previous section. One suggestion is to undertake analysis of existing PAM data to provide a cross-check that our simulated acoustic encounter rates are realistic.
- One method to improve power is to accept a higher false-positive detection rate. We used a nominal false-positive rate of 5% and a target power of 80%, but these values are merely conventions and consideration could be given to using other values.

5 References

Aschettino, J. M., D. T. Engelhaupt, A. G. Engelhaupt, A. DiMatteo, T. Pusser, M. F. Richlen, and J. T. Bell. 2020. Satellite telemetry reveals spatial overlap between vessel high-traffic areas and humpback whales (Megaptera novaeangliae) near the mouth of the Chesapeake Bay. Frontiers in Marine Science:121.

Bailey, H., B. Senior, D. Simmons, J. Rusin, G. Picken, and P. M. Thompson. 2010. Assessing underwater noise levels during pile-driving at an offshore windfarm and its potential effects on marine mammals. Marine Pollution Bulletin **60**:888-897.

Barco, S. G., W. A. McLellan, J. M. Allen, R. A. Asmutis-Silvia, R. Mallon-Day, E. M. Meagher, D. A. Pabst, J. Robbins, R. E. Seton, and W. M. Swingle. 2002. Population identity of humpback whales (Megaptera novaeangliae) in the waters of the US mid-Atlantic states. J. Cetacean Res. Manage. **4**:135-141.

Baumgartner, M. F., and D. M. Fratantoni. 2008. Diel periodicity in both sei whale vocalization rates and the vertical migration of their copepod prey observed from ocean gliders. Limnology and Oceanography **53**:2197-2209.

Benhemma-Le Gall, A., I. M. Graham, N. D. Merchant, and P. M. Thompson. 2021. Broad-scale responses of harbor porpoises to pile-driving and vessel activities during offshore windfarm construction. Frontiers in Marine Science **8**:664724.

BOEM. 2023. Bureau of Ocean Energy Management (BOEM) and National Oceanic and Atmospheric Administration (NOAA). MarineCadastre.gov. .

BOEM, and NOAA. 2022. DRAFT BOEM and NOAA Fisheries North Atlantic Right Whale and Offshore Wind Strategy.

Booth, C., A. K. D. O' Hara, and L. Thomas. 2023. North Atlantic Right Whale Expert Elicitation. Assessing population effects of offshore wind development on North Atlantic Right Whales (Eubalaena glacials). BOEM Contract 140M0121C008 report, version 1.0.

Brandt, M. J., A.-C. Dragon, A. Diederichs, M. A. Bellmann, V. Wahl, W. Piper, J. Nabe-Nielsen, and G. Nehls. 2018. Disturbance of harbour porpoises during construction of the first seven offshore wind farms in Germany. Marine Ecology Progress Series **596**:213-232.

Brandt, M. J., A. Dragon, A. Diederichs, A. Schubert, V. Kosarev, G. Nehls, V. Wahl, A. Michalik, A. Braasch, C. Hinz, C. Katzer, D. Todeskino, M. Gauger, M. Laczny, and W. Piper. 2016. Effects of offshore pile driving on harbour porpoise abundance in the German Bight. Report prepared for Offshore Forum Windenergie.

Buckland, S. T., D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas. 2001. Introduction to Distance Sampling. Estimating abundance of biological populations. Oxford University Press, Oxford.

Calderan, S., B. Miller, K. Collins, P. Ensor, M. Double, R. Leaper, and J. Barlow. 2014. Low-frequency vocalizations of sei whales (Balaenoptera borealis) in the Southern Ocean. The Journal of the Acoustical Society of America **136**:EL418-EL423.

CETAP. 1982. A characterization of marine mammals and turtles in the mid- and North Atlantic areas of the U.S. outer continental shelf, final report. Washington, DC: Bureau of Land Management.

Cholewiak, D., C. W. Clark, D. Ponirakis, A. Frankel, L. T. Hatch, D. Risch, J. E. Stanistreet, M. Thompson, E. Vu, and S. M. Van Parijs. 2018. Communicating amidst the noise: modeling the aggregate influence of ambient and vessel noise on baleen whale communication space in a national marine sanctuary. Endangered Species Research **36**:59-75.

Chudzinska, M., J. Wood, M. Ryder, and L. Thomas. 2023. Power analysis for optimal design of passive acoustic monitoring network in the Virginia Offshore Wind area. Report number SMRUC - XXX to RWSC (Unpublished). .

Claisse, J. T., D. J. Pondella, M. Love, L. A. Zahn, C. M. Williams, J. P. Williams, and A. S. Bull. 2014. Oil platforms off California are among the most productive marine fish habitats globally. Proceedings of the National Academy of Sciences **111**:15462-15467.

Cohen, J. 1988. Statistical Power Analysis for the Behavioral Sciences. 2nd edn. Hillsdale, New Jersey: Lawrence Erlbaum. .

Dähne, M., A. Gilles, K. Lucke, V. Peschko, S. Adler, K. Krugel, J. Sundermeyer, and U. Siebert. 2013. Effects of pile-driving on harbour porpoises (*Phocoena phocoena*) at the first offshore wind farm in Germany. Environmental Research Letters **8**:025002.

Dähne, M., J. Tougaard, J. Carstensen, A. Rose, and J. Nabe-Nielsen. 2017. Bubble curtains attenuate noise from offshore wind farm construction and reduce temporary habitat loss for harbour porpoises. Marine Ecology Progress Series **580**:221-237.

Davis, G. E., M. F. Baumgartner, J. M. Bonnell, J. Bell, C. Berchok, J. Bort Thornton, S. Brault, G. Buchanan, R. A. Charif, and D. Cholewiak. 2017. Long-term passive acoustic recordings track the changing distribution of North Atlantic right whales (Eubalaena glacialis) from 2004 to 2014. Scientific Reports **7**:1-12.

Davis, G. E., M. F. Baumgartner, P. J. Corkeron, J. Bell, C. Berchok, J. M. Bonnell, J. Bort Thornton, S. Brault, G. A. Buchanan, and D. M. Cholewiak. 2020. Exploring movement patterns and changing distributions of baleen whales in the western North Atlantic using a decade of passive acoustic data. Global change biology **26**:4812-4840.

Dunlop, R. A., R. D. McCauley, and M. J. Noad. 2020. Ships and air guns reduce social interactions in humpback whales at greater ranges than other behavioral impacts. Marine Pollution Bulletin **154**:111072.

Dunlop, R. A., M. J. Noad, R. D. McCauley, E. Kniest, R. Slade, D. Paton, and D. H. Cato. 2018. A behavioural dose-response model for migrating humpback whales and seismic air gun noise. Marine Pollution Bulletin **133**:506-516.

Dunlop, R. A., M. J. Noad, R. D. McCauley, L. Scott-Hayward, E. Kniest, R. Slade, D. Paton, and D. H. Cato. 2017. Determining the behavioural dose–response relationship of marine mammals to air gun noise and source proximity. Journal of Experimental Biology **220**:2878-2886.

Durbach, I. N., C. M. Harris, C. Martin, T. A. Helble, E. E. Henderson, G. Ierley, L. Thomas, and S. W. Martin. 2021. Changes in the movement and calling behavior of minke whales (Balaenoptera acutorostrata) in response to navy training. Frontiers in Marine Science **8**:660122.

Ellison, W., B. Southall, C. Clark, and A. Frankel. 2012. A new context-based approach to assess marine mammal behavioral responses to anthropogenic sounds. Conservation Biology **26**:21-28.

Erbe, C., C. Reichmuth, K. Cunningham, K. Lucke, and R. Dooling. 2016. Communication masking in marine mammals: A review and research strategy. Marine Pollution Bulletin.

Estabrook, B. J., K. Hodge, D. P. Salisbury, A. Rahaman, D. Ponirakis, D. Harris, J. Zeh, A. N. Park, and A. Rice. 2021. Final Report for New York Bight Whale Monitoring Passive Acoustic Surveys October 2017- October 2020. Contract C009925. New York State Department of Environmental Conservation. East Setauket, NY.

Fernandez-Betelu, O., I. M. Graham, and P. M. Thompson. 2022. Reef effect of offshore structures on the occurrence and foraging activity of harbour porpoises. Frontiers in Marine Science **9**.

Fregosi, S., D. V. Harris, H. Matsumoto, D. K. Mellinger, S. W. Martin, B. Matsuyama, J. Barlow, and H. Klinck. 2022. Detection probability and density estimation of fin whales by a Seaglider. The Journal of the Acoustical Society of America **152**:2277-2291.

Gervaise, C., Y. Simard, F. Aulanier, and N. Roy. 2019. Optimal passive acoustic systems for real-time detection and localization of North Atlantic right whales in their feeding ground off Gaspé in the Gulf of St. Lawrence. Department of Fisheries and Oceans Ottawa, Canada, Can. Tech. Rep. Fish. Aquat. Sci. 3345: ix + 58 p.

Graham, I. M., D. Gillespie, K. C. Gkikopoulou, G. D. Hastie, and P. M. Thompson. 2023. Directional hydrophone clusters reveal evasive responses of small cetaceans to disturbance during construction at offshore windfarms. Biol Lett **19**:20220101.

Graham, I. M., N. D. Merchant, A. Farcas, T. R. C. Barton, B. Cheney, S. Bono, and P. M. Thompson. 2019. Harbour porpoise responses to pile-driving diminish over time. Royal Society Open Science **6**:190335.

Harris, C. M., L. Thomas, E. A. Falcone, J. Hildebrand, D. Houser, P. H. Kvadsheim, F. P. A. Lam, P. J. O. Miller, D. J. Moretti, A. J. Read, H. Slabbekoorn, B. L. Southall, P. L. Tyack, D. Wartzok, and V. M. Janik. 2018. Marine mammals and sonar: dose-response studies, the risk-disturbance hypothesis and the role of exposure context. Journal of Applied Ecology.

Hatch, L. T., C. W. Clark, S. M. Van Parijs, A. S. Frankel, and D. W. Ponirakis. 2012. Quantifying loss of acoustic communication space for right whales in and around a US National Marine Sanctuary. Conservation Biology **26**:983-994.

Heaney, K. D., M. A. Ainslie, M. B. Halvorsen, K. D. Seger, R. Muller, M. Nijhof, and T. Lippert. 2020. A Parametric Analysis and Sensitivity Study of the Acoustic Propagation for Renewable Energy Sources. Sterling (VA), U.S. Department of the Interior, Bureau of Ocean Energy Management. Prepared by CSA Ocean Sciences Inc. OCS Study BOEM 2020-011. 165 p.

Katona, S. K., and J. A. Beard. 1990. Population size, migrations and feeding aggregations of the humpback whale (Megaptera novaeangliae) in the western North Atlantic Ocean. Pages 295-306 Report of the International Whaling Commission (Special Issue 12).

Kenney, R. D., C. A. Mayo, and H. E. Winn. 2020. Migration and foraging strategies at varying spatial scales in western North Atlantic right whales: a review of hypotheses. J. Cetacean Res. Manage.:251-260.

Kowarski, K. A., B. J. Gaudet, A. J. Cole, E. E. Maxner, S. P. Turner, S. B. Martin, H. D. Johnson, and J. E. Moloney. 2020. Near real-time marine mammal monitoring from gliders: Practical challenges, system development, and management implications. J Acoust Soc Am **148**:1215.

Kraus, S., R. D. Kenney, and L. Thomas. 2019. A Framework for Studying the Effects of Offshore Wind Development on Marine Mammals and Turtles. Report prepared for the Massachusetts Clean Energy Center, Boston, MA 02110, and the Bureau of Ocean Energy Management.

Mackenzie, M., L. Scott-Hayward, C. S. Oedekoven, H. Skov, E. Humphreys, and E. Rexstad. 2013. Statistical Modelling of Seabird and Cetacean data: Guidance Document. University of St. Andrews contract for Marine Scotland; SB9 (CR/2012/05).

Malme, C., P. Miles, C. Clark, P. Tyack, and J. Bird. 1984. Investigations of the Potential Effects of Underwater Noise from Petroleum Industry Activities on Migrating Gray Whale Behavior—Phase II. U-S. Department of the Interior Minerals Management Service.

Marques, T. A., L. Thomas, S. W. Martin, D. K. Mellinger, J. A. Ward, D. J. Moretti, D. Harris, and P. L. Tyack. 2013. Estimating animal population density using passive acoustics. Biological Reviews **88**:287-309.

Marques, T. A., L. Thomas, J. Ward, N. DiMarzio, and P. L. Tyack. 2009. Estimating cetacean population density using fixed passive acoustic sensors: An example with Blainville's beaked whales. The Journal of the Acoustical Society of America **125**:1982-1994.

Martin, S. W., T. A. Marques, L. Thomas, R. P. Morrissey, S. Jarvis, N. DiMarzio, D. Moretti, and D. K. Mellinger. 2013. Estimating minke whale (*Balaenoptera acutorostrata*) boing sound density using passive acoustic sensors. Marine Mammal Science **29**:142-158.

Methratta, E. T. 2021. Distance-Based Sampling Methods for Assessing the Ecological Effects of Offshore Wind Farms: Synthesis and Application to Fisheries Resource Studies. Frontiers in Marine Science **8**.

Meyer-Gutbrod, E. L., C. H. Greene, and K. T. Davies. 2018. Marine species range shifts necessitate advanced policy planning: The case of the North Atlantic right whale. Oceanography **31**:19-23.

Miller, D. L., and L. Thomas. 2015. Mixture models for distance sampling detection functions. PLoS ONE **10**:e0118726.

Mitchell, B. 1975. Preliminary report on Nova Scotia fishery for sei whales (Balaenoptera borealis). Report of the International Whaling Commission, 25, 218–225.

Morano, J. L., D. P. Salisbury, A. N. Rice, K. L. Conklin, K. L. Falk, and C. W. Clark. 2012. Seasonal and geographical patterns of fin whale song in the western North Atlantic Ocean. The Journal of the Acoustical Society of America **132**:1207-1212.

Muirhead, C. A., A. M. Warde, I. S. Biedron, A. Nicole Mihnovets, C. W. Clark, and A. N. Rice. 2018. Seasonal acoustic occurrence of blue, fin, and North Atlantic right whales in the New York Bight. Aquatic Conservation: Marine and Freshwater Ecosystems **28**:744-753.

NROC. 2009. Northeast Regional Ocean Council. Northeast Ocean Data Portal, www.northeastoceandata.org. Date accessed: 10/01/2023 https://northeastoceandata.org/a9rtNd6M.

PACM. 2023. Passive Acoustic Cetacean Map. Woods Hole (MA): NOAA Northeast Fisheries Science Center v1.1.5 [accessed date]. https://apps-nefsc.fisheries.noaa.gov/pacm

Palmer, K. J., S. Tabbutt, D. Gillespie, J. Turner, P. King, D. Tollit, J. Thompson, and J. Wood. 2022. Evaluation of a coastal acoustic buoy for cetacean detections, bearing accuracy and exclusion zone monitoring. Methods in Ecology and Evolution.

Parks, S. E., A. Searby, A. Celerier, M. P. Johnson, D. P. Nowacek, and P. L. Tyack. 2011. Sound production behavior of individual North Atlantic right whales: implications for passive acoustic monitoring. Endanger Species Res **15**:63-76.

Pyć, C., D. Zeddies, S. L. Denes, and M. Weirathmueller. 2018. Appendix III-M: REVISED DRAFT - Supplemental Information for the Assessment of Potential Acoustic and Non-acoustic Impact Producing Factors on Marine Fauna during Construction of the Vineyard Wind Project. Document 001639, Version 3.1. Technical report by JASCO Applied Sciences (USA) Inc. for Vineyard Wind.

Quintana-Rizzo, E., S. Leiter, T. Cole, M. Hagbloom, A. Knowlton, P. Nagelkirk, O. Brien, C. Khan, A. Henry, and P. Duley. 2021. Residency, demographics, and movement patterns of North Atlantic right whales Eubalaena glacialis in an offshore wind energy development area in southern New England, USA. Endangered Species Research **45**:251-268.

R Core Team. 2022. R: A language to environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Ramp, C., J. Delarue, P. J. Palsbøll, R. Sears, and P. S. Hammond. 2015. Adapting to a warmer ocean—seasonal shift of baleen whale movements over three decades. PLoS ONE **10**:e0121374.

Risch, D., M. Castellote, C. W. Clark, G. E. Davis, P. J. Dugan, L. E. Hodge, A. Kumar, K. Lucke, D. K. Mellinger, and S. L. Nieukirk. 2014. Seasonal migrations of North Atlantic minke whales: novel insights from large-scale passive acoustic monitoring networks. Movement Ecology **2**:24.

Risch, D., C. W. Clark, P. J. Dugan, M. Popescu, U. Siebert, and S. M. Van Parijs. 2013. Minke whale acoustic behavior and multi-year seasonal and diel vocalization patterns in Massachusetts Bay, USA. Marine Ecology Progress Series **489**:279-295.

