Part IV. Recommendations for Boat-based and Aeria Surveys

Note: This is an excerpt from "*Guidance for Pre- and Post-Construction Monitoring to Detect Changes in Marine Bird Distributions and Habitat Use Related to Offshore Wind Development".* The full guidance document is available at www.nyetwg.com/avian-displacement-guidance

Developed by the Avian Displacement Guidance Committee of the Environmental Technical Working Group, with support from the Biodiversity Research Institute

Citation: Avian Displacement Guidance Committee. 2024. Guidance for Pre- and Post-Construction Monitoring to Detect Changes in Marine Bird Distributions and Habitat Use Related to Offshore Wind Development. Report to the Offshore Wind Environmental Technical Working Group. 100 pp. Available at www.nyetwg.com/avian-displacement-guidance.

Photo credit: Nicholas Doherty

Part IV. Recommendations for Boat-based and Aerial Surveys

As indicated in Section 6, observational surveys are best suited to answer the following types of effects questions:

- Are changes in habitat use (e.g., displacement/attraction) of marine birds occurring, and if so, what is the magnitude and distance from the OSW facility at which it occurs?
- Does the occurrence, magnitude, and distance of changes in habitat use vary temporally (e.g., does habituation occur)?
- Are there changes in foraging or roosting activity of marine birds in relation to the OSW facility?

In contrast, observational surveys are not well suited to answer effects questions related to individual movements. Surveys to detect effects from OSW facilities are typically focused at the spatial scale of a single OSW project, with a "buffer area" around the project footprint (except in cases where effects of neighboring wind facilities are studied with a single survey effort). Such surveys are typically conducted both prior to and following OSW construction and must be designed to have adequate statistical power to detect responses. The recommendations below build from existing BOEM avian survey guidelines (BOEM 2020; references where relevant) but have been expanded upon to focus specifically on surveys to answer the above types of research questions. The recommendations in this section are intended to be widely applicable across effect studies conducted at the site-level using observational surveys. However, recognizing that project-level considerations will play a role in study design, deviations from these recommendations should be carefully considered and justified based on statistically and scientifically robust analysis in consultation with federal agencies.

9.0 Connection Between Site Assessment Surveys and Pre-Construction Surveys to Detect Effects

Before OSW facilities are built, observational surveys are conducted for several purposes, including (1) to inform the siting of wind energy areas, (2) for site characterization to inform permitting processes and monitoring plans, and (3) to pair with post-construction surveys to detect effects of OSW development ("effects surveys"; above). Government-funded offshore surveys to inform siting are often regional in spatial scale, and thus may lack the fine-scale spatial resolution to adequately detect effects at the project scale. Both site characterization surveys and effects surveys occur at a finer spatial scale, focused in and around an OSW facility. **The primary focus of this effort is to provide recommendations for conducting surveys to detect effects from OSW development on marine birds, including surveys conducted both pre- and post-construction**. However, it is important to consider the degree to which surveys conducted at an OSW project site prior to construction may inform site characterization efforts as well as the assessment of OSW effects.

The primary question that site characterization surveys should be designed to answer is: What are exposure levels for different species/taxa at the project site and how does exposure vary spatiotemporally? With this exposure information, the following questions can then be explored to inform risk assessments and project design: (1) Do existing vulnerability data suggest any of these species could be at high risk from OSW development given considerations of population status and sensitivity to effects (see Section 5 for definitions)? And, if so, (2) Where should avoidance and minimization efforts be focused, based on the greatest potential effects to different species across the annual cycle?

In some locations, existing survey data for a site can be used in place of new site characterization surveys (see Avian Displacement Guidance Committee 2023) for additional guidance on when new site characterization surveys are needed). The existing BOEM avian survey guidelines (2020) are explicitly focused on recommendations for conducting site characterization surveys, and the methods recommended therein are thus inadequate for effects studies focused on understanding changes in distribution and abundance patterns due to the presence of OSW facilities[5]. This issue is further discussed by this Committee in the site characterization recommendations (Avian Displacement Guidance Committee 2023).

Given that both site characterization surveys and pre-construction effects surveys occur prior to construction of a wind facility, it is theoretically possible that the two types of surveys could be combined into a single survey effort prior to OSW construction. However, pre-construction surveys have stricter study design limitations than site characterization surveys, to ensure they have sufficient power to detect change (see Section 9), and post-construction surveys should be initiated within five years of the completion of pre-construction surveys, to minimize the chance of non-OSW variables (e.g., decadal shifts in marine ecosystems due to climate change) influencing distributions and abundance in ways that could be conflated with OSW effects (Kinlan et al. 2012). It is unlikely that post-construction surveys could be initiated within five years of the completion of site characterization surveys (which should be conducted prior to development of a Construction and Operations Plan), particularly given the length of current permitting and construction timelines. As such, in cases where there are insufficient preexisting survey data for a proposed OSW location for site characterization purposes, and additional data are needed to characterize the site, we recommend that separate site assessment and pre-construction surveys to detect effects are conducted (Avian Displacement Guidance Committee 2023), given differences in the objectives of each survey as well as challenges associated with timing under current permitting timelines. Site assessment data (either pre-existing or collected during site characterization surveys for the project) on species presence and abundance at the site should be used to inform the choice of focal taxa and the design of effects surveys.

10.0 Survey Design and Methodology Recommendations

Surveys can be used for many different types of research questions, but the recommendations below are focused on effectively quantifying effects of displacement and attraction from OSW energy development (see Section 6). If the intent is for observational surveys to serve multiple objectives, careful consideration is needed to ensure that all objectives are met effectively. Some of the below recommendations apply broadly to observational surveys. Others may be specific to boat-based or digital aerial surveys or may be specific to certain focal taxa, as indicated.

10.1 Define Clear Study Goals

Given that observational surveys can be used for multiple purposes, it is important to define clear study goals and research questions (Section 7.1). In addition to defining research questions (Section 4), it is also important to define focal species (Section 5). While one of the strengths of observational surveys is the ability to simultaneously collect data across a range of taxa, key aspects of study design and methodology (e.g., choice of buffer size) rely on the choice of focal species. As such, existing data from the area (either from previous site characterization surveys, or other data sources such as tracking data and incidental observations), should be used to define the full list of species likely to be found in the area, and then

[5] See Atlantic Marine Bird Cooperative Marine Spatial Planning Workgroup's 2021 recommendations to BOEM on these avian survey guidelines.

categorized into “high”, “medium”, and “low” priority species (Section 5). The goal should then be to design surveys to adequately answer research questions for the high priority species with careful consideration of the amount of existing data available to inform the design and the level of likely exposure and sensitivity to effects of these focal taxa, as these considerations will be key in refining study methods.

10.2 Use of Gradient Study Design

It is recommended that observational surveys to detect effects utilize Before-After-Gradient (BAG) study designs (Cook et al. 2018). Effect studies using observational surveys in Europe have used various study designs, including Before-After-Control-Impact (BACI), Stratified BACI, After-Control-Impact (ACI), and Before-After-Gradient (BAG) designs (see Appendix C for summary). BACI designs sample a treatment site (e.g., the OSW facility) and a control site away from the facility before and after “intervention” (e.g., when the OSW project is built) and statistically compare across locations and time periods (Green 1979). Stratified BACI and ACI are variations on this design whereby the impact area is stratified into concentric areas for comparison with the control, or a comparison only occurs after impact, respectively. While study designs involving a control are commonly used in the study of effects from OSW development (Methratta 2021), there are challenges associated with these designs whereby it is difficult, if not impossible, to choose adequate control sites (Vanermen et al. 2015b). In contrast, a BAG design includes data collection at relative distances from the OSW facility both pre- and post-construction (Ellis and Schneider 1997). Combining the before-after sampling design with distance-based methods is a powerful approach that accounts for both spatial and temporal variation in response (Methratta 2021). While often more powerful than BACI-type designs, the spatial and temporal scale of BAG designs must still be carefully selected (Section 10.3).

10.3 Assessment of Spatial and Temporal Coverage

Before-After-Gradient survey designs require that surveys be conducted in the entirety of the wind facility, plus a buffer area of some distance outside of the project footprint. Appropriate survey design must consider the necessary size of this buffer zone and the proportion of the “survey area” (the wind facility plus buffer area) that is covered by survey effort, as well as the ratio of the “effect area” (e.g., the wind facility footprint) to the full survey area. All three of these aspects interact to affect statistical power and therefore should be carefully considered. In addition to spatial coverage, the temporal scale of surveys, both in terms of the length of the overall data collection period pre- and post-construction, and frequency of surveys throughout the period, require careful consideration. Below, we provide general recommendations on aspects of spatial and temporal coverage based on existing knowledge, but **strongly recommend that existing data are used in site-specific power analyses to inform the choice of spatial and temporal coverage of surveys based on the focal taxa at each site**. There are various tools, such as the R package MRSeaPower (Scott-Hayward et al. 2014) that can aid in this type of analysis.

It is important to note that regardless of choice of spatial and temporal coverage, zero inflation (e.g., as dictated by species abundance and distribution) and effect size (e.g., the magnitude of change in these distributions due to the presence of the OSW facility) play important roles in determining a study’s statistical power to detect an effect if the effect exists. Surveys of species that are uncommon or lower in abundance at a site will have large numbers of zeroes in the data, which has a strong negative effect on statistical power (Vanermen et al. 2015b; LaPeña et al. 2011). **As such, we encourage the choice of focal species with relatively high exposure** (Section 5). Similarly, small changes in abundance (e.g., 10%) are difficult to detect even with high intensity survey effort (Donovan & Caneco 2020; Leirness & Kinlan 2018), so **selection of focal species with expected greater magnitude of response will increase the chance of detecting that response if it occurs** (Section 5). **For species where potential effect size is unknown, effect size should be estimated conservatively to ensure the study is designed with a higher chance of detecting effects, should they occur.**

10.3.1 Buffer Size and Ratio of Effect: Overall Area

While we can draw from European studies regarding potential species-specific displacement and attraction distances, there have been relatively few well-designed studies to date. There is a high level of variation in effects among species and studies in the existing literature, and the degree to which results are applicable to U.S. populations and ecosystems is unknown. However, for species where there is evidence of displacement in Europe (e.g., auks, loons, gannets, sea ducks), populations were displaced anywhere between 500 m and 16.5 km (see Appendix C and Lamb et al. 2024).