Roberts, J., and T. Yack. 2022a. Density Model for Common Minke Whale (Balaenoptera acutorostrata) for the U.S. East Coast, Version 10, 2022-06-20, and Supplementary Report. Marine Geospatial

Ecology Laboratory, Duke University, Durham, North Carolina. https://seamap.env.duke.edu/models/Duke/EC/EC_Common_minke_whale_history.html.

Roberts, J., and T. Yack. 2022b. Density Model for North Atlantic Right Whale (Eubalaena glacialis) for the U.S. East Coast, Version 12, 2022-02-14, and Supplementary Report. Marine Geospatial

Ecology Laboratory, Duke University, Durham, North Carolina. https://seamap.env.duke.edu/models/Duke/EC/EC_North_Atlantic_right_whale_history.html.

Roberts, J., and T. Yack. 2022c. Density Model for Sei Whale (Balaenoptera borealis) for the U.S. East Coast, Version 10, 2022-06-20, and Supplementary Report. Marine Geospatial Ecology Laboratory, Duke University, Durham, North Carolina. https://seamap.env.duke.edu/models/Duke/EC/EC_Sei_whale_history.html.

Roberts, J., T. Yack, and P. N. Halpin. 2022. Density Model for Fin whale (Balaenoptera physalus) for the U.S. East Coast, Version 12, and Supplementary Report. Marine Geospatial Ecology Laboratory, Duke University, Durham, North Carolina.Prepared for Naval Facilities Engineering Command, Atlantic by the Duke University Marine Geospatial Ecology Lab, Durham, NC. https://seamap.env.duke.edu/models.

Roberts, J., T. Yack, and P. N. Halpin. 2023. Marine mammal density models for the U.S. Navy Atlantic Fleet Training and Testing (AFTT) study area for the Phase IV Navy Marine Species Density Database (NMSDD). Document version 1.3. Report prepared for Naval Facilities Engineering Systems Command, Atlantic by the Duke University Marine Geospatial Ecology Lab, Durham, North Carolina.

Roberts, J. J., B. D. Best, L. Mannocci, E. Fujioka, P. N. Halpin, D. L. Palka, L. P. Garrison, K. D. Mullin, T. V. Cole, and C. B. Khan. 2016. Habitat-based cetacean density models for the US Atlantic and Gulf of Mexico. . Scientific Reports **6**:1-12.

Rolland, R. M., R. S. Schick, H. M. Pettis, A. R. Knowlton, P. K. Hamilton, J. S. Clark, and S. D. Kraus. 2016. Health of North Atlantic right whales *Eubalaena glacialis* over three decades: From individual health to demographic and population health trends. Marine Ecology Progress Series **542**:265-282.

Salisbury, D. P., B. J. Estabrook, H. Klinck, and A. Rice. 2018. Understanding marine mammal presence in the Virginia Offshore Wind Energy Area. Sterling (VA): US Department of the Interior, Bureau of Ocean Energy Management. OCS Study BOEM 2019-007. 103 p.

Scott-Hayward, L., C. S. Oedekoven, M. L. Mackenzie, C. Walker, and E. Rexstad. 2017. MRSea package (version 1.0-beta): Statistical Modelling of bird and cetacean distributions in offshore renewables development areas. Retrieved from http://creem2.st-and.ac.uk/software.aspx.

Sivle, L. D., P. H. Kvadsheim, C. Curé, S. Isojunno, P. J. Wensveen, F.-P. A. Lam, F. Visser, L. Kleivane, P. L. Tyack, and C. M. Harris. 2015. Severity of expert-identified behavioural responses of humpback whale, minke whale, and northern bottlenose whale to naval sonar. Aquatic Mammals **41**:469.

Sivle, L. D., P. J. Wensveen, P. Kvadsheim, F.-P. A. Lam, F. Visser, C. Curé, C. M. Harris, P. L. Tyack, and P. Miller. 2016. Naval sonar disrupts foraging behaviour in humpback whales. Marine Ecology Progress Series.

Steidl, R. J., and L. Thomas. 2001. Power analysis and experimental design. Design and analysis of ecological experiments:14-36.

Stepanuk, J. E., E. I. Heywood, J. F. Lopez, R. A. DiGiovanni Jr, and L. H. Thorne. 2021. Age-specific behavior and habitat use in humpback whales: implications for vessel strike. Marine Ecology Progress Series **663**:209-222.

Stimpert, A. K., S. L. DeRuiter, E. A. Falcone, J. Joseph, A. B. Douglas, D. J. Moretti, A. S. Friedlaender, J. Calambokidis, G. Gailey, and P. L. Tyack. 2015. Sound production and associated behavior of tagged fin whales (Balaenoptera physalus) in the Southern California Bight. Animal Biotelemetry **3**:23.

Stöber, U., and F. Thomsen. 2021. How could operational underwater sound from future offshore wind turbines impact marine life? The Journal of the Acoustical Society of America **149**:1791-1795.

Teilmann, J., and J. Carstensen. 2012. Negative long term effects on harbour porpoises from a large scale offshore wind farm in the Baltic — evidence of slow recovery. **045101**.

Tyack, P., and L. Thomas. 2019. Using dose–response functions to improve calculations of the impact of anthropogenic noise. Aquatic Conservation Marine and Freshwater Ecosystems. **29(S1)**:242-253.

Tyack, P. L. 2011. Using digital acoustic recording tags to detect marine mammals on Navy ranges and study their responses to Naval sonar.

Van Parijs, S. M., K. Baker, J. Carduner, J. Daly, G. E. Davis, C. Esch, S. Guan, A. Scholik-Schlomer, N. B. Sisson, and E. Staaterman. 2021. NOAA and BOEM Minimum Recommendations for Use of Passive Acoustic Listening Systems in Offshore Wind Energy Development Monitoring and Mitigation Programs. Frontiers in Marine Science **8**:760840.

Vu, E. T., D. Risch, C. W. Clark, S. Gaylord, L. T. Hatch, M. A. Thompson, D. N. Wiley, and S. M. Van Parijs. 2012. Humpback whale song occurs extensively on feeding grounds in the western North Atlantic Ocean. Aquatic Biology **14**:175-183.

Whitt, A. D., K. Dudzinski, and J. R. Laliberté. 2013. North Atlantic right whale distribution and seasonal occurrence in nearshore waters off New Jersey, USA, and implications for management. Endangered Species Research **20**:59-69.

Wood, J., B. Southall, and D. Tollit. 2012. PG&E offshore 3-D seismic survey project EIR – Marine Mammal Technical Report., SMRU Ltd Report.

Wood, S. N. 2017. Generalized Additive Models: An Introduction with R (2nd

edition). Chapman and Hall/CRC.

Zoidis, A. M., K. S. Lomac-MacNair, D. S. Ireland, M. E. Rickard, K. A. McKown, and M. D. Schlesinger. 2021. Distribution and density of six large whale species in the New York Bight from monthly aerial surveys 2017 to 2020. Continental Shelf Research **230**:104572.

Appendix A. Example Simulated Baseline Densities by Species, Wind Energy Area, and Month

Densities are shown as individuals/25 km^2 and are displayed on a 5 x 5 km grid. Numbers over each sub-panel indicates month of the year (1 = January, etc.).