- **We recommend a buffer zone of 4–20 km be surveyed around the OSW project footprint with a consistent buffer distance in all directions. The choice of buffer size should be based on the suite of species present in the area, selection of specific focal species (Section 5), and their known or suspected sensitivity to displacement (based on** best available knowledge from the literature). For example, if primarily focused on species such as auks, a 4–6 km buffer would likely suffice, whereas if species with high displacement distances (e.g., loons, sea ducks) are focal species of the survey, a larger buffer (10+ km) is needed (NatureScot 2023). See Appendix C and Lamb et al. 2024 for current literature on displacement distances. In cases where sensitivity is unknown, a precautionary approach (e.g., larger buffer) should be used. If data are also intended to contribute to regional understanding of distributions, not capturing the areas where birds are displaced to introduces additional bias into overall density estimates.
- The choice of buffer size should be informed by 1) power analyses of existing data, 2) abundance of focal species at the site, as an increase in species abundance helps to reduce skewness of the distribution and in turn increases statistical power (LaPeña et al. 2011), and 3) ratio of effect area to overall area surveyed. A reduced ratio (e.g., increased area surveyed outside of the effect area), with density of observations held constant, decreases variance and reduces spatial autocorrelation, thereby increasing statistical power (LaPeña et al. 2010). As a rule of thumb, the choice of survey area should be informed by the spatial extent at which changes are predicted to occur, such that the total survey area includes the wind farm footprint, as well as a buffer zone that incorporates the predicted effect distance for focal taxa plus 10%.
- For adjacent lease areas, we encourage coordinated survey efforts, to the degree feasible given differences in construction timelines, to maximize efficiency and treat the area as a continuous habitat for marine birds. Such coordination should be supported by regulators and by regional groups, such as the Regional Wildlife Science Collaborative.

As data from the U.S. Atlantic become available from initial offshore wind project studies, the recommended buffer size should be revisited to confirm that studies to detect displacement effects are designed to have adequate statistical power and are incorporating updated information on effect distances for species in the region.

10.3.2 Spatial Coverage

The percentage of the total survey area that is covered by the survey is calculated as sampled area/total survey area) *100. The **effective strip width** is calculated differently for boat-based and digital aerial surveys. For data collected by digital aerial surveys, the assumption is that image reviewers detect every target within the surveyed area and estimate seabird relative abundance by dividing the number of individuals sighted by the area of ocean surface surveyed (Hyrenbach et al. 2007). For strip transects, the effective strip width is a single value representing the sum of the digital aerial survey cameras' width of coverage at sea level, while accounting for the actual (rather than planned) altitude of the aircraft. For boat-based surveys, **line transects** utilize distance sampling methods to handle imperfections of the observation process such as decaying detectability with increasing distance from the observer (Buckland et al. 2001), the overall detectability at zero distance (Buckland et al. 2001) and the effect of environmental conditions on detectability (Marques and Buckland 2003). The effective strip width with line transect methodology varies by species, as detectability of those species varies with distance from the observer. Distance (u) from the transect line where the number of animals detected beyond u are equal to the number that are missed within u. This value estimates the effective area of the survey and can be used to correct density estimates or estimates of survey coverage (Buckland et al. 2001). From both a logistical feasibility and statistical analysis standpoint there may be tradeoffs between buffer size and percent spatial coverage as these are interacting spatial factors in study design.

- Generally, **we recommend at least 20% spatial coverage of the survey area for surveys to detect effects in order to achieve adequate statistical power**, as is common in European OSW studies (Harker et al. 2022, HiDef 2021), has been achieved in some U.S. Atlantic regional studies (Mid Atlantic Baseline Studies; Williams et al. 2015). However, power analyses with existing data should be used to inform this choice, taking into consideration both the abundance and spatial distribution of focal species. In general, increasing spatial coverage leads to an increase in power due to improved ability to estimate means and reduced variance (e.g., reducing transect spacing from 3 km to 1 km increased power from 0.55 to 0.84 in La Pena et al. 2011). While there may be instances where a study can achieve adequate statistical power to detect change with 10% or less spatial coverage, this is likely only true for abundant and consistently distributed species with high effect size (>20% change; Donovan & Caneco 2020). If focal species are rare (e.g., low exposure, high population sensitivity) or highly aggregated in space, additional spatial coverage beyond 20% may be required to achieve adequate statistical power.
- **Percent spatial coverage for boat-based line transects should be calculated based on effective strip width for focal species**. If the study is focused on detecting effects across multiple species, the minimum effective strip width across focal taxa can be used to calculate percent spatial coverage based on previous detection probability curves (ideally weighted from existing data in the region or, if none are available, from the literature). If there is a single focal species, the detection probability curve of that species should be used.
- For smaller areas, 20% spatial coverage may be difficult to achieve while ensuring that sampling units are independent (e.g., avoiding double-counting issues). Generally, **transect lines should be a distance apart that is >2 times the effective strip width** (Buckland et al. 2001; Jackson & Whitfield 2011). If focal species are known to be influenced by vessel activity, then boat-based survey transects should also be spaced >2 times the distance at which this behavioral effect is known to occur.
- **Stratified sampling:** Transects should be placed/oriented such that important environmental gradients are fully represented within sampling designs (e.g., water depth, benthic complexity, etc.).

The financial cost of increasing coverage versus the scientific and management value of additional data likely varies based on factors including species exposure levels and effect size. Additional research is needed to refine the 20% coverage recommendation outlined above (see Section 11 for additional details). In the case of digital aerial surveys, it may be possible to collect data at a higher spatial coverage, analyze a subset of the data initially, and then use detection rates and other metrics from the initial dataset to determine if additional data need to be analyzed in order to reliably detect change if it occurs.

10.3.3 Temporal Resolution

In addition to spatial scale considerations, the temporal resolution of surveys requires careful consideration to ensure that surveys are statistically independent while capturing adequate variability in the abundance and distribution of marine birds over time. Previous analyses using data from the Northwest Atlantic Seabird Catalog found that surveys conducted 3+ days apart can be considered independent (Kinlan et al. 2012). However, this should be balanced with consideration of spacing to capture seasonal variability (AMBC 2021).

- For studies to detect effects, **12–16 surveys per year for at least two years pre-construction should be conducted to adequately capture variation in distributions** (Kinlan et al. 2012). Two years of monthly surveys are currently recommended in the BOEM avian survey guidelines (BOEM 2020). In addition, pre-construction surveys need to commence early enough (minimum of two years) to allow for completion prior to the start of construction. There should be no more than five years between pre-construction data collection and the first post-construction data collection to avoid introduction of additional sources of variation.
- **The duration and frequency of post-construction surveys should depend on the question (e.g., interest in temporal patterns of displacement/habituation) and levels of variability in site-level data but should include no less than 3 years of 12–16 surveys per year** (Percival 2013). Particularly for low abundance species and/or those with low effect sizes, additional surveys may be needed to achieve sufficient statistical power (Vanermen et al. 2015b). Studies focused on temporal patterns/habituation should aim to survey periodically throughout the lifespan of the project.
- **The distribution of surveys within a particular year should take into consideration seasonal patterns of focal species**, as increases in power can be achieved if effort is concentrated in seasons in which species of interest are most abundant (Maclean *et al*. 2013).

10.4 Data Collection Methods

In addition to the above survey design topics, there are several other key considerations to obtain high-quality data from surveys. Some of these are applicable across multiple types of observational survey, while others are specific to boat-based or digital aerial surveys. Conducting surveys in the same way pre- and post-construction is not always possible, but care should be taken to make post-construction surveys as similar as possible to pre-construction surveys to allow for strong comparison of the two datasets. Generally, **to the degree possible, survey methods, including data collection methods, should be consistent across pre- and post-construction surveys so as not to introduce biases relating to changes in survey methods that are unnecessary or unaccounted for (BOEM 2020).** Upgrades in survey capabilities (i.e., new camera systems for digital aerial surveys) should still be pursued for integration into survey designs post-construction, if they are available, especially if they provide significant improvements in data quality or safety. **If** substantial aspects of the study design or survey methods change between survey periods, however, calibration studies must be conducted to understand the effect of these changes on detection rates, identification rates, and the behavior of the animals being surveyed, to inform viable approaches for data analysis (Matthiopoulos et al. 2022).

10.4.1 Sampling Method

Sampling methods should be used that allow for correction of potential biases and follow established methods. Specific characteristics of survey platforms are discussed in "Platform height and other characteristics," below.

Boat-based surveys: As noted in the BOEM avian survey guidelines (2020), line transects with distance-sampling methods should be used for boat-based surveys (Buckland et al. 2001; Camphuysen et al. 2004; Ballance and Force 2016). The observer should search within a 90-degree bow to beam arc either to port or starboard of the track line (ideally the side with the best visibility) to detect individuals prior to their response to the survey platform (Buckland et al. 2001). Individual birds and groups of birds should in turn be identified with an estimate of distance and bearing along with behavior (see "Data Collection," below). Before surveys, observers should calibrate distance estimates using a laser rangefinder on inanimate objects (e.g., buoys; BOEM 2020). Observers should aim to detect and record all birds with no *a priori* truncation of distance (Buckland et al. 2001; Camphuysen et al. 2004; Ballance and Force 2016; Bolduc and Fifield 2017). While data analysts may need to truncate the maximum distance to optimize model fit, it is best to leave them with the decision of how to implement that with continuous distance estimates. If this is infeasible due to unusual survey constraints or exceptionally high bird densities, we recommend ignoring the collection of distance data, as the detection process can be assumed from the recorded data in other locations in some situations (Goyert et al. 2016).