A.1 Fin Whale

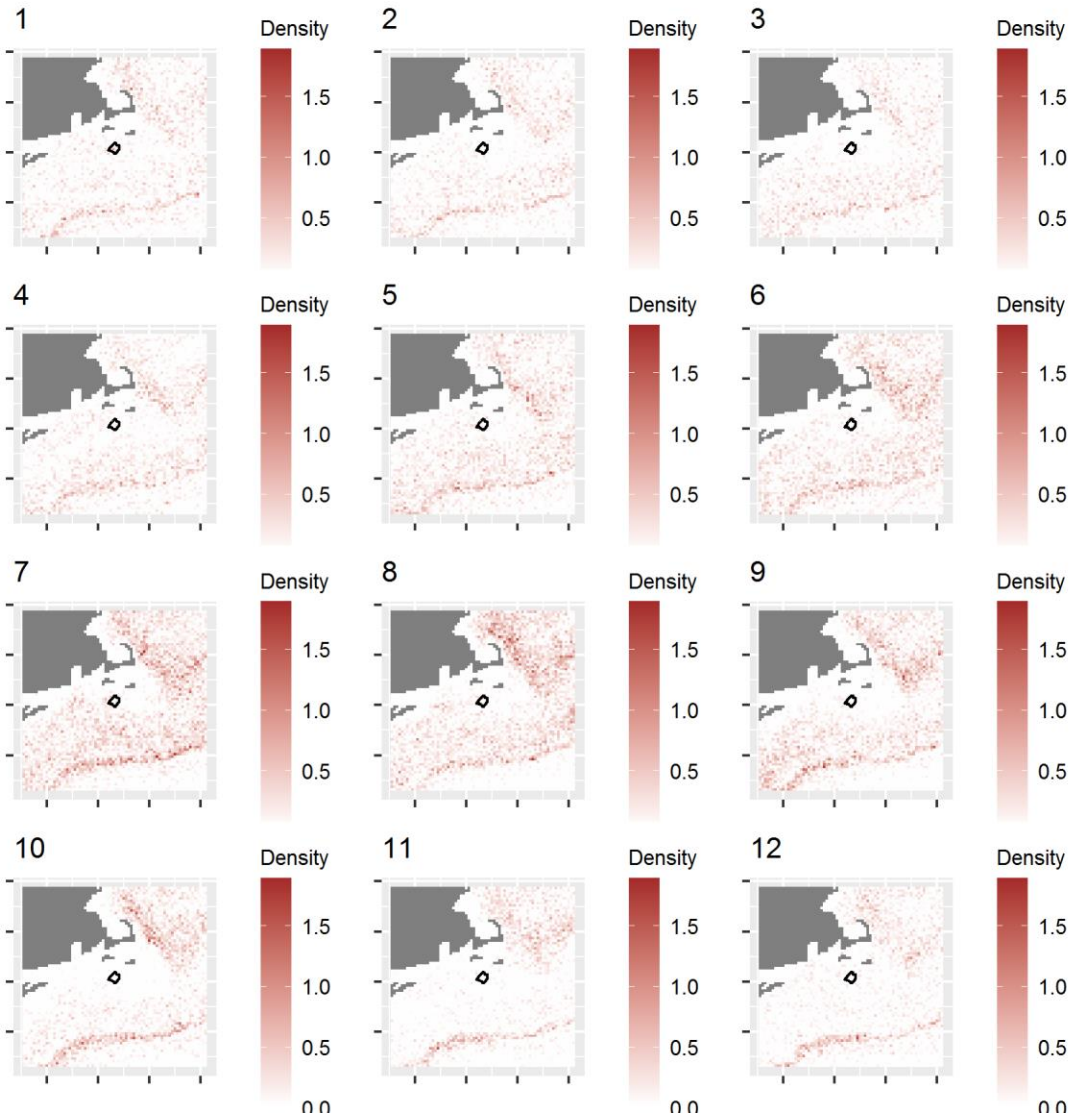


Figure A-1. Fin whale VYWA

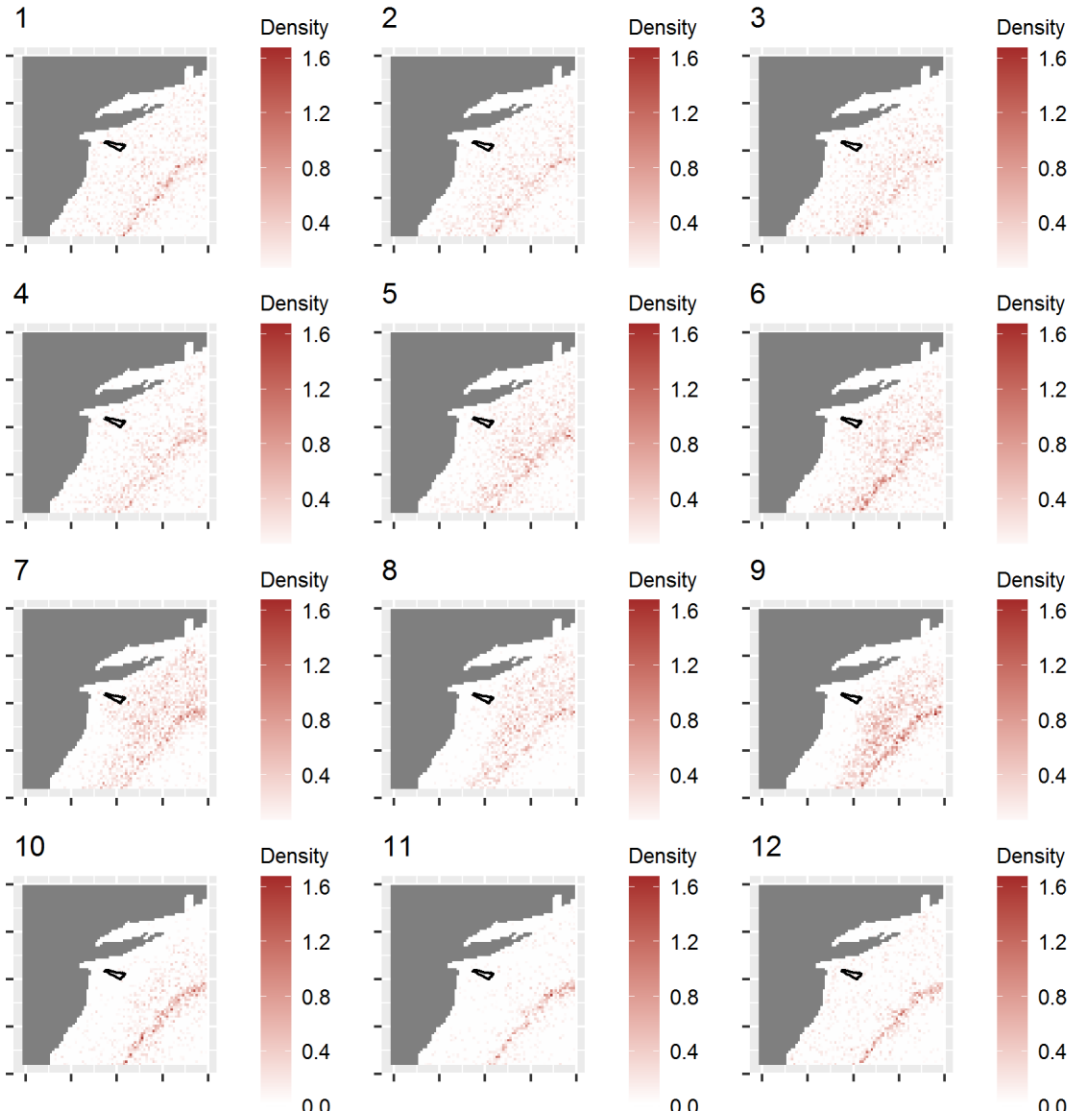


Figure A-2. Fin whale EOWA

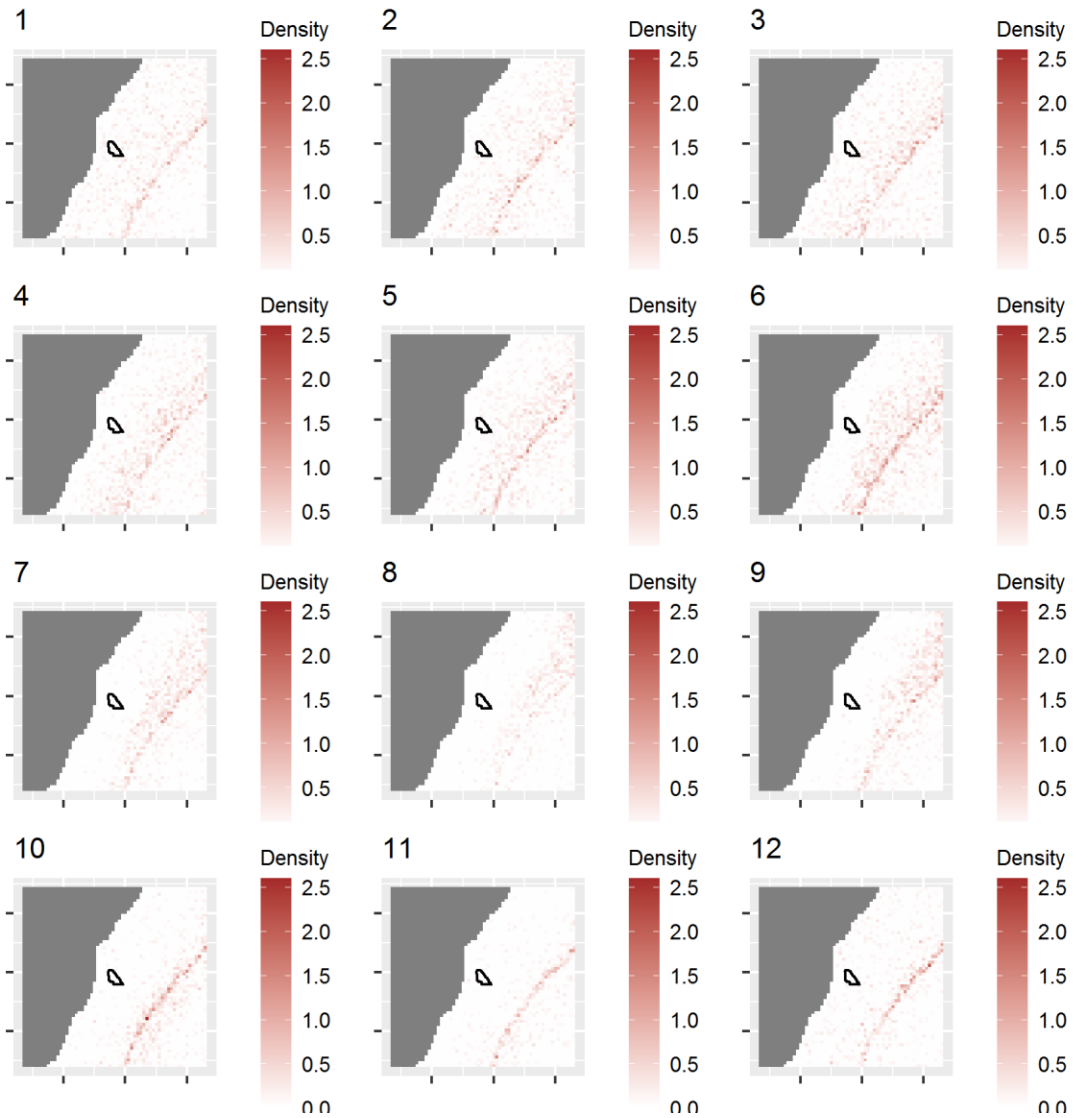


Figure A-3. Fin whale MAWA

A.2 Sei Whale

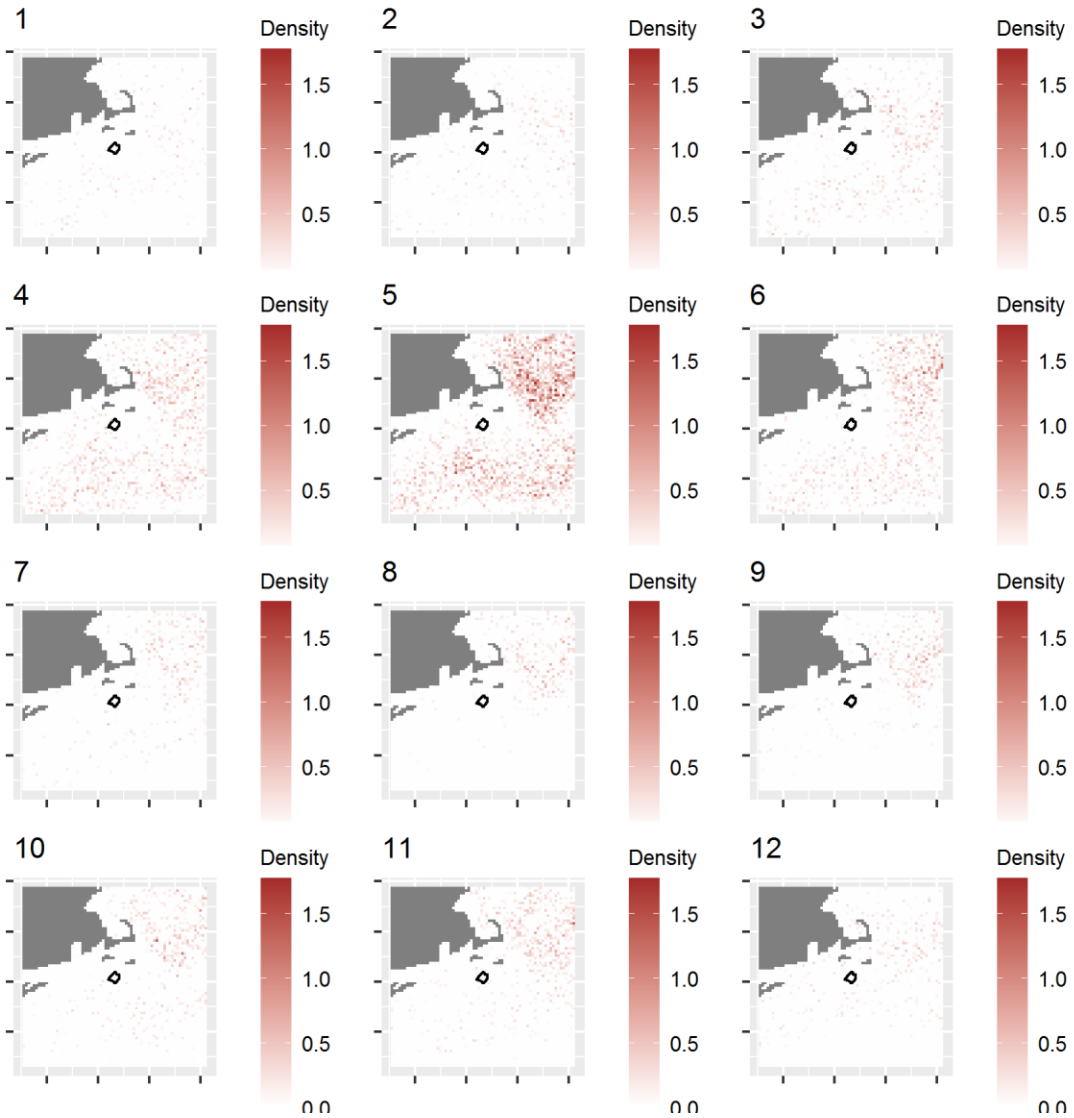


Figure A-4. Sei whale VYWA

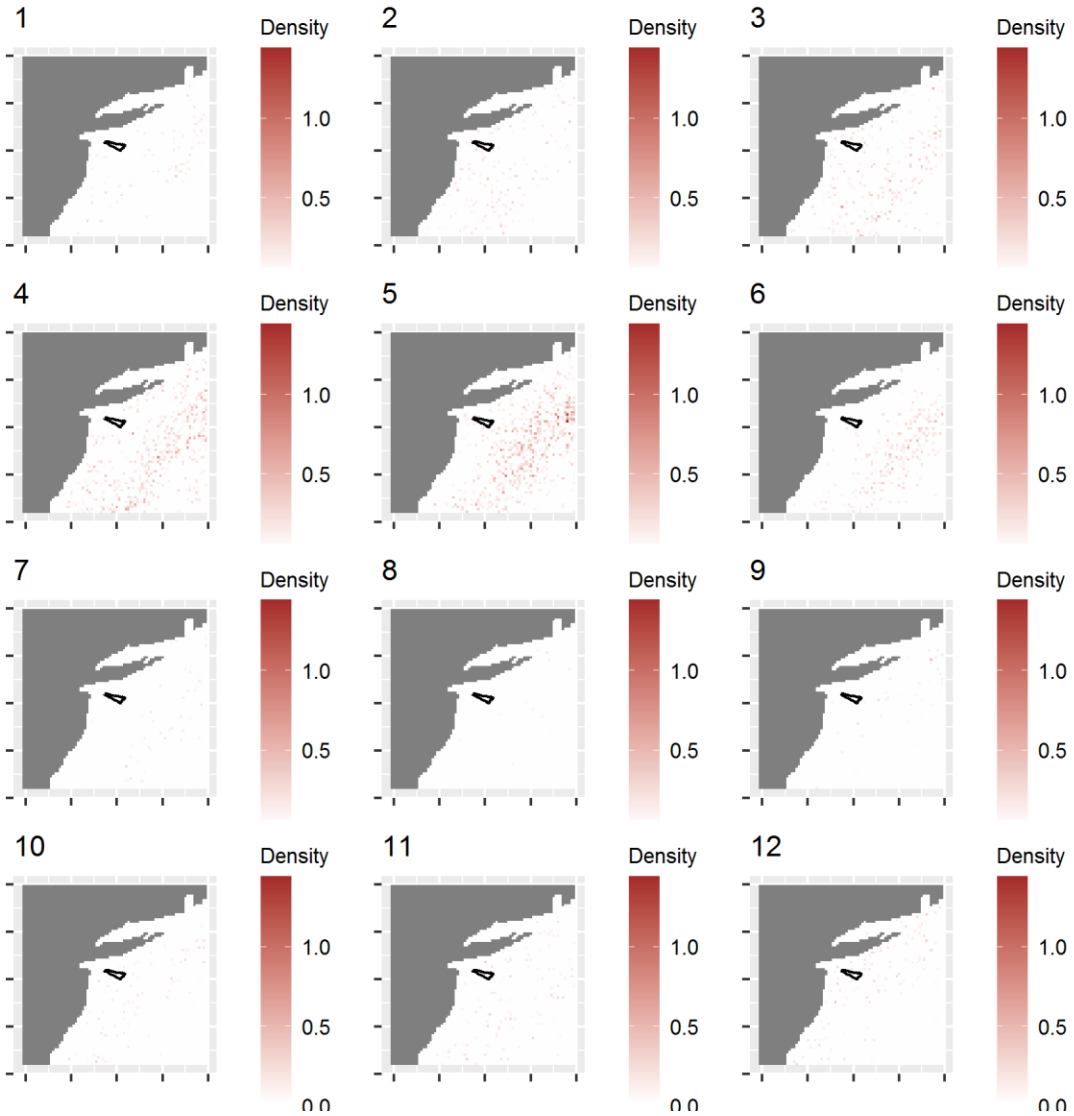