If expected detection rates or study design do not allow for a true line transect approach, predefining distance bands (e.g., 0–100m, 100–200m, etc.) and assignment of birds to each band during observation can be an acceptable alternative approach. However, distance bands must be carefully selected *a priori* and must be useful to all the study species of interest. Regardless of whether line transects or distance bands are used, boat-based surveys conducted pre- and post-construction to evaluate changes in marine bird distributions must address fundamental requirements to 1) use a standardized, replicable sampling protocol, 2) allow for extension of inference from the sampled population to a clearly delineated biological population, and 3) adjust for detectability bias that arises from distance, movement, environmental covariates, and other relevant factors.

Digital aerial surveys: follow existing guidelines (BOEM 2020) to use strip-transect or grid sampling methods. Either of these methods may be used in a model-based analysis (e.g., before-after-gradient design). Continuous strip transects, such as digital video or abutting digital still imagery, may better capture sampling gradients, but may have high variance due to autocorrelated distributions of aggregated (e.g., flocking or schooling) species. Grid samples, often used with digital still photography, may better handle aggregated species by reducing autocorrelation, but are generally more expensive per unit of observation data.

10.4.2 Consistency in Survey Platform

If possible, the same platform (e.g., the specific boat or plane as well as camera setup for digital aerial surveys) should be used for pre- and post-construction surveys to control for detection differences that may be caused by different platforms. If a different platform is used for pre- and post-construction surveys, the potential biases caused in the resulting dataset due to variation in size, platform height and field of view, etc. must be explicitly addressed in the study plan and data analysis (Section 7) or via targeted calibration studies (see Munson et al. 2010 and Matthiopoulos et al. 2022).

10.4.3 Platform Speed

- **Boat-based surveys: A speed of 7–10 knots is recommended for boat-based marine bird surveys**. Platforms moving <4 knots (7.4 km/h) or >19 knots (235.2 km/h) are not appropriate for collecting marine bird survey data (Gjerdrum et al. 2012). The existing BOEM guidelines for site characterization surveys recommend 10 knots (BOEM 2020).
- Digital aerial surveys: follow existing guidelines and fly surveys between 220–350 km/hr (ground speed; BOEM 2020). Speed should not be significantly varied between surveys, or within surveys (less than+/-10% fluctuation), during periods when imagery is being collected for analysis.

10.4.4 Platform Height and Other Characteristics

The choice of survey platform, and specific location from which observations are conducted/images are recorded, can have a large influence on the quality of resulting data. For boat-based surveys, in general, observers should be located high above the water's surface in a location with a wide forward field of view. Larger boats can also conduct surveys safely in a wider range of weather conditions. However, vessel availability is also a consideration; if a slightly smaller vessel will be more readily available for surveys when there is a weather window, which might be preferable to a larger vessel that has more limited availability for surveys. In addition, a vantage point that is too high can negatively influence detection for some species. Surveyors should also consider safety and observation efficacy when selecting a survey platform on the vessel. The location of survey observers on the vessel should be:

- At a position above sea level that enables detectability within a minimum of 300 m of the trackline for focal taxa, ideally ~10 m (range: 5–25 m; Camphuysen et al. 2004). A vantage point that is too high or low can negatively influence the detection of some birds, particularly small, dark birds near the water's surface. Positions within a couple meters above sea level (e.g., small recreational boats) can limit the depth of field for distance estimation, such that farther distances (e.g., > 100m) are indistinguishable. Taller platforms (e.g., > 5m above sea level) are recommended to better distinguish farther distances but may require careful selection of observation points to prevent the ship breadth from blocking the view alongside the vessel.
- Have a clear (>90 degree) field of view to the front and side of the vessel.
- Be a safe location from which to conduct surveys (e.g., without having to hold onto railings or other infrastructure).
- Be a stable location from which to conduct surveys (e.g., a crow's nest or similar platform that tilts back and forth with wave action is generally not going to be an effective location from which to conduct surveys).

For digital aerial surveys, there is a key tradeoff between flight height of the plane (i.e., higher flights increase crew safety, make it easier to conduct surveys using the same methods pre- and post-construction, and reduce wildlife disturbance caused by the plane) and image resolution (i.e., higher flights may result in lower image resolution and fewer birds identified to species).