Figure A-5. Sei whale EOWA

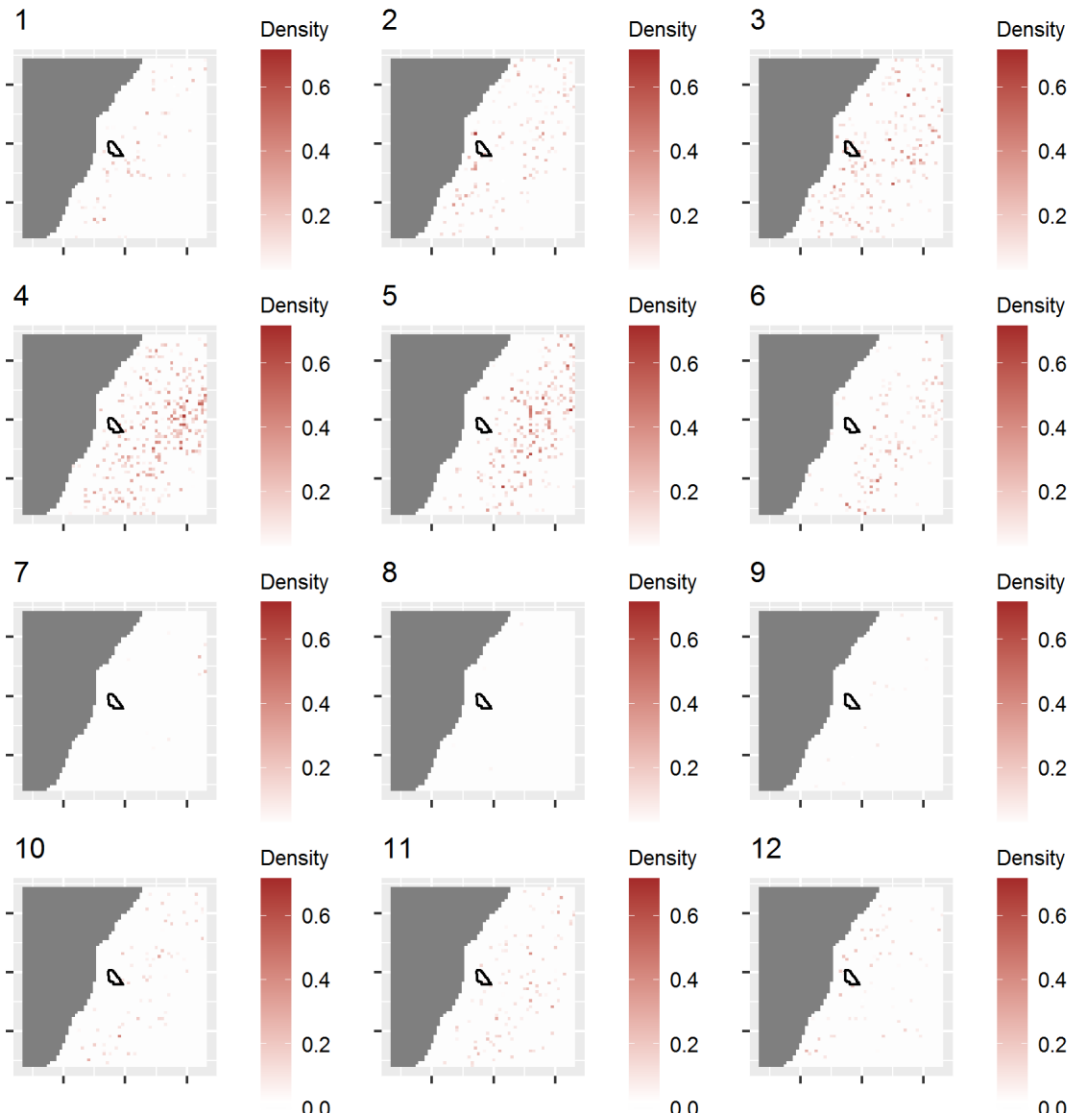


Figure A-6. Sei whale MAWA

A.3 Minke Whale

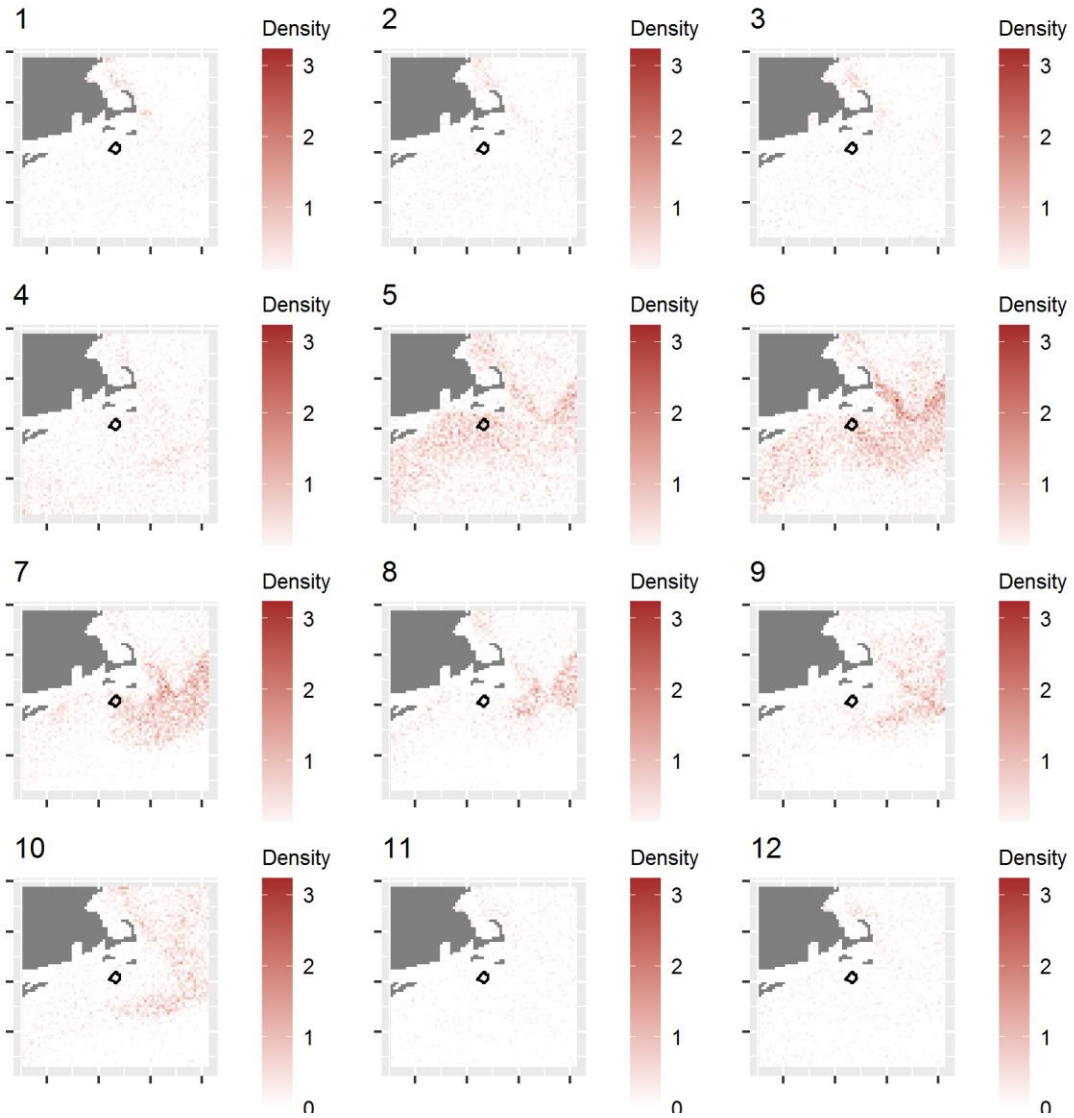


Figure A-7. Minke whale VYWA

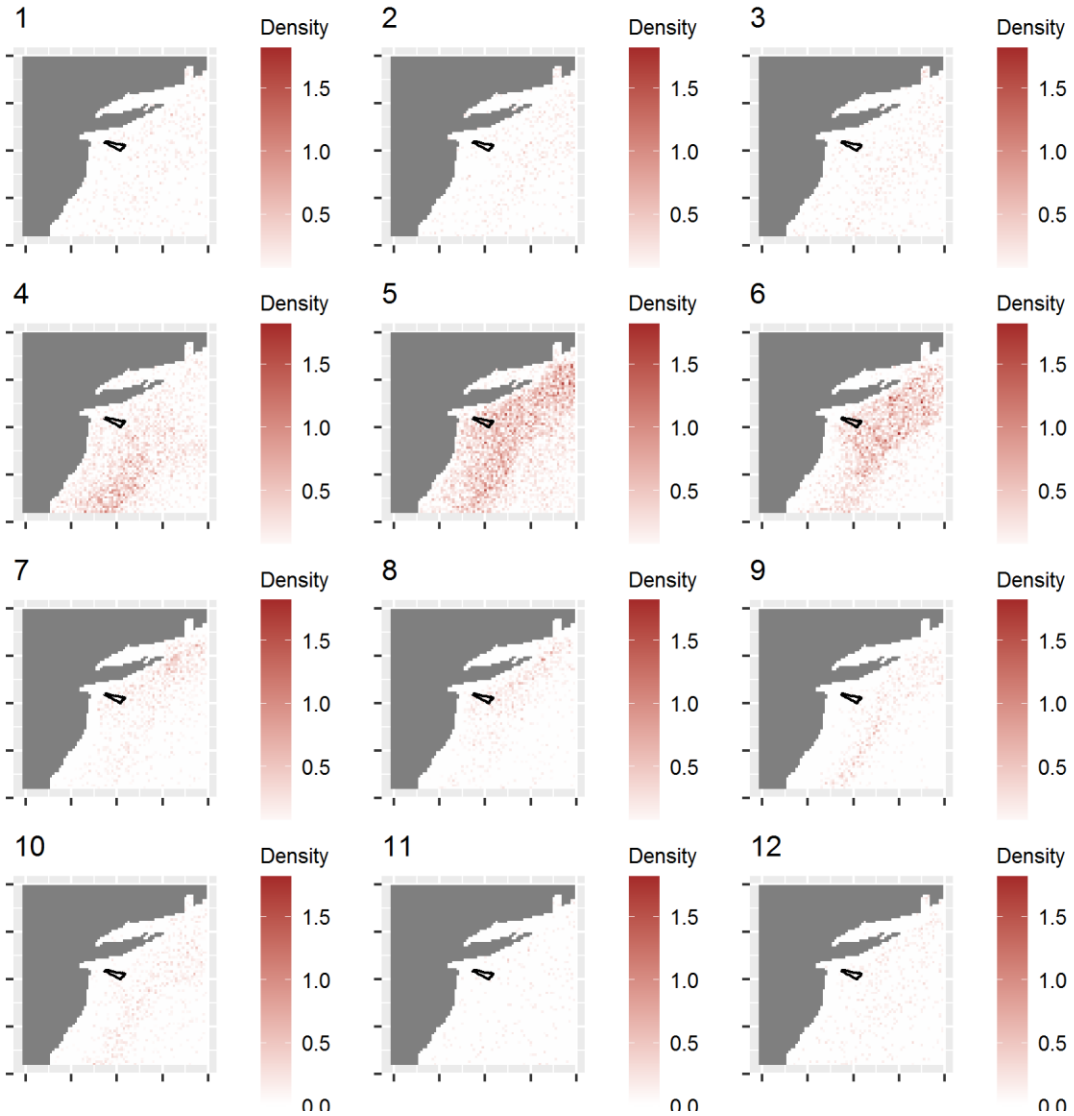


Figure A-8. Minke whale EOWA

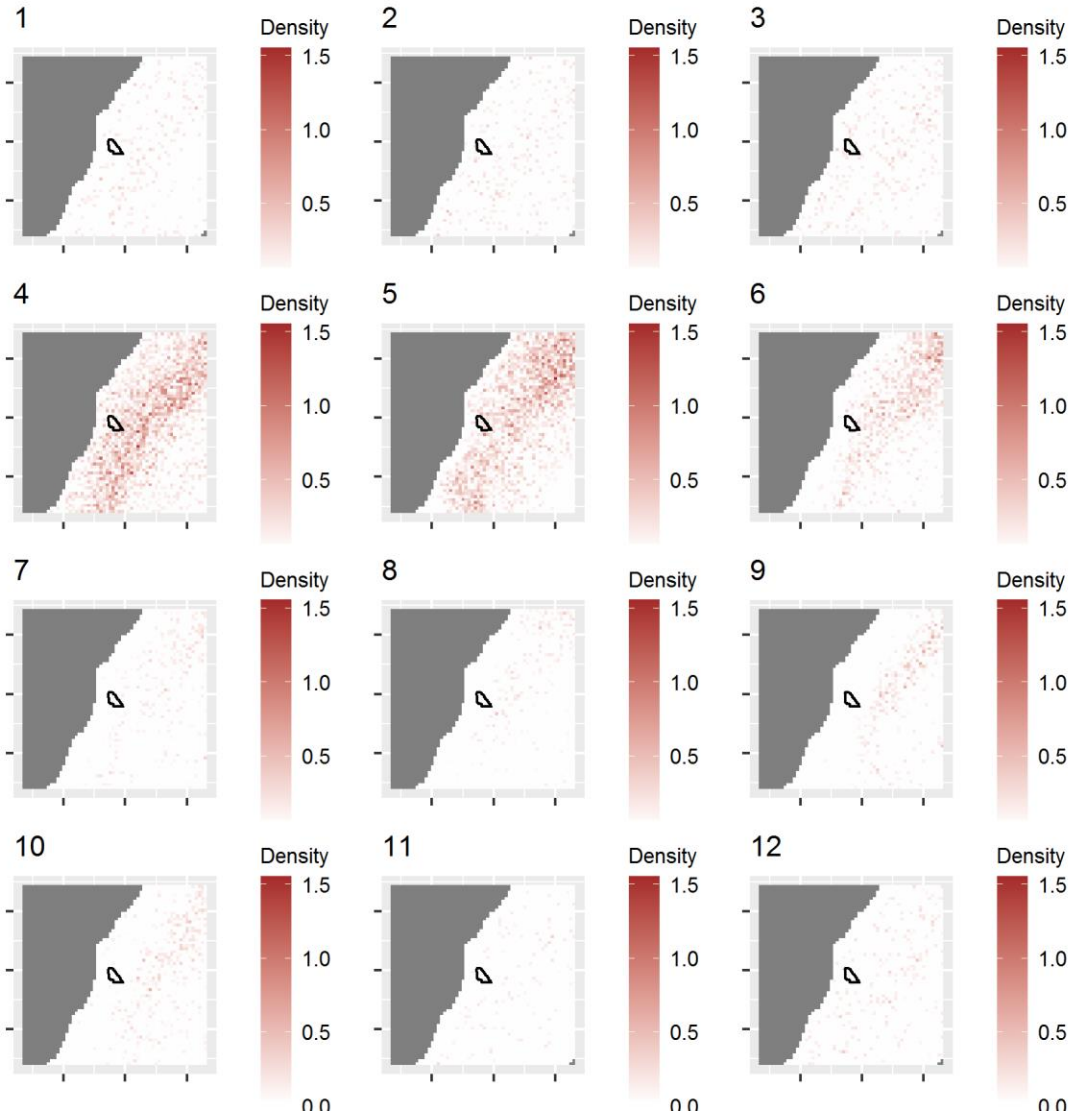


Figure A-9. Minke whale MAWA

A.4 North Atlantic Right Whale (NARW)