- For digital aerial surveys, surveys should ideally be flown at the same altitude pre- and post-construction, but at minimum should have consistent image resolution between these survey periods to provide the most comparable data between these two periods (see data collection section below for additional recommendations on image resolution). The optimal flight height for a given situation will be a balance between (1) obtaining the necessary image resolution (see data collection section below), and (2) flying at heights that eliminate disturbance to wildlife (500 m minimum; AMBC 2021[6]) and allow safe flying above turbine rotors. However, flight height may evolve as camera resolution and technology improves (e.g., by the time post-construction surveys are flown for a project, it may be possible to fly higher while retaining the same image resolution as pre-construction surveys). Given current Federal Aviation Administration guidelines, for safety reasons, planes will likely be required to fly at least 500 feet above the upper edge of the rotor-swept zone (14 CFR 91.119).

- In many cases, exact turbine height will not be known at the time that pre-construction surveys are flown. In this situation, the most conservative estimate of turbine height should be used (e.g., higher end of the design envelope identified in the Construction and Operations Plan) to identify a safe flight height for surveys.

10.4.5 Surveyor Qualifications

The value of data is directly related to its quality, which depends on the capabilities of the surveyors as well as the quality of training provided (Environment and Climate Change Canada 2020). Current BOEM avian survey guidelines recommend the use of "qualified biologists specializing in seabirds" for surveys (BOEM 2020), but how qualification is determined is not clearly defined. In the UK, commercial and volunteer boat-based surveyors are assessed by accredited instructors on five key standards – bird identification, visual acuity, application of methods, recording stamina, and navigation (Lewis & Dunn 2020). Based on these standards and the Eastern Canada Seabirds at Sea standardized protocol for pelagic seabird surveys, we recommend the following:

- **Observers/biologists conducting boat-based surveys or identifying images from digital aerial surveys must have documented experience observing and counting seabirds with a good understanding of seabird behavior and ecology. Experience includes at least 50–100 hours of training with qualified observers/biologists** (Environment and Climate Change Canada 2020, Jackson & Whitfield 2011).
- **Observers/biologists should have demonstrated ability to rapidly identify seabirds at sea/from images in the region in all plumages, in various lighting conditions, under reduced visibility, and in rough sea conditions.**

10.4.6 Survey Conditions

The weather conditions (e.g., visibility, sea state, glare) during which surveys can be conducted should be defined based on human safety considerations as well as quality of data collection. Conditions can significantly impact detection rates, leading to biases in resulting data. An improved understanding of the relationship between survey conditions and species detection and identification could aid in developing a correction to allow for a broader range of conditions to be acceptable for conducting surveys. Unless there are data available with which to correct detection probabilities based on differing conditions, and

[6] From the AMBC 2021 letter to BOEM: "Published studies suggest that digital aerial surveys should be flown above 460 m, preferably a minimum altitude of 500 m, to avoid disturbance (Thaxter et al. 2016), but operators have reported minimal disturbance or flushing of target species during surveys conducted at 415 m. More empirical support is needed to determine the ideal minimum survey altitude, and whether it should range depending on environmental conditions."

these differing conditions remain safe for those conducting the surveys, we recommend that surveys are conducted in the following weather conditions:

- **Boat-based surveys: In general, surveys should be conducted at no higher sea state than Beaufort 4 and with >1 km visibility** (with the exception of large research vessels specifically designed for survey work that can remain safe and provide a stable viewing platform in conditions up to sea state 5–6). **As much as possible, transect orientation and observer orientation during surveys should be designed to minimize glare-related effects on detections (BOEM 2020).** Following existing BOEM guidelines (2020), surveys should commence when there is enough light to identify birds to species. Boat size and platform height, and conditions in which surveys were conducted, should always be noted in metadata such that these variables can be included in future data analyses.
- **Digital aerial surveys: Surveys generally should be conducted at no higher than Beaufort 4 (BOEM 2020).** Higher sea conditions may lead to both safety concerns and the potential to miss smaller species depending on the region and species present. Glare, likewise, can affect species detection and identification. **Prior to initiating surveys, transect orientation should be designed to minimize glare (while also designing surveys to cover important environmental gradients).** Additionally, when conducting surveys, the angle and height of the sun should be carefully considered when assessing survey conditions for glare, and cameras that can be rotated (e.g., away from the sun) are an effective way to avoid glare. **Light conditions should be adequate for species identification in imagery (BOEM 2020).** Flight altitude and speed, and conditions in which surveys were conducted (such as sea state and glare), should always be noted in metadata to inform future data analyses.

10.4.7 Data Collection

Data collection on each survey should encompass information on survey conditions, timing, level of effort, and bird observations. The general information collected during surveys should be consistent with existing guidelines (BOEM 2020, Normandeau 2012).