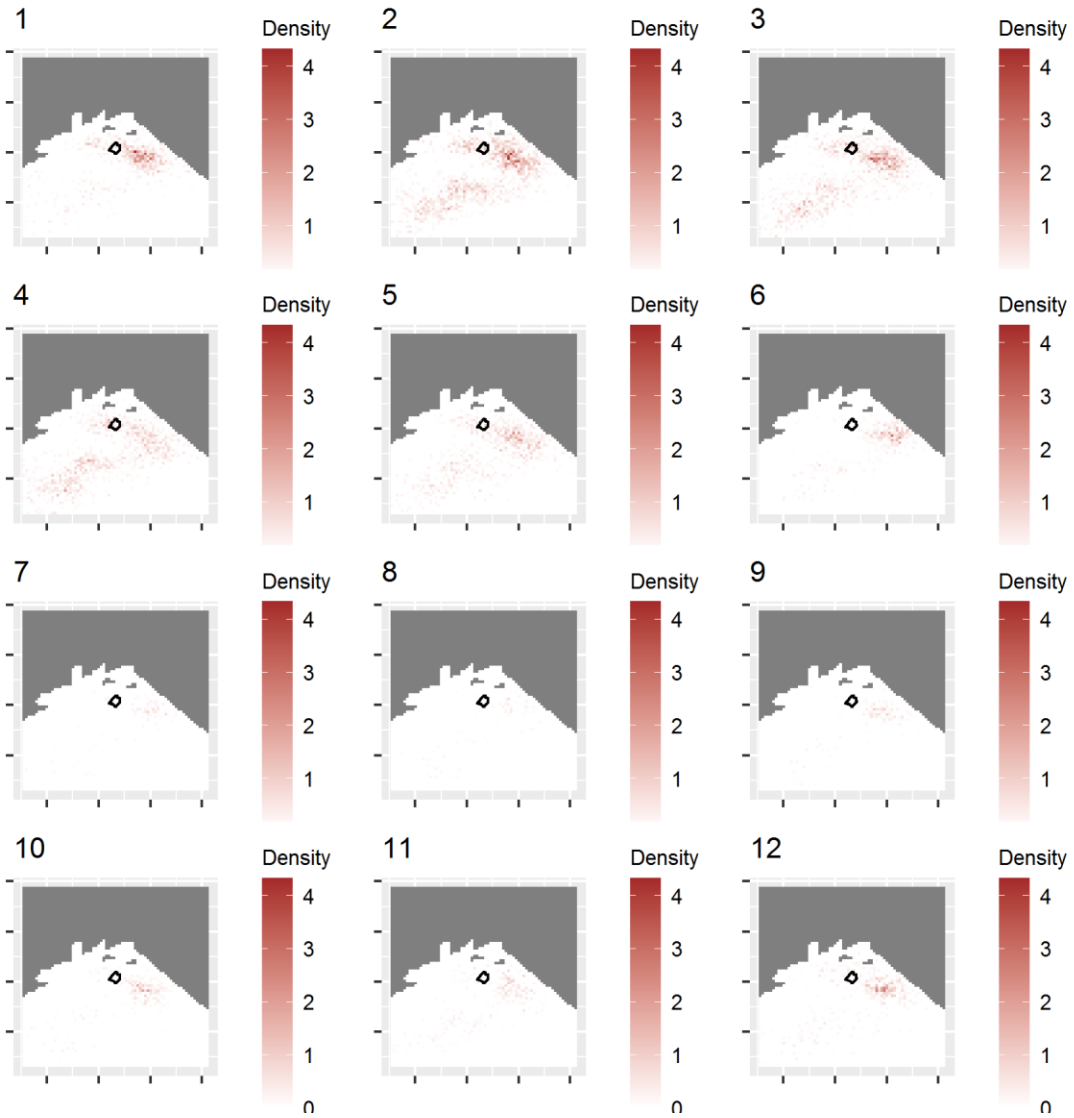


Figure A-10. North Atlantic right whale VYWA

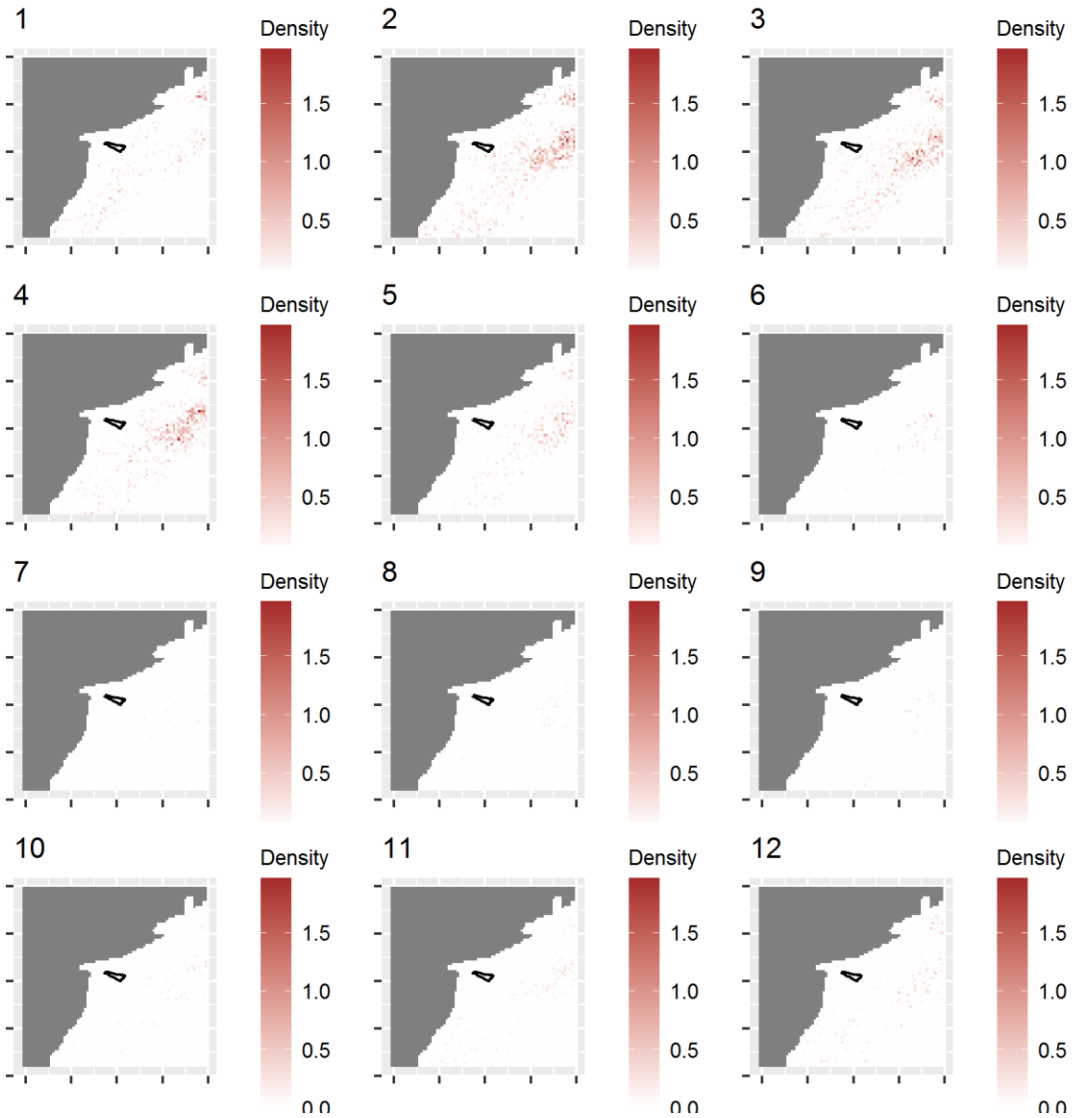


Figure A-11. North Atlantic right whale EOWA

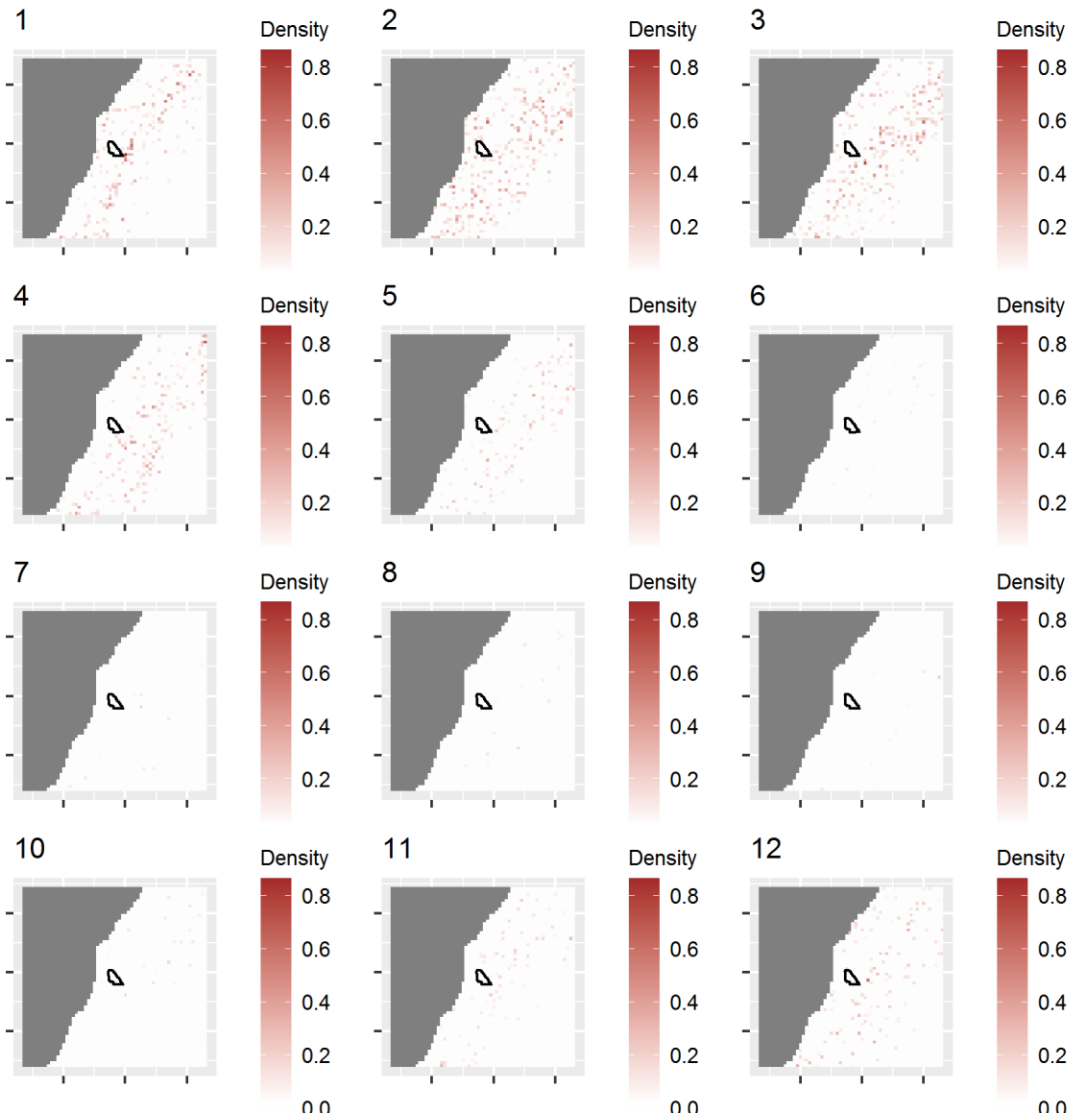


Figure A-12. North Atlantic right whale MAWA

Appendix B. Van Parijs et al. PAM Design and Species' Effective Detection Ranges by Wind Energy Area

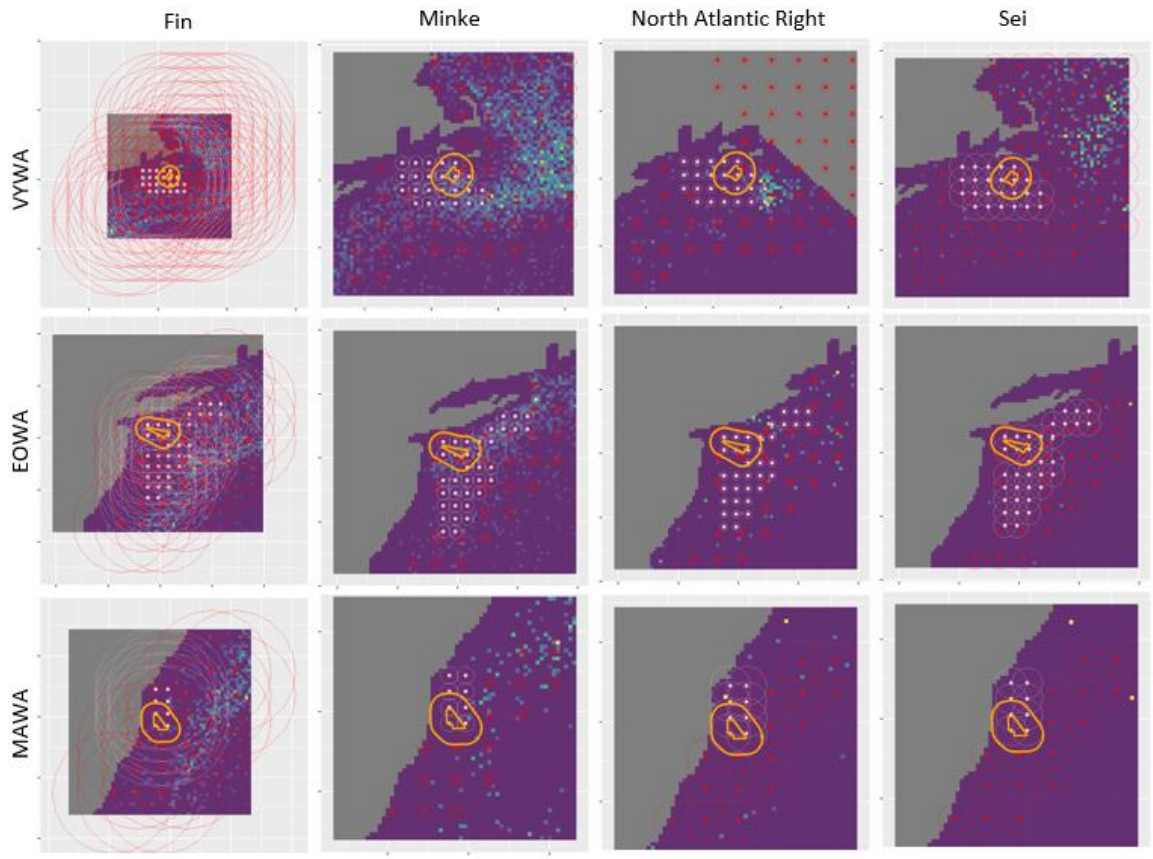


Figure B-1. Area covered by all the PAM stations as suggested by Van Parijs et al. for the large monitoring site for all wind farm areas
Red circles show effective detection area for all four species (

Table 5). Pink dots and corresponding buffers show the small PAM grid and red dots show the large PAM grid. Orange lines show the small monitoring area.

Appendix C. Mean Whale Density and Number Responding by Wind Energy Area

Table C-1. Mean number of responding whales over 500 realizations of the density surfaces for each month of construction at each wind farms

Two dose-response functions (DR1, and DR2), defining the probability of response as a function of distance from construction, were used. Effect size for each DR was calculated as proportion of all whales at the small monitoring area which responded. Densities are in number of whales per 25 km^2. The mean number of detected cues is calculated for baseline density for all PAM grids at the large monitoring area and is calculated as mean number of cues (cues per month) at the grids overlapping with EDR of each species.

Species	Wind Farm	Month of Construction	Mean Responding (DR1)	Mean Responding (DR2)	Effect Size (DR1)	Effect Size (DR2)	Mean Density in Large Monitoring Area	Mean Density in Small Monitoring Area	Mean Number of Detected Cues
Fin	VYWA	5	0.32	0.91	0.112	0.04	0.1	0.05	8,490,000
Fin	VYWA	6	0.36	0.95	0.135	0.051	0.13	0.04	10,380,000
Fin	VYWA	7	0.46	1.15	0.118	0.047	0.16	0.06	13,260,000
Fin	VYWA	8	0.54	1.35	0.157	0.063	0.14	0.06	12,150,000
Fin	VYWA	9	0.12	0.33	0.08	0.029	0.12	0.03	9,810,000
Fin	VYWA	10	0.05	0.12	0.098	0.037	0.09	0.01	6,750,000
Minke	VYWA	5	2.59	7.39	0.112	0.039	0.23	0.43	153,000
Minke	VYWA	6	3.39	8.89	0.144	0.055	0.27	0.39	172,500
Minke	VYWA	7	1.62	4.18	0.146	0.056	0.15	0.18	96,750
Minke	VYWA	8	0.78	2.00	0.138	0.054	0.09	0.09	57,300
Minke	VYWA	9	0.46	1.21	0.101	0.038	0.1	0.08	58,500
Minke	VYWA	10	0.33	0.89	0.115	0.044	0.1	0.05	56,850
Sei	VYWA	5	0.20	0.59	0.11	0.038	0.13	0.03	255,000
Sei	VYWA	6	0.06	0.16	0.134	0.05	0.05	0.01	105,000
Sei	VYWA	7	0.01	0.04	0.107	0.041	0.01	0.005	25,590
Sei	VYWA	8	0.01	0.03	0.142	0.06	0.01	0.005	15,690
Sei	VYWA	9	0.02	0.04	0.112	0.046	0.01	0.005	27,210
Sei	VYWA	10	0.02	0.06	0.105	0.036	0.02	0.005	48,300
NARW	VYWA	5	0.73	2.10	0.115	0.04	0.06	0.12	3,600
NARW	VYWA	6	0.14	0.37	0.131	0.05	0.02	0.02	1,530
NARW	VYWA	7	0.07	0.18	0.141	0.056	0.01	0.01	780
NARW	VYWA	8	0.06	0.14	0.138	0.056	0	0.01	1,242
NARW	VYWA	9	0.08	0.22	0.115	0.044	0.01	0.01	1,440
NARW	VYWA	10	0.17	0.44	0.15	0.059	0.01	0.02	840
Fin	EOWA	4	0.25	0.63	0.087	0.034	0.06	0.04	5,400,000
Fin	EOWA	5	0.12	0.37	0.038	0.013	0.07	0.05	6,690,000
Fin	EOWA	6	0.44	1.16	0.111	0.043	0.08	0.05	7,500,000
Fin	EOWA	7	0.19	0.54	0.05	0.017	0.09	0.05	8,790,000
Fin	EOWA	8	0.08	0.26	0.037	0.012	0.06	0.03	6,540,000
Fin	EOWA	9	0.07	0.18	0.039	0.015	0.09	0.02	8,940,000
Fin	EOWA	10	0.05	0.15	0.068	0.024	0.05	0.01	3,870,000

Species	Wind Farm	Month of Construction	Mean Responding (DR1)	Mean Responding (DR2)	Effect Size (DR1)	Effect Size (DR2)	Mean Density in Large Monitoring Area	Mean Density in Small Monitoring Area	Mean Number of Detected Cues
Fin	EOWA	11	0.05	0.14	0.067	0.023	0.03	0.01	2,124,000
Fin	EOWA	12	0.15	0.43	0.066	0.022	0.04	0.03	3,480,000
Minke	EOWA	4	1.58	4.01	0.097	0.038	0.14	0.2	97,200
Minke	EOWA	5	1.46	4.26	0.056	0.019	0.24	0.37	184,500
Minke	EOWA	6	1.92	5.09	0.107	0.04	0.17	0.23	147,750
Minke	EOWA	7	0.31	0.89	0.071	0.025	0.04	0.06	36,900
Minke	EOWA	8	0.25	0.71	0.061	0.021	0.03	0.06	30,600
Minke	EOWA	9	0.08	0.22	0.030	0.011	0.04	0.03	33,450
Minke	EOWA	10	0.08	0.24	0.045	0.015	0.04	0.03	25,950
Minke	EOWA	11	0.04	0.11	0.073	0.025	0.01	0.01	5,640
Minke	EOWA	12	0.07	0.22	0.068	0.023	0.01	0.02	9,525
Sei	EOWA	4	0.17	0.41	0.096	0.038	0.03	0.02	54,990
Sei	EOWA	5	0.05	0.15	0.04	0.014	0.03	0.02	72,810
Sei	EOWA	6	0.02	0.05	0.082	0.031	0.01	0.01	23,700
Sei	EOWA	7	0	0	0.058	0.021	0.01	0.005	2,640
Sei	EOWA	8	0	0	0.065	0.021	0.01	0.005	1,140
Sei	EOWA	9	0.01	0.02	0.074	0.028	0.01	0.005	3,690
Sei	EOWA	10	0.02	0.04	0.05	0.018	0.01	0.005	9,330
Sei	EOWA	11	0.05	0.15	0.08	0.028	0.01	0.005	14,700
Sei	EOWA	12	0.04	0.11	0.049	0.016	0.01	0.005	13,710
NARW	EOWA	4	0.25	0.64	0.116	0.046	0.03	0.03	3,030
NARW	EOWA	5	0.03	0.08	0.043	0.014	0.01	0.01	4,890
NARW	EOWA	6	0.02	0.04	0.104	0.039	0.005	0.005	909
NARW	EOWA	7	0.01	0.02	0.071	0.023	0.005	0.005	489
NARW	EOWA	8	0.01	0.02	0.074	0.026	0.005	0.005	468
NARW	EOWA	9	0.01	0.02	0.07	0.027	0.005	0.005	612
NARW	EOWA	10	0.01	0.02	0.047	0.016	0.005	0.005	996
NARW	EOWA	11	0.04	0.10	0.087	0.031	0.005	0.01	1,509
NARW	EOWA	12	0.08	0.22	0.070	0.024	0.01	0.02	3,480
Fin	MAWA	4	0.21	0.53	0.091	0.036	0.07	0.03	7,320,000
Fin	MAWA	5	0.10	0.28	0.063	0.023	0.07	0.03	7,140,000
Fin	MAWA	6	0.17	0.44	0.088	0.033	0.09	0.03	8,850,000
Fin	MAWA	7	0.09	0.22	0.110	0.045	0.05	0.01	4,860,000
Fin	MAWA	8	0.06	0.15	0.125	0.046	0.03	0.01	3,060,000
Fin	MAWA	9	0.07	0.18	0.094	0.036	0.06	0.01	5,070,000
Minke	MAWA	4	1.02	2.47	0.105	0.043	0.14	0.14	156,600
Minke	MAWA	5	0.82	2.22	0.078	0.029	0.18	0.17	208,800
Minke	MAWA	6	0.25	0.72	0.059	0.021	0.09	0.07	98,700
Minke	MAWA	7	0.13	0.31	0.118	0.047	0.02	0.01	18,120

Species	Wind Farm	Month of Construction	Mean Responding (DR1)	Mean Responding (DR2)	Effect Size (DR1)	Effect Size (DR2)	Mean Density in Large Monitoring Area	Mean Density in Small Monitoring Area	Mean Number of Detected Cues
Minke	MAWA	8	0.08	0.20	0.115	0.045	0.01	0.01	13,770
Minke	MAWA	9	0.02	0.06	0.063	0.025	0.02	0.01	19,680
Sei	MAWA	4	0.08	0.19	0.078	0.032	0.02	0.01	46,410
Sei	MAWA	5	0.03	0.07	0.067	0.024	0.02	0.01	30,300
Sei	MAWA	6	0.01	0.02	0.06	0.024	0.01	0.005	12,000
Sei	MAWA	7	0	0	0.119	0.046	0.005	0.005	780
Sei	MAWA	8	0	0	0.119	0.043	0.005	0.005	450
Sei	MAWA	9	0	0.01	0.083	0.025	0.005	0.005	1,470
NARW	MAWA	4	0.07	0.17	0.067	0.025	0.02	0.01	14,100
NARW	MAWA	5	0.01	0.03	0.054	0.019	0.005	0.005	8,700
NARW	MAWA	6	0	0.01	0.066	0.024	0.005	0.005	11,580
NARW	MAWA	7	0	0.01	0.093	0.037	0.005	0.005	4,890
NARW	MAWA	8	0	0.01	0.105	0.038	0.005	0.005	4,530
NARW	MAWA	9	0	0.01	0.077	0.026	0.005	0.005	7,530

Table C-2. Mean number of responding whales per month given operation of all the wind farms
Mean and median were calculated over 500 realizations of the density surfaces.

Species	Month of Operation	Mean Responding
Fin	1	17.43
Fin	2	16.60
Fin	3	13.98
Fin	4	14.29
Fin	5	18.41
Fin	6	18.76
Fin	7	24.92
Fin	8	18.43
Fin	9	13.70
Fin	10	4.63
Fin	11	4.29
Fin	12	11.87
Minke	1	7.86
Minke	2	8.30
Minke	3	8.70
Minke	4	62.23
Minke	5	121.93
Minke	6	97.40
Minke	7	38.03

Species	Month of Operation	Mean Responding
Minke	8	21.37
Minke	9	20.69
Minke	10	14.36
Minke	11	3.47
Minke	12	6.23
Sei	1	2.19
Sei	2	2.22
Sei	3	4.58
Sei	4	7.21
Sei	5	9.92
Sei	6	2.44
Sei	7	0.48
Sei	8	0.28
Sei	9	0.61
Sei	10	1.60
Sei	11	2.96
Sei	12	3.64
NARW	1	19.51
NARW	2	26.93
NARW	3	24.76
NARW	4	21.00
NARW	5	16.82
NARW	6	2.29
NARW	7	1.04
NARW	8	0.82
NARW	9	1.55
NARW	10	2.59
NARW	11	3.57
NARW	12	9.74

Appendix D. Power Estimates for Van Parijs et al. Design.