- **Survey data collection should include effort data and information on weather conditions at the scale of the transect segment, where a new transect segment is defined by a change in any one of the conditions listed below. Effort/conditions data should include, at minimum:**
    - Full time-location track information, including the start and end date and time
    - GPS track of transect with associated time of each position
    - Sampling method (e.g., line transect, strip transect, grid sampling)
    - Sea state (boat-based surveys)
    - Visibility
    - Glare (digital aerial)
    - Observer ID
    - Altitude of plane (digital aerial) or height above sea level of observer (boat-based)
- **Data collected for each observation should include, at minimum:**
    - Date and time
    - Location (latitude and longitude)
    - Species identification
    - Number of individuals in group

- Behavior (such as flying, on water, foraging)
    - Distance and angle (with certain short-term exceptions based on conditions; see above)
    - Non-bird objects/events that could influence distributions (e.g., fishing vessels, debris, sea turtles, fish, and marine mammal observations). If the observer can collect data on other animals observed during surveys, they should do so consistently. If data on non-bird animals is only collected during portions of the survey, or for certain non-avian taxa, this effort-related information should be included with the observation data. Unless systematically recorded, these observations should be treated as opportunistic.
- **Data collected for each observation should also, where possible, include:**
    - Bird flight direction
    - Flight height, collected using the best available science. In the case of boat-based surveys, ornithodolites/laser rangefinders paired with inclinometers should be used to the degree possible for flight height estimation of all individuals, due to lesser accuracy of purely visual flight height estimates from vessels (Largey et al. 2021). At minimum, such systems should be used for calibration and training of observers (Harwood et al. 2018). If binning flight height data, categories should be carefully considered (based on project's proposed rotor swept zone) and consistent across observers, surveys, and studies. For example, AMAPPS[7] surveys use 0–10 m, 10–25 m, 25–50 m, 50–100 m, 100–200 m, >200 m bins. For digital aerial surveys, recent advances in LiDAR and digital aerial imaging also offer the potential to collect estimates of the altitude of birds in flight (Cook et al. 2018, Humphries et al. 2023) and should be used whenever possible. Biases associated with the chosen method for estimating flight height should be carefully considered and explicitly stated in study design and reporting.
- **Birds should be identified to species whenever possible (but only when confidence in identification is high); if this cannot be done, then birds should be identified to lowest distinguishable taxonomic group, as recommended in the BOEM guidelines (BOEM 2020)**. While confidence in identification is subjective, a common set of identification criteria should be used by all observers.
- For digital aerial surveys, color images should be collected with a ground spatial resolution of 2 cm or finer. Image resolution is a key factor influencing species identification for digital aerial surveys and should be somewhat dependent on species of interest. The recommendation to use 2 cm resolution or finer is applicable regardless of survey intent, finer resolution may be obtained to allow for distinction among similar small-bodied species of particular interest (e.g., auks, terns). For boat-based surveys, color images using a digital camera with telephoto lens should be collected, where possible, of birds, with a particular focus on (1) rare species and (2) species that are difficult to distinguish (e.g., tern species).
- **Survey data should be collected and recorded in a standardized way that can seamlessly be incorporated into the Northwest Atlantic Seabird Catalog and other data repositories**. To improve data standardization and workflow, boat-based surveys should collect data using a survey application, such as SeaScribe (Gilbert et al. 2016) or Sealog (Swingley et al. 2023).
- **Careful consideration should be given to the collection of *in situ* environmental and prey data simultaneous with bird observations, continuously or at regular intervals (e.g., hourly or per**

[7] Atlantic Marine Assessment Program for Protected Species: https://www.fisheries.noaa.gov/new-england-mid-atlantic/population-assessments/atlantic-marine-assessment-program-protected

transect) to inform data modeling and mechanisms of potential effects from OSW development on marine bird habitat use, abundance, and distribution**.** Environmental data could include weather conditions for each observation, water temperature and salinity (for boat-based surveys), and prey information including hydroacoustic surveys of fish biomass (for boat-based surveys) or the location and size of fish shoals identified in images from digital aerial surveys (Goetsch et al. 2023).

10.5 Review of Data

Data collected on each survey should be reviewed for quality control purposes.

- **Boat-based surveys: data should be summarized and reviewed by one or more of the observers for obviously erroneous information**, with a particular focus on species and counts to ascertain incorrect information was not recorded (for example, the standard 4-letter species code is ROST for Roseate Tern and ROYT for Royal Tern). Preliminary data review should be carried out as soon as possible (within 48 hours of survey completion) to prevent any potential errors being overlooked. Any unidentified individuals for which images were taken should be identified from the photographs, if possible.
- **Digital aerial surveys:** following the BOEM avian survey guidelines, **qualified biologists specializing in seabirds should assess images, and at least 20% of images should be independently audited by an expert during both the detection and identification stages of the review process** (see Buckland et al. 2012).