D.1 Local Analysis

Table D-1. Statistical power (%) to detect change in cue rates under eight hypotheses (see Table 2) for each studied species, wind farm area and combination of monitoring size and PAM grid
H1 (no effect) was calculated both for construction phase (H1_co) and operation (H1_op). For hypotheses related to construction of the windfarms (H2–H5), power to detect change was calculated for two dose-response functions: DR1 and DR2. Scenarios with power >= 80% are marked in bold and with an asterisk.

Design	Hypothesis	Species	Wind Farm	Large Monitoring Area & Both PAM Grids	Large Monitoring Area & Small PAM Grid Only	Small Monitoring Area & Small PAM Grid Only
Van Parijs et al.	H1_co	Minke	MAWA	5.2	6.2	8.7
Van Parijs et al.	H1_op	Minke	MAWA	7.2	7.6	10.2
Van Parijs et al.	H1_co	Sei	MAWA	12.2	15.6	16.9
Van Parijs et al.	H1_op	Sei	MAWA	11.7	17.6	21.1
Van Parijs et al.	H1_co	Fin	MAWA	7.5	8.3	14.4
Van Parijs et al.	H1_op	Fin	MAWA	8.2	9.2	17.1
Van Parijs et al.	H1_co	NARW	MAWA	10.2	12.4	13.9
Van Parijs et al.	H1_op	NARW	MAWA	14.2	14.8	19.5
Van Parijs et al.	H1_co	Minke	VYWA	5.2	6.4	10.8
Van Parijs et al.	H1_op	Minke	VYWA	6.2	8.2	15.8
Van Parijs et al.	H1_co	Sei	VYWA	12.3	14.3	14.7
Van Parijs et al.	H1_op	Sei	VYWA	14.2	16.8	23.8
Van Parijs et al.	H1_co	Fin	VYWA	7.8	8.3	14.5
Van Parijs et al.	H1_op	Fin	VYWA	15.6	16	19.3
Van Parijs et al.	H1_co	NARW	VYWA	12.4	12.8	20.0
Van Parijs et al.	H1_op	NARW	VYWA	13.2	14.7	20.4
Van Parijs et al.	H1_co	Minke	EOWA	8.9	12	17.1
Van Parijs et al.	H1_op	Minke	EOWA	9.3	11.3	19.0
Van Parijs et al.	H1_co	Sei	EOWA	11.8	13.2	18.6
Van Parijs et al.	H1_op	Sei	EOWA	12.7	16.8	21.5
Van Parijs et al.	H1_co	Fin	EOWA	8.6	9.2	16.0
Van Parijs et al.	H1_op	Fin	EOWA	11.6	13.2	15.3
Van Parijs et al.	H1_co	NARW	EOWA	10.4	12.6	13.8
Van Parijs et al.	H1_op	NARW	EOWA	11.4	12.8	14.3
Van Parijs et al.	H2_50_DR1	Minke	MAWA	5	11	13.9
Van Parijs et al.	H2_50_DR2	Minke	MAWA	3.8	5.8	12.2
Van Parijs et al.	H2_100_DR1	Minke	MAWA	10.6	13.6	14.4
Van Parijs et al.	H2_100_DR2	Minke	MAWA	11.2	17.2	21.4
Van Parijs et al.	H2_50_DR1	Sei	MAWA	15	17	18.2
Van Parijs et al.	H2_50_DR2	Sei	MAWA	14.2	21.2	27.3
Van Parijs et al.	H2_100_DR1	Sei	MAWA	1.4	6.4	9.6
Van Parijs et al.	H2_100_DR2	Sei	MAWA	8.2	14.2	14.7

Design	Hypothesis	Species	Wind Farm	Large Monitoring Area & Both PAM Grids	Large Monitoring Area & Small PAM Grid Only	Small Monitoring Area & Small PAM Grid Only
Van Parijs et al.	H2_50_DR1	Fin	MAWA	7.1	14.1	19.8
Van Parijs et al.	H2_50_DR2	Fin	MAWA	3.4	6.4	11.4
Van Parijs et al.	H2_100_DR1	Fin	MAWA	0.3	4.3	4.6
Van Parijs et al.	H2_100_DR2	Fin	MAWA	0.7	6.7	14.2
Van Parijs et al.	H2_50_DR1	NARW	MAWA	2.8	7.8	15.7
Van Parijs et al.	H2_50_DR2	NARW	MAWA	3.4	10.4	12.4
Van Parijs et al.	H2_100_DR1	NARW	MAWA	13.6	19.6	25.7
Van Parijs et al.	H2_100_DR2	NARW	MAWA	18.2	23.2	27.6
Van Parijs et al.	H2_50_DR1	Minke	VYWA	10.2	16.2	23.5
Van Parijs et al.	H2_50_DR2	Minke	VYWA	5.2	12.2	13.0
Van Parijs et al.	H2_100_DR1	Minke	VYWA	16.7	22.7	25.5
Van Parijs et al.	H2_100_DR2	Minke	VYWA	22.1	25.1	25.3
Van Parijs et al.	H2_50_DR1	Sei	VYWA	2.6	8.6	15.4
Van Parijs et al.	H2_50_DR2	Sei	VYWA	2.8	5.8	9.1
Van Parijs et al.	H2_100_DR1	Sei	VYWA	8.9	12.9	17.1
Van Parijs et al.	H2_100_DR2	Sei	VYWA	12.3	18.3	25.2
Van Parijs et al.	H2_50_DR1	Fin	VYWA	6	9	12.7
Van Parijs et al.	H2_50_DR2	Fin	VYWA	6	8	14.3
Van Parijs et al.	H2_100_DR1	Fin	VYWA	4	9	10.8
Van Parijs et al.	H2_100_DR2	Fin	VYWA	5.4	12.4	18.9
Van Parijs et al.	H2_50_DR1	NARW	VYWA	4.4	6.4	12.7
Van Parijs et al.	H2_50_DR2	NARW	VYWA	4.2	8.2	10.7
Van Parijs et al.	H2_100_DR1	NARW	VYWA	14.6	20.6	25.4
Van Parijs et al.	H2_100_DR2	NARW	VYWA	17.2	22.2	26.7
Van Parijs et al.	H2_50_DR1	Minke	EOWA	8.2	12.2	17.9
Van Parijs et al.	H2_50_DR2	Minke	EOWA	9.4	14.4	18.3
Van Parijs et al.	H2_100_DR1	Minke	EOWA	16.7	23.7	26.9
Van Parijs et al.	H2_100_DR2	Minke	EOWA	17.2	21.2	22.9
Van Parijs et al.	H2_50_DR1	Sei	EOWA	6.4	12.4	20.3
Van Parijs et al.	H2_50_DR2	Sei	EOWA	4.5	10.5	14.8
Van Parijs et al.	H2_100_DR1	Sei	EOWA	4.6	9.6	11.2
Van Parijs et al.	H2_100_DR2	Sei	EOWA	8.2	10.2	12.3
Van Parijs et al.	H2_50_DR1	Fin	EOWA	9.4	14.4	22.2
Van Parijs et al.	H2_50_DR2	Fin	EOWA	9.2	13.2	20.6
Van Parijs et al.	H2_100_DR1	Fin	EOWA	8.6	15.6	20.0
Van Parijs et al.	H2_100_DR2	Fin	EOWA	8.8	12.8	13.9
Van Parijs et al.	H2_50_DR1	NARW	EOWA	5.4	12.4	16.9
Van Parijs et al.	H2_50_DR2	NARW	EOWA	4.6	6.6	9.8
Van Parijs et al.	H2_100_DR1	NARW	EOWA	14.6	21.6	28.5
Van Parijs et al.	H2_100_DR2	NARW	EOWA	15.8	20.8	22.5

Design	Hypothesis	Species	Wind Farm	Large Monitoring Area & Both PAM Grids	Large Monitoring Area & Small PAM Grid Only	Small Monitoring Area & Small PAM Grid Only
Van Parijs et al.	H3_DR1	Minke	MAWA	40.2	44.2	47.3
Van Parijs et al.	H3_DR2	Minke	MAWA	44.7	46.7	49.3
Van Parijs et al.	H3_DR1	Sei	MAWA	14.8	17.8	25.5
Van Parijs et al.	H3_DR2	Sei	MAWA	13.8	19.8	24.5
Van Parijs et al.	H3_DR1	Fin	MAWA	0.1	0.6	6.0
Van Parijs et al.	H3_DR2	Fin	MAWA	0.7	0.2	6.4
Van Parijs et al.	H3_DR1	NARW	MAWA	43	50	55.7
Van Parijs et al.	H3_DR2	NARW	MAWA	43.2	49.2	50.8
Van Parijs et al.	H3_DR1	Minke	VYWA	38.2	41.2	46.6
Van Parijs et al.	H3_DR2	Minke	VYWA	41.3	46.3	47.6
Van Parijs et al.	H3_DR1	Sei	VYWA	23	29	34.3
Van Parijs et al.	H3_DR2	Sei	VYWA	23	27	31.7
Van Parijs et al.	H3_DR1	Fin	VYWA	5.8	0.5	6.6
Van Parijs et al.	H3_DR2	Fin	VYWA	5.4	1.1	4.7
Van Parijs et al.	H3_DR1	NARW	VYWA	24.4	26.4	33.6
Van Parijs et al.	H3_DR2	NARW	VYWA	28.1	31.1	36.5
Van Parijs et al.	H3_DR1	Minke	EOWA	40.7	46.7	47.6
Van Parijs et al.	H3_DR2	Minke	EOWA	42.2	44.2	47.1
Van Parijs et al.	H3_DR1	Sei	EOWA	15.4	20.4	21.7
Van Parijs et al.	H3_DR2	Sei	EOWA	18.2	22.2	22.5
Van Parijs et al.	H3_DR1	Fin	EOWA	8.4	7.9	12.1
Van Parijs et al.	H3_DR2	Fin	EOWA	8.4	2	9.6
Van Parijs et al.	H3_DR1	NARW	EOWA	15.8	21.8	25.1
Van Parijs et al.	H3_DR2	NARW	EOWA	17.9	21.9	23.6
Van Parijs et al.	H4_DR1	Minke	MAWA	38.7	44.7	50.6
Van Parijs et al.	H4_DR2	Minke	MAWA	41.2	47.2	49.9
Van Parijs et al.	H4_DR1	Sei	MAWA	15	17	22.4
Van Parijs et al.	H4_DR2	Sei	MAWA	14	20	23.6
Van Parijs et al.	H4_DR1	Fin	MAWA	0.4	0	2.2
Van Parijs et al.	H4_DR2	Fin	MAWA	2	0	2.4
Van Parijs et al.	H4_DR1	NARW	MAWA	43	49	52.0
Van Parijs et al.	H4_DR2	NARW	MAWA	43.2	47.2	53.8
Van Parijs et al.	H4_DR1	Minke	VYWA	38.5	43.5	46.2
Van Parijs et al.	H4_DR2	Minke	VYWA	42.5	49.5	52.9
Van Parijs et al.	H4_DR1	Sei	VYWA	23	28	34.3
Van Parijs et al.	H4_DR2	Sei	VYWA	23	29	33.4
Van Parijs et al.	H4_DR1	Fin	VYWA	0.7	0	5.7
Van Parijs et al.	H4_DR2	Fin	VYWA	5.4	0	5.1
Van Parijs et al.	H4_DR1	NARW	VYWA	17.4	20.4	23.2
Van Parijs et al.	H4_DR2	NARW	VYWA	18.2	23.2	23.5

Design	Hypothesis	Species	Wind Farm	Large Monitoring Area & Both PAM Grids	Large Monitoring Area & Small PAM Grid Only	Small Monitoring Area & Small PAM Grid Only
Van Parijs et al.	H4_DR1	Minke	EOWA	44.7	50.7	55.5
Van Parijs et al.	H4_DR2	Minke	EOWA	47.2	50.2	55.1
Van Parijs et al.	H4_DR1	Sei	EOWA	15.4	18.4	23.0
Van Parijs et al.	H4_DR2	Sei	EOWA	18.2	20.2	23.2
Van Parijs et al.	H4_DR1	Fin	EOWA	8.4	2	7.6
Van Parijs et al.	H4_DR2	Fin	EOWA	8.6	2	9.3
Van Parijs et al.	H4_DR1	NARW	EOWA	15.2	18.2	19.2
Van Parijs et al.	H4_DR2	NARW	EOWA	17.3	22.3	25.3
Van Parijs et al.	H5_DR1	Minke	MAWA	43.6	49.6	54.0
Van Parijs et al.	H5_DR2	Minke	MAWA	48.2	51.2	58.2
Van Parijs et al.	H5_DR1	Sei	MAWA	15	19	20.1
Van Parijs et al.	H5_DR2	Sei	MAWA	14	18	20.4
Van Parijs et al.	H5_DR1	Fin	MAWA	1.1	1.5	3.9
Van Parijs et al.	H5_DR2	Fin	MAWA	2.3	0.8	7.7
Van Parijs et al.	H5_DR1	NARW	MAWA	43	49	52.3
Van Parijs et al.	H5_DR2	NARW	MAWA	43.2	50.2	52.6
Van Parijs et al.	H5_DR1	Minke	VYWA	51.3	56.3	60.8
Van Parijs et al.	H5_DR2	Minke	VYWA	55.8	60.8	65.0
Van Parijs et al.	H5_DR1	Sei	VYWA	23	28	33.4
Van Parijs et al.	H5_DR2	Sei	VYWA	23	26	27.6
Van Parijs et al.	H5_DR1	Fin	VYWA	6	3.4	5.3
Van Parijs et al.	H5_DR2	Fin	VYWA	5.4	1.2	2.4
Van Parijs et al.	H5_DR1	NARW	VYWA	16.8	18.8	25.3
Van Parijs et al.	H5_DR2	NARW	VYWA	18.2	20.2	21.2
Van Parijs et al.	H5_DR1	Minke	EOWA	49.6	55.6	62.1
Van Parijs et al.	H5_DR2	Minke	EOWA	51.2	56.2	61.3
Van Parijs et al.	H5_DR1	Sei	EOWA	15.4	20.4	25.6
Van Parijs et al.	H5_DR2	Sei	EOWA	15.2	20.2	23.4
Van Parijs et al.	H5_DR1	Fin	EOWA	8.4	2	4.6
Van Parijs et al.	H5_DR2	Fin	EOWA	8.6	2	3.9
Van Parijs et al.	H5_DR1	NARW	EOWA	18	23	26.7
Van Parijs et al.	H5_DR2	NARW	EOWA	21.3	24.3	29.0
Van Parijs et al.	H6	Minke	VYWA	15.8	21.8	27.2
Van Parijs et al.	H6	Sei	VYWA	11.8	15.8	18.5
Van Parijs et al.	H6	Fin	VYWA	1.6	7.6	8.3
Van Parijs et al.	H6	NARW	VYWA	9.8	11.8	19.3
Van Parijs et al.	H6	Minke	EOWA	0.6	7.6	14.3
Van Parijs et al.	H6	Sei	EOWA	1.6	6.6	12.8
Van Parijs et al.	H6	Fin	EOWA	13.8	18.8	24.1
Van Parijs et al.	H6	NARW	EOWA	12.2	19.2	24.7

Design	Hypothesis	Species	Wind Farm	Large Monitoring Area & Both PAM Grids	Large Monitoring Area & Small PAM Grid Only	Small Monitoring Area & Small PAM Grid Only
Van Parijs et al.	H7	Minke	MAWA	24.4	30.4	34.3
Van Parijs et al.	H7	Sei	MAWA	18.6	20.6	27.5
Van Parijs et al.	H7	Fin	MAWA	2.5	4.5	5.0
Van Parijs et al.	H7	NARW	MAWA	6.6	9.6	9.8
Van Parijs et al.	H7	Minke	VYWA	25.2	31.2	33.3
Van Parijs et al.	H7	Sei	VYWA	7.8	12.8	18.1
Van Parijs et al.	H7	Fin	VYWA	1.6	6.6	11.2
Van Parijs et al.	H7	NARW	VYWA	19.8	24.8	27.0
Van Parijs et al.	H7	Minke	EOWA	1	8	15.2
Van Parijs et al.	H7	Sei	EOWA	0.4	4.4	10.0
Van Parijs et al.	H7	Fin	EOWA	8.4	12.4	17.4
Van Parijs et al.	H7	NARW	EOWA	17	24	25.9
Van Parijs et al.	H8	Minke	MAWA	*** 97**	*** 99.8**	*** 95.6**
Van Parijs et al.	H8	Sei	MAWA	*** 97.4**	*** 87**	*** 88.2**
Van Parijs et al.	H8	Fin	MAWA	69	*** 93.1**	*** 96.2**
Van Parijs et al.	H8	NARW	MAWA	*** 88.8**	78	*** 81.6**
Van Parijs et al.	H8	Minke	VYWA	*** 94**	*** 86**	*** 89.9**
Van Parijs et al.	H8	Sei	VYWA	*** 89.8**	*** 82.4**	*** 54.0**
Van Parijs et al.	H8	Fin	VYWA	69.8	*** 99**	*** 95.7**
Van Parijs et al.	H8	NARW	VYWA	*** 84**	*** 94.1**	*** 94.8**
Van Parijs et al.	H8	Minke	EOWA	*** 91.2**	*** 95.8**	*** 91.8**
Van Parijs et al.	H8	Sei	EOWA	*** 94.8**	*** 91.6**	*** 96.3**
Van Parijs et al.	H8	Fin	EOWA	79	*** 99.8**	*** 94.7**
Van Parijs et al.	H8	NARW	EOWA	*** 88.4**	*** 84.8**	*** 90.4**

D.2 Regional Analysis

Table D-2. Statistical power (%) to detect change in cue rates under two hypotheses (see Table 2) related to operation of wind farms for each studied species, combination of monitoring size and PAM grid and two lengths of monitoring: 1 and 5 years
Scenarios with power >= 80% are marked in bold and with an asterisk.

Design	Hypothesis	Species	Windfarm	Both PAM Grids	Small PAM Grid Only
Van Parijs et al.	H1_op_1y	Minke	Regional	7.2	9.1
Van Parijs et al.	H1_op_1y	Fin	Regional	9.7	15.5
Van Parijs et al.	H1_op_1y	Sei	Regional	6.9	7.2
Van Parijs et al.	H1_op_1y	NARW	Regional	5.3	9.3
Van Parijs et al.	H6_1y	Minke	Regional	***88.3**	***90**
Van Parijs et al.	H6_1y	Sei	Regional	45.4	47.5
Van Parijs et al.	H6_5y	Sei	Regional	56.2	66
Van Parijs et al.	H6_1y	Fin	Regional	4.2	3.5
Van Parijs et al.	H6_5y	Fin	Regional	4.1	8.3
Van Parijs et al.	H6_1y	NARW	Regional	41.8	41.3
Van Parijs et al.	H6_5y	NARW	Regional	52.7	52.8
Van Parijs et al.	H7_1y	Minke	Regional	***82.4**	***87.2**
Van Parijs et al.	H7_1y	Sei	Regional	55.1	55.7
Van Parijs et al.	H7_5y	Sei	Regional	67.1	69.2
Van Parijs et al.	H7_1y	Fin	Regional	11.3	15.2
Van Parijs et al.	H7_5y	Fin	Regional	12.4	14.3
Van Parijs et al.	H7_1y	NARW	Regional	44.5	51.3
Van Parijs et al.	H7_5y	NARW	Regional	68.2	66

Appendix E. Power Estimates for Alternative PAM Designs

E.1 Local Analysis

E.1.1 10 x 10 km Grid

Table E-1. Statistical power (%) to detect change in cue rates under eight hypotheses (see Table 2) for each studied species, wind farm area and combination of monitoring size and PAM grid assuming modified small PAM grid using 10 x 10 km grid

For hypotheses related to construction of the windfarms (H2–H5), power to detect change was calculated for two dose-response functions: DR1 and DR2. For alternative designs estimating power for small monitoring area and modified small PAM only was not conducted. Scenarios with power >= 80% are marked in bold and with an asterisk.

Design	Hypothesis	Species	Wind Farm	Large Monitoring Area & Both PAM Grids	Large Monitoring Area & Modified Small PAM Grid Only
10 x 10 km	H1_co	Minke	MAWA	5.7	8.1
10 x 10 km	H1_op	Minke	MAWA	7.3	10.5
10 x 10 km	H1_co	Sei	MAWA	10.1	15.8
10 x 10 km	H1_op	Sei	MAWA	13.0	17.1
10 x 10 km	H1_co	Fin	MAWA	5.0	9.4
10 x 10 km	H1_op	Fin	MAWA	5.6	13.9
10 x 10 km	H1_co	NARW	MAWA	10.9	14.3
10 x 10 km	H1_op	NARW	MAWA	15.3	24.8
10 x 10 km	H1_co	Minke	VYWA	5.2	8.9
10 x 10 km	H1_op	Minke	VYWA	6.3	11.7
10 x 10 km	H1_co	Sei	VYWA	13.9	20.1
10 x 10 km	H1_op	Sei	VYWA	14.4	17.6
10 x 10 km	H1_co	Fin	VYWA	5.1	9.3
10 x 10 km	H1_op	Fin	VYWA	8.0	14.6
10 x 10 km	H1_co	NARW	VYWA	12.8	16.2
10 x 10 km	H1_op	NARW	VYWA	12.4	20.5
10 x 10 km	H1_co	Minke	EOWA	8.9	13.2
10 x 10 km	H1_op	Minke	EOWA	11.1	15.6
10 x 10 km	H1_co	Sei	EOWA	12.1	16.8
10 x 10 km	H1_op	Sei	EOWA	13.1	17.0
10 x 10 km	H1_co	Fin	EOWA	8.7	10.8
10 x 10 km	H1_op	Fin	EOWA	5.7	11.0
10 x 10 km	H1_co	NARW	EOWA	11.2	16.4
10 x 10 km	H1_op	NARW	EOWA	11.1	19.2
10 x 10 km	H2_50_DR1	Minke	MAWA	42.0	49.7
10 x 10 km	H2_50_DR2	Minke	MAWA	44.6	52.9
10 x 10 km	H2_100_DR1	Minke	MAWA	49.3	53.0
10 x 10 km	H2_100_DR2	Minke	MAWA	51.6	60.3
10 x 10 km	H2_50_DR1	Sei	MAWA	42.0	52.2
10 x 10 km	H2_50_DR2	Sei	MAWA	42.0	49.9

Design	Hypothesis	Species	Wind Farm	Large Monitoring Area & Both PAM Grids	Large Monitoring Area & Modified Small PAM Grid Only
10 x 10 km	H2_100_DR1	Sei	MAWA	41.5	46.0
10 x 10 km	H2_100_DR2	Sei	MAWA	43.1	53.6
10 x 10 km	H2_50_DR1	Fin	MAWA	2.2	5.0
10 x 10 km	H2_50_DR2	Fin	MAWA	8.5	18.5
10 x 10 km	H2_100_DR1	Fin	MAWA	1.9	4.8
10 x 10 km	H2_100_DR2	Fin	MAWA	8.0	15.9
10 x 10 km	H2_50_DR1	NARW	MAWA	38.0	45.6
10 x 10 km	H2_50_DR2	NARW	MAWA	25.9	36.4
10 x 10 km	H2_100_DR1	NARW	MAWA	26.8	37.5
10 x 10 km	H2_100_DR2	NARW	MAWA	35.2	44.6
10 x 10 km	H2_50_DR1	Minke	VYWA	50.3	57.6
10 x 10 km	H2_50_DR2	Minke	VYWA	52.8	58.8
10 x 10 km	H2_100_DR1	Minke	VYWA	56.9	67.3
10 x 10 km	H2_100_DR2	Minke	VYWA	58.4	65.9
10 x 10 km	H2_50_DR1	Sei	VYWA	25.3	31.1
10 x 10 km	H2_50_DR2	Sei	VYWA	23.8	28.7
10 x 10 km	H2_100_DR1	Sei	VYWA	24.3	27.2
10 x 10 km	H2_100_DR2	Sei	VYWA	23.6	30.7
10 x 10 km	H2_50_DR1	Fin	VYWA	4.5	12.3
10 x 10 km	H2_50_DR2	Fin	VYWA	5.2	12.1
10 x 10 km	H2_100_DR1	Fin	VYWA	7.8	13.1
10 x 10 km	H2_100_DR2	Fin	VYWA	1.0	6.0
10 x 10 km	H2_50_DR1	NARW	VYWA	33.8	37.7
10 x 10 km	H2_50_DR2	NARW	VYWA	27.1	34.7
10 x 10 km	H2_100_DR1	NARW	VYWA	28.2	34.4
10 x 10 km	H2_100_DR2	NARW	VYWA	27.1	30.1
10 x 10 km	H2_50_DR1	Minke	EOWA	44.3	49.3
10 x 10 km	H2_50_DR2	Minke	EOWA	48.8	56.5
10 x 10 km	H2_100_DR1	Minke	EOWA	52.3	54.8
10 x 10 km	H2_100_DR2	Minke	EOWA	61.6	69.6
10 x 10 km	H2_50_DR1	Sei	EOWA	26.8	35.3
10 x 10 km	H2_50_DR2	Sei	EOWA	27.0	33.0
10 x 10 km	H2_100_DR1	Sei	EOWA	18.0	20.7
10 x 10 km	H2_100_DR2	Sei	EOWA	25.0	29.4
10 x 10 km	H2_50_DR1	Fin	EOWA	7.8	16.8
10 x 10 km	H2_50_DR2	Fin	EOWA	4.3	14.2
10 x 10 km	H2_100_DR1	Fin	EOWA	5.2	9.3
10 x 10 km	H2_100_DR2	Fin	EOWA	3.0	13.9
10 x 10 km	H2_50_DR1	NARW	EOWA	25.8	36.1
10 x 10 km	H2_50_DR2	NARW	EOWA	27.3	31.1

Design	Hypothesis	Species	Wind Farm	Large Monitoring Area & Both PAM Grids	Large Monitoring Area & Modified Small PAM Grid Only
10 x 10 km	H2_100_DR1	NARW	EOWA	29.5	33.6
10 x 10 km	H2_100_DR2	NARW	EOWA	33.9	41.6
10 x 10 km	H3_DR1	Minke	MAWA	73.0	80.2
10 x 10 km	H3_DR2	Minke	MAWA	76.4	81.9
10 x 10 km	H3_DR1	Sei	MAWA	41.4	50.5
10 x 10 km	H3_DR2	Sei	MAWA	42.8	49.9
10 x 10 km	H3_DR1	Fin	MAWA	0.7	2.8
10 x 10 km	H3_DR2	Fin	MAWA	5.5	13.7
10 x 10 km	H3_DR1	NARW	MAWA	58.2	63.0
10 x 10 km	H3_DR2	NARW	MAWA	55.9	64.0
10 x 10 km	H3_DR1	Minke	VYWA	***81.1**	***84.4**
10 x 10 km	H3_DR2	Minke	VYWA	***80.6**	***80.3**
10 x 10 km	H3_DR1	Sei	VYWA	24.8	31.3
10 x 10 km	H3_DR2	Sei	VYWA	23.1	33.1
10 x 10 km	H3_DR1	Fin	VYWA	6.2	9.1
10 x 10 km	H3_DR2	Fin	VYWA	6.7	13.4
10 x 10 km	H3_DR1	NARW	VYWA	33.8	37.0
10 x 10 km	H3_DR2	NARW	VYWA	27.6	35.4
10 x 10 km	H3_DR1	Minke	EOWA	75.2	78.3
10 x 10 km	H3_DR2	Minke	EOWA	75.0	84.9
10 x 10 km	H3_DR1	Sei	EOWA	56.8	67.6
10 x 10 km	H3_DR2	Sei	EOWA	65.7	71.5
10 x 10 km	H3_DR1	Fin	EOWA	6.7	9.0
10 x 10 km	H3_DR2	Fin	EOWA	8.3	12.8
10 x 10 km	H3_DR1	NARW	EOWA	55.1	63.4
10 x 10 km	H3_DR2	NARW	EOWA	51.0	53.7
10 x 10 km	H4_DR1	Minke	MAWA	79.6	87.0
10 x 10 km	H4_DR2	Minke	MAWA	79.7	88.2
10 x 10 km	H4_DR1	Sei	MAWA	41.9	49.9
10 x 10 km	H4_DR2	Sei	MAWA	42.5	45.3
10 x 10 km	H4_DR1	Fin	MAWA	1.7	9.6
10 x 10 km	H4_DR2	Fin	MAWA	0.7	9.3
10 x 10 km	H4_DR1	NARW	MAWA	52.1	59.6
10 x 10 km	H4_DR2	NARW	MAWA	66.4	71.6
10 x 10 km	H4_DR1	Minke	VYWA	***82.0**	***84.8**
10 x 10 km	H4_DR2	Minke	VYWA	***80.4**	***87.0**
10 x 10 km	H4_DR1	Sei	VYWA	25.3	34.9
10 x 10 km	H4_DR2	Sei	VYWA	23.0	31.0
10 x 10 km	H4_DR1	Fin	VYWA	0.7	5.2
10 x 10 km	H4_DR2	Fin	VYWA	2.3	6.6

Design	Hypothesis	Species	Wind Farm	Large Monitoring Area & Both PAM Grids	Large Monitoring Area & Modified Small PAM Grid Only
10 x 10 km	H4_DR1	NARW	VYWA	33.8	38.7
10 x 10 km	H4_DR2	NARW	VYWA	27.2	31.1
10 x 10 km	H4_DR1	Minke	EOWA	54.7	59.4
10 x 10 km	H4_DR2	Minke	EOWA	59.3	64.2
10 x 10 km	H4_DR1	Sei	EOWA	36.5	42.6
10 x 10 km	H4_DR2	Sei	EOWA	36.5	45.8
10 x 10 km	H4_DR1	Fin	EOWA	2.6	5.1
10 x 10 km	H4_DR2	Fin	EOWA	3.4	5.9
10 x 10 km	H4_DR1	NARW	EOWA	64.7	69.9
10 x 10 km	H4_DR2	NARW	EOWA	61.6	70.9
10 x 10 km	H5_DR1	Minke	MAWA	44.8	55.3
10 x 10 km	H5_DR2	Minke	MAWA	47.2	52.7
10 x 10 km	H5_DR1	Sei	MAWA	41.6	46.2
10 x 10 km	H5_DR2	Sei	MAWA	42.2	45.3
10 x 10 km	H5_DR1	Fin	MAWA	4.9	13.0
10 x 10 km	H5_DR2	Fin	MAWA	3.8	14.4
10 x 10 km	H5_DR1	NARW	MAWA	68.2	71.8
10 x 10 km	H5_DR2	NARW	MAWA	66.5	72.8
10 x 10 km	H5_DR1	Minke	VYWA	***87.5**	***89.6**
10 x 10 km	H5_DR2	Minke	VYWA	***88.8**	***84.9**
10 x 10 km	H5_DR1	Sei	VYWA	25.5	33.3
10 x 10 km	H5_DR2	Sei	VYWA	29.2	31.7
10 x 10 km	H5_DR1	Fin	VYWA	4.9	14.3
10 x 10 km	H5_DR2	Fin	VYWA	5.0	14.6
10 x 10 km	H5_DR1	NARW	VYWA	33.9	44.6
10 x 10 km	H5_DR2	NARW	VYWA	27.4	36.8
10 x 10 km	H5_DR1	Minke	EOWA	55.4	63.5
10 x 10 km	H5_DR2	Minke	EOWA	58.2	62.9
10 x 10 km	H5_DR1	Sei	EOWA	56.4	58.4
10 x 10 km	H5_DR2	Sei	EOWA	66.8	72.2
10 x 10 km	H5_DR1	Fin	EOWA	3.5	6.6
10 x 10 km	H5_DR2	Fin	EOWA	4.2	10.4
10 x 10 km	H5_DR1	NARW	EOWA	54.9	58.6
10 x 10 km	H5_DR2	NARW	EOWA	50.8	56.9
10 x 10 km	H6	Minke	VYWA	59.2	61.2
10 x 10 km	H6	Sei	VYWA	56.0	59.2
10 x 10 km	H6	Fin	VYWA	5.4	9.7
10 x 10 km	H6	NARW	VYWA	50.2	54.7
10 x 10 km	H6	Minke	EOWA	59.6	66.8
10 x 10 km	H6	Sei	EOWA	60.1	62.6

Design	Hypothesis	Species	Wind Farm	Large Monitoring Area & Both PAM Grids	Large Monitoring Area & Modified Small PAM Grid Only
10 x 10 km	H6	Fin	EOWA	5.8	9.5
10 x 10 km	H6	NARW	EOWA	63.1	71.0
10 x 10 km	H7	Minke	MAWA	54.3	63.9
10 x 10 km	H7	Sei	MAWA	46.0	49.0
10 x 10 km	H7	Fin	MAWA	6.0	13.7
10 x 10 km	H7	NARW	MAWA	68.9	78.4
10 x 10 km	H7	Minke	VYWA	64.5	68.0
10 x 10 km	H7	Sei	VYWA	70.3	74.4
10 x 10 km	H7	Fin	VYWA	1.1	4.7
10 x 10 km	H7	NARW	VYWA	52.9	57.8
10 x 10 km	H7	Minke	EOWA	70.0	75.2
10 x 10 km	H7	Sei	EOWA	70.7	78.9
10 x 10 km	H7	Fin	EOWA	7.1	16.2
10 x 10 km	H7	NARW	EOWA	55.5	60.0
10 x 10 km	H8	Minke	MAWA	***99.1**	***99.2**
10 x 10 km	H8	Sei	MAWA	***93.3**	***98.3**
10 x 10 km	H8	Fin	MAWA	***94.3**	***94.1**
10 x 10 km	H8	NARW	MAWA	***89.2**	***96.5**
10 x 10 km	H8	Minke	VYWA	***94.3**	***98.1**
10 x 10 km	H8	Sei	VYWA	***93.9**	***94.1**
10 x 10 km	H8	Fin	VYWA	***97.7**	***98.2**
10 x 10 km	H8	NARW	VYWA	***98.8**	***98.3**
10 x 10 km	H8	Minke	EOWA	***94.7**	***96.2**
10 x 10 km	H8	Sei	EOWA	***95.0**	***96.5**
10 x 10 km	H8	Fin	EOWA	***97.9**	***98.2**
10 x 10 km	H8	NARW	EOWA	***95.8**	***98.2**

E.1.2 T-design Grid

Table E-2. Statistical power (%) to detect change in cue rates under eight hypotheses (see Table 2) for each studied species, wind farm area and combination of monitoring size and PAM grid assuming modified small PAM grid using T-design grid

For hypotheses related to construction of the windfarms (H2–H5), power to detect change was calculated for two dose-response functions: DR1 and DR2. Scenarios with power >= 80% are marked in bold and with an asterisk.