10.6 Data Analysis

The current BOEM avian survey guidelines (2020) provide useful guidance for analysis regardless of whether surveys are intended to inform site assessment or to assess effects of OSW on marine bird distributions. **The development of a clearly defined analysis plan (See Section 7) should include specific models and statistical tests along with the following considerations specific to surveys:**

- **Accounting for biases**: Following existing BOEM avian survey guidelines, for line transect sampling from boats, distance sampling data should be used to model species-level distance functions (see Buckland et al. 2001) to correct density and abundance estimates. Analyses should use formulations of distance models that allow for inclusion of covariates (observer, sea state, etc.). While detectability is assumed to be constant across the captured area for digital aerial surveys, species-level and condition-dependent detectability should be considered, as appropriate. Availability bias is an additional important consideration, perhaps particularly for digital aerial surveys that move much faster than boat surveys and therefore may have a higher availability bias for some diving species (e.g., Winiarski et al. 2014). Data on activity budgets from tracking studies (existing or future) may be required to adequately characterize species-level availability biases to allow for corrections. In addition, accounting for uncertainty in species identification can be achieved using various analytical methods, including multiple simulation approaches (see Johnston et al. 2014 for details on approaches).
- **Choosing the appropriate modeling framework**: There are multiple modeling approaches that provide methods to examine displacement and attraction effects for gradient study designs comparing pre-construction and post-construction distributions, including generalized linear mixed models (GLMM), generalized additive mixed models (GAMMs), Poisson point processes,

and Complex Regional Spatial Smoother models (CreSS). All have strengths and limitations given data and research questions, but in an analysis comparing analytical methods for offshore renewable energy surveys, CreSS performed better than GAMMs at assessing whether effects were present and at identifying spatially explicit differences (Mackenzie et al. 2013). Comparisons between spatial modeling approaches will be needed during analysis to identify the best choice for a given study.

- **Accounting for autocorrelation.** Spatial and temporal autocorrelation is highly likely to be present in observational survey datasets and should be adequately accounted for in study design and/or analysis. Observations collected close together in space and time may be more similar than those collected further apart, resulting in autocorrelation among count data. If similarities are not accounted for in analysis, it can lead to an underestimation of uncertainty and thus an overestimate of effect size. Correlograms or variograms, for example calculating Moran's I, may be used to test for spatial and temporal autocorrelation in the data and the residuals of a model. Autocorrelation may be minimized through the use of design-based studies (e.g., grid sampling) or model-based analyses. For example, inclusion of autocorrelated predictors in models may remove some of this non-independence, in which case model tests should indicate no residual autocorrelation. Where predictors do not sufficiently account for such autocorrelation, other methods, such as conditional auto-regressive (CAR) models or Generalized Estimating Equation (GEE; Hardin & Hilbe 2002) can be used to account for this type of autocorrelation.
- **Comprehensive identification of covariates helps ensure successful model selection** as these covariates help control for variability in response to the underlying environment (e.g., changes in distributions/abundance) that is not attributable to OSW development. The choice of covariates will vary depending on research questions, focal taxa, biological relevance, and data availability.
  - **Potential covariates should include, to the extent available, environmental variables (e.g., bathymetric features, flow dynamics) as well as existing anthropogenic pressures (e.g., vessel traffic)** based on existing information about the biological relevance and influence of these variables on abundance/distribution of focal taxa (Mackenzie et al. 2013).
  - **To describe effects across small spatial scales (10s of km), a relatively high spatial resolution of covariates is most appropriate (e.g., at the resolution of turbine spacing or higher).**

10.7 Data Reporting

Standardized reporting should include information on data collection methods (including boat size and platform height), spatial and temporal coverage, effect size, uncertainty, and assumptions, such that survey data can be integrated into future meta-analyses and other assessments (Section 8). For observational surveys in particular, key aspects of reporting include the following:

- **Report study design information including spatial and temporal coverage of surveys** (% spatial coverage, distance between transects, buffer size/area, overall survey area in km^2, timing of surveys).
- **Following existing BOEM avian survey guidelines (BOEM 2020), provide spatially explicit density estimates and associated variance (95% confidence intervals) by species/taxonomic groups in map and tabular formats**. Uncertainty about estimated parameters is crucial when drawing conclusions from a model. 95% confidence intervals can be used as best- and worst-case

scenarios, as well as provide key information about uncertainty of effects for future meta-analyses.

- **Provide information on site characteristics** including latitude and longitude, OSW project footprint size, distance between turbines, number of turbines, height of turbines, minimum and maximum water depth, and minimum and maximum distance to shore.
- **Make observation datasets publicly available** via the Northwest Atlantic Seabird Catalog and/or OBIS-SEAMAP (BOEM 2020, NYSERDA 2021). This should include final processed dataset(s) (following QA/QC), co-collected environmental covariate data, complete effort data, and comprehensive metadata (NYSERDA 2021). Until a suitable database or archive for digital aerial survey imagery is developed, projects should aim to at least make clipped 'snag' images available publicly online via searchable websites. Full images, image metadata, and image annotations (e.g., observation data associated with each frame) should be archived for the life of the OSW wind project, and in such a manner that they can be easily made available on request of federal and state regulatory agencies for machine-learning applications or other purposes. For guidance on formatting requirements and archiving of digital aerial imagery, contact Kyle Landolt at klandolt@usgs.gov at the Upper Midwest Environmental Science Center.
- Make data publicly available as soon as possible, but within a maximum of two years following collection, if feasible. For multi-year data collection, subsets of data should be released as they are finalized to ensure that the data can be incorporated in a timely way into broader efforts.

Additional recommendations for data transparency and reporting are discussed in Section 8, above.