Design	Hypothesis	Species	Wind Farm	Large Monitoring Area & Both PAM Grids	Large Monitoring Area & Modified Small PAM Grid Only
T-design	H1_co	Minke	MAWA	6.6	6.5
T-design	H1_op	Minke	MAWA	7.2	6.5
T-design	H1_co	Sei	MAWA	6.4	6.3
T-design	H1_op	Sei	MAWA	6.5	6.4
T-design	H1_co	Fin	MAWA	7.8	7.4
T-design	H1_op	Fin	MAWA	11.3	13.0
T-design	H1_co	NARW	MAWA	8.8	11.0
T-design	H1_op	NARW	MAWA	9.9	11.2
T-design	H1_co	Minke	VYWA	11.6	13.2
T-design	H1_op	Minke	VYWA	11.8	12.3
T-design	H1_co	Sei	VYWA	6.1	7.7
T-design	H1_op	Sei	VYWA	8.0	9.1
T-design	H1_co	Fin	VYWA	8.7	9.6
T-design	H1_op	Fin	VYWA	8.9	9.0
T-design	H1_co	NARW	VYWA	5.7	5.8
T-design	H1_op	NARW	VYWA	7.9	9.1
T-design	H1_co	Minke	EOWA	8.0	8.4
T-design	H1_op	Minke	EOWA	9.3	9.3
T-design	H1_co	Sei	EOWA	10.9	11.0
T-design	H1_op	Sei	EOWA	10.6	10.5
T-design	H1_co	Fin	EOWA	9.5	10.2
T-design	H1_op	Fin	EOWA	10.4	10.4
T-design	H1_co	NARW	EOWA	5.2	6.8
T-design	H1_op	NARW	EOWA	6.0	6.5
T-design	H2_50_DR1	Minke	MAWA	72.7	75.9
T-design	H2_50_DR2	Minke	MAWA	72.7	75.4
T-design	H2_100_DR1	Minke	MAWA	78.1	***82.1**
T-design	H2_100_DR2	Minke	MAWA	79.6	***82.5**
T-design	H2_50_DR1	Sei	MAWA	40.5	46.0
T-design	H2_50_DR2	Sei	MAWA	39.1	42.6
T-design	H2_100_DR1	Sei	MAWA	39.9	46.3
T-design	H2_100_DR2	Sei	MAWA	38.9	45.6
T-design	H2_50_DR1	Fin	MAWA	5.2	12.4
T-design	H2_50_DR2	Fin	MAWA	6.4	13.2

Design	Hypothesis	Species	Wind Farm	Large Monitoring Area & Both PAM Grids	Large Monitoring Area & Modified Small PAM Grid Only
T-design	H2_100_DR1	Fin	MAWA	9.0	12.3
T-design	H2_100_DR2	Fin	MAWA	10.3	14.4
T-design	H2_50_DR1	NARW	MAWA	66.2	68.6
T-design	H2_50_DR2	NARW	MAWA	66.7	74.5
T-design	H2_100_DR1	NARW	MAWA	66.4	72.8
T-design	H2_100_DR2	NARW	MAWA	66.5	73.3
T-design	H2_50_DR1	Minke	VYWA	68.4	72.2
T-design	H2_50_DR2	Minke	VYWA	69.2	75.6
T-design	H2_100_DR1	Minke	VYWA	***81.4**	***85.3**
T-design	H2_100_DR2	Minke	VYWA	***80.6**	***83.0**
T-design	H2_50_DR1	Sei	VYWA	70.5	74.6
T-design	H2_50_DR2	Sei	VYWA	71.1	76.2
T-design	H2_100_DR1	Sei	VYWA	76.2	78.2
T-design	H2_100_DR2	Sei	VYWA	76.7	77.1
T-design	H2_50_DR1	Fin	VYWA	16.2	19.6
T-design	H2_50_DR2	Fin	VYWA	10.5	12.5
T-design	H2_100_DR1	Fin	VYWA	13.0	14.0
T-design	H2_100_DR2	Fin	VYWA	11.1	17.2
T-design	H2_50_DR1	NARW	VYWA	65.1	72.6
T-design	H2_50_DR2	NARW	VYWA	60.7	64.2
T-design	H2_100_DR1	NARW	VYWA	74.0	79.5
T-design	H2_100_DR2	NARW	VYWA	75.7	81.7
T-design	H2_50_DR1	Minke	EOWA	67.4	74.5
T-design	H2_50_DR2	Minke	EOWA	67.3	74.5
T-design	H2_100_DR1	Minke	EOWA	***82.8**	***85.4**
T-design	H2_100_DR2	Minke	EOWA	***85.4**	***91.0**
T-design	H2_50_DR1	Sei	EOWA	66.5	68.2
T-design	H2_50_DR2	Sei	EOWA	64.7	70.1
T-design	H2_100_DR1	Sei	EOWA	64.8	67.3
T-design	H2_100_DR2	Sei	EOWA	64.3	66.8
T-design	H2_50_DR1	Fin	EOWA	14.2	19.4
T-design	H2_50_DR2	Fin	EOWA	12.3	17.8
T-design	H2_100_DR1	Fin	EOWA	16.4	19.7
T-design	H2_100_DR2	Fin	EOWA	18.6	25.4
T-design	H2_50_DR1	NARW	EOWA	62.8	67.7
T-design	H2_50_DR2	NARW	EOWA	60.5	67.2
T-design	H2_100_DR1	NARW	EOWA	60.7	64.0
T-design	H2_100_DR2	NARW	EOWA	60.2	61.8
T-design	H3_DR1	Minke	MAWA	***81.6**	***87.4**
T-design	H3_DR2	Minke	MAWA	***85.4**	***88.6**
T-design	H3_DR1	Sei	MAWA	39.7	46.1

Design	Hypothesis	Species	Wind Farm	Large Monitoring Area & Both PAM Grids	Large Monitoring Area & Modified Small PAM Grid Only
T-design	H3_DR2	Sei	MAWA	39.9	41.9
T-design	H3_DR1	Fin	MAWA	12.1	14.6
T-design	H3_DR2	Fin	MAWA	14.9	16.2
T-design	H3_DR1	NARW	MAWA	66.4	72.1
T-design	H3_DR2	NARW	MAWA	66.4	68.9
T-design	H3_DR1	Minke	VYWA	***82.2**	***84.9**
T-design	H3_DR2	Minke	VYWA	***87.2**	***92.2**
T-design	H3_DR1	Sei	VYWA	***82.9**	***87.0**
T-design	H3_DR2	Sei	VYWA	***85.6**	***88.3**
T-design	H3_DR1	Fin	VYWA	15.5	17.9
T-design	H3_DR2	Fin	VYWA	17.4	19.1
T-design	H3_DR1	NARW	VYWA	***84.9**	***91.1**
T-design	H3_DR2	NARW	VYWA	***84.1**	***86.1**
T-design	H3_DR1	Minke	EOWA	***84.3**	***91.3**
T-design	H3_DR2	Minke	EOWA	***88.8**	***94.5**
T-design	H3_DR1	Sei	EOWA	79.9	***85.2**
T-design	H3_DR2	Sei	EOWA	79.8	***82.1**
T-design	H3_DR1	Fin	EOWA	13.3	15.3
T-design	H3_DR2	Fin	EOWA	17.8	22.6
T-design	H3_DR1	NARW	EOWA	62.8	65.9
T-design	H3_DR2	NARW	EOWA	60.2	62.5
T-design	H4_DR1	Minke	MAWA	***84.6**	***87.4**
T-design	H4_DR2	Minke	MAWA	***88.8**	***90.7**
T-design	H4_DR1	Sei	MAWA	40.1	47.6
T-design	H4_DR2	Sei	MAWA	39.1	41.3
T-design	H4_DR1	Fin	MAWA	13.4	19.2
T-design	H4_DR2	Fin	MAWA	16.1	18.1
T-design	H4_DR1	NARW	MAWA	66.8	73.1
T-design	H4_DR2	NARW	MAWA	67.0	70.5
T-design	H4_DR1	Minke	VYWA	***84.4**	***89.8**
T-design	H4_DR2	Minke	VYWA	***89.7**	***94.1**
T-design	H4_DR1	Sei	VYWA	***81.4**	***81.4**
T-design	H4_DR2	Sei	VYWA	***81.0**	***87.3**
T-design	H4_DR1	Fin	VYWA	12.7	15.8
T-design	H4_DR2	Fin	VYWA	15.2	17.3
T-design	H4_DR1	NARW	VYWA	***80.3**	***81.8**
T-design	H4_DR2	NARW	VYWA	***82.4**	***88.4**
T-design	H4_DR1	Minke	EOWA	***82.2**	***85.7**
T-design	H4_DR2	Minke	EOWA	***85.4**	***88.7**
T-design	H4_DR1	Sei	EOWA	***81.9**	***87.2**
T-design	H4_DR2	Sei	EOWA	***81.9**	***84.3**

Design	Hypothesis	Species	Wind Farm	Large Monitoring Area & Both PAM Grids	Large Monitoring Area & Modified Small PAM Grid Only
T-design	H4_DR1	Fin	EOWA	14.2	20.5
T-design	H4_DR2	Fin	EOWA	15.2	22.1
T-design	H4_DR1	NARW	EOWA	62.2	67.9
T-design	H4_DR2	NARW	EOWA	60.2	64.4
T-design	H5_DR1	Minke	MAWA	***85.4**	***89.8**
T-design	H5_DR2	Minke	MAWA	***89.5**	***95.5**
T-design	H5_DR1	Sei	MAWA	40.6	47.5
T-design	H5_DR2	Sei	MAWA	39.3	43.8
T-design	H5_DR1	Fin	MAWA	10.6	17.3
T-design	H5_DR2	Fin	MAWA	18.0	22.9
T-design	H5_DR1	NARW	MAWA	66.5	71.3
T-design	H5_DR2	NARW	MAWA	67.5	70.4
T-design	H5_DR1	Minke	VYWA	***83.8**	***88.2**
T-design	H5_DR2	Minke	VYWA	***82.9**	***86.8**
T-design	H5_DR1	Sei	VYWA	***81.9**	***88.0**
T-design	H5_DR2	Sei	VYWA	***83.5**	***90.9**
T-design	H5_DR1	Fin	VYWA	13.7	19.5
T-design	H5_DR2	Fin	VYWA	15.8	21.0
T-design	H5_DR1	NARW	VYWA	***85.9**	***92.8**
T-design	H5_DR2	NARW	VYWA	***85.8**	***92.8**
T-design	H5_DR1	Minke	EOWA	***83.4**	***85.9**
T-design	H5_DR2	Minke	EOWA	***88.2**	***93.7**
T-design	H5_DR1	Sei	EOWA	65.5	70.8
T-design	H5_DR2	Sei	EOWA	65.0	71.4
T-design	H5_DR1	Fin	EOWA	13.3	18.4
T-design	H5_DR2	Fin	EOWA	17.6	22.4
T-design	H5_DR1	NARW	EOWA	62.0	69.0
T-design	H5_DR2	NARW	EOWA	59.5	65.5
T-design	H6	Minke	VYWA	***81.2**	***85.0**
T-design	H6	Sei	VYWA	74.9	78.8
T-design	H6	Fin	VYWA	9.3	14.8
T-design	H6	NARW	VYWA	75.8	81.3
T-design	H6	Minke	EOWA	***81.5**	***86.4**
T-design	H6	Sei	EOWA	74.4	78.5
T-design	H6	Fin	EOWA	9.2	14.6
T-design	H6	NARW	EOWA	78.5	75.2
T-design	H7	Minke	MAWA	79.6	***86.7**
T-design	H7	Sei	MAWA	47.3	53.1
T-design	H7	Fin	MAWA	10.5	14.2
T-design	H7	NARW	MAWA	73.1	70.6
T-design	H7	Minke	VYWA	76.7	***81.7**

Design	Hypothesis	Species	Wind Farm	Large Monitoring Area & Both PAM Grids	Large Monitoring Area & Modified Small PAM Grid Only
T-design	H7	Sei	VYWA	71.0	72.8
T-design	H7	Fin	VYWA	11.2	16.3
T-design	H7	NARW	VYWA	78.1	73.1
T-design	H7	Minke	EOWA	78.7	***83.6**
T-design	H7	Sei	EOWA	75.5	78.2
T-design	H7	Fin	EOWA	10.4	12.7
T-design	H7	NARW	EOWA	69.1	72.9
T-design	H8	Minke	MAWA	***98.9**	***94.4**
T-design	H8	Sei	MAWA	***88.1**	***94.2**
T-design	H8	Fin	MAWA	***91.2**	***94.8**
T-design	H8	NARW	MAWA	***74.4**	***79.5**
T-design	H8	Minke	VYWA	***88.6**	***92.1**
T-design	H8	Sei	VYWA	***96.2**	***99.2**
T-design	H8	Fin	VYWA	***94.9**	***97.1**
T-design	H8	NARW	VYWA	***94.4**	***91.2**
T-design	H8	Minke	EOWA	***96.3**	***90.3**
T-design	H8	Sei	EOWA	***69.8**	***73.8**
T-design	H8	Fin	EOWA	***95.4**	***99.4**
T-design	H8	NARW	EOWA	***84.3**	***85.8**

E.2 Regional Analysis

Table E-3. Statistical power (%) to detect change in cue rates under two hypotheses (see Table 2) related to operation of wind farms for each studied species, combination of monitoring size and PAM grid and two lengths of monitoring: 1 and 5 years
Scenarios with power >= 80% are marked in bold and with an asterisk.

Design	Hypothesis	Species	Wind Farm	Both PAM Grids	Modified Small PAM Grid Only
T-design	H1_op_1y	Minke	Regional	7.6	8.4
T-design	H1_op_1y	Sei	Regional	11.3	12.5
T-design	H1_op_1y	Fin	Regional	11.2	15.11
T-design	H1_op_1y	NARW	Regional	8.8	9.2
T-design	H6_1y	Minke	Regional	***89.2**	***97.2**
T-design	H6_1y	Sei	Regional	***87.2**	***97.0**
T-design	H6_1y	Fin	Regional	43	52.3
T-design	H6_5y	Fin	Regional	55.7	16.2
T-design	H6_1y	NARW	Regional	***88.6**	***89.3**
T-design	H7_1y	Minke	Regional	***88.2**	***88.7**
T-design	H7_1y	Sei	Regional	***86.1**	***85.1**
T-design	H7_1y	Fin	Regional	44.1	48.6
T-design	H7_5y	Fin	Regional	48.7	52.4
T-design	H7_1y	NARW	Regional	***87.8**	***89.2**

U.S. Department of the Interior (DOI)

DOI protects and manages the Nation's natural resources and cultural heritage; provides scientific and other information about those resources; and honors the Nation’s trust responsibilities or special commitments to American Indians, Alaska Natives, and affiliated island communities.

Bureau of Ocean Energy Management (BOEM)

BOEM’s mission is to manage development of U.S. Outer Continental Shelf energy and mineral resources in an environmentally and economically responsible way.

BOEM Environmental Studies Program

The mission of the Environmental Studies Program is to provide the information needed to predict, assess, and manage impacts from offshore energy and marine mineral exploration, development, and production activities on human, marine, and coastal environments. The proposal, selection, research, review, collaboration, production, and dissemination of each of BOEM’s Environmental Studies follows the DOI Code of Scientific and Scholarly Conduct, in support of a culture of scientific and professional integrity, as set out in the DOI Departmental Manual (305 DM 3).