The Scottish Government
Riaghaltas na h-Alba

Marine Scotland Science

Marine Scotland Science Report 05/14

Statistical Modelling of Seabird and Cetacean Data: Literature Review

Marine Scotland Science Report 05/14

Statistical Modelling of Seabird and Cetacean Data: Literature Review

C.S. Oedekoven, M.L. Mackenzie, L.A. Scott-Hayward and E. Rexstad

This report constitutes work carried out at the Centre for Research Into Ecological and Environmental Modelling (CREEM) at the University of St. Andrews, performed under contract for Marine Scotland.

Published by Marine Scotland Science

This report presents the results of marine and freshwater scientific work carried out for Marine Scotland under external commission.

Marine Scotland Science
Marine Laboratory
375 Victoria Road
Aberdeen
AB11 9DB

Copies of this report are available from the Marine Scotland website at www.scotland.gov.uk/marinescotland

Statistical Modelling of Seabird and Cetacean data: Literature Review

C.S. Oedekoven, M.L. Mackenzie, L.A.S. Scott-Hayward,
E. Rexstad

This review constitutes work carried out at the Centre for Research Into Ecological and Environmental Modelling (CREEM) at the University of St. Andrews, performed under contract for Marine Scotland.

Please reference this document as: Oedekoven, C.S., M.L. Mackenzie, L.A.S. Scott-Hayward, E. Rexstad (2013). Statistical Modelling of Seabird and Cetacean data: Literature Review. University of St. Andrews contract for Marine Scotland; SB9 (CR/2012/05).

Contents

1 *Introduction*

[1] This document forms an overview of the survey and analysis methods used by the marine renewables industry for baseline monitoring and quantitative environmental impact assessment. The information contained within this report is derived from environmental statements submitted to Marine Scotland but also contains other documents that have been made available for review. This review is not intended to be an exhaustive or detailed discussion of all of the potential survey and analysis methods appropriate for (or used by) groups with an interest in marine renewables.

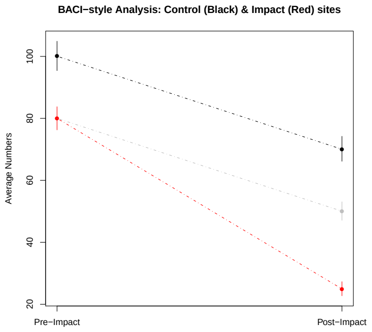

Figure 1: 95% Confidence Intervals for average numbers before and after impact for a control and impact site. Black line is control site, grey line represents expected behaviour of impact site without an impact, the red line represents behaviour in impact site in the presence of an impact.

[2] We begin with a brief overview about how survey and analysis methods are typically conducted for studies of this type and then discuss particulars about the survey and analysis methods. Examples from industry are mentioned throughout the document and the majority of the packages and functions referred to in this document relate to R software (**?**).

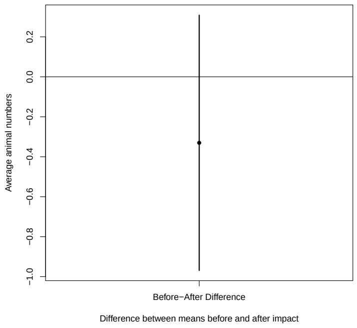

Figure 2: 95% Confidence interval (vertical line) for the average difference (black dot) in animal numbers before and after impact.

2 *Overview*

[3] Environmental impact assessment has traditionally taken the form of Before-After-Control-Impact (BACI) designs where at least one potentially impacted site and at least one control site (that cannot receive the impact) are sampled before and after the potential impact. Control and potentially impacted sites are surveyed and the number of animals is compared at one or more time points (**??**). The BACI approach assumes naturally occurring changes will appear in both the control and impact sites. Hence unusually large post-impact differences between the control and impact sites can be attributed to the impact. The analysis of BACI data tend to involve simple comparisons of average numbers (or differences in average numbers) at the control and impact sites before and after impact with associated measures of precision, examining the interaction term (depicted as difference in slopes between the black and red lines in Figure 1).

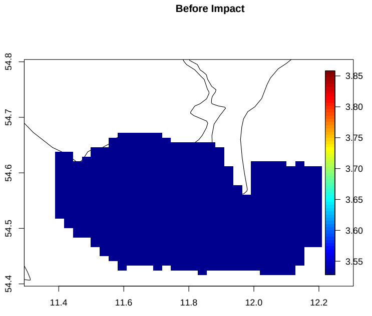

Figure 3: A spatial illustration of a single before/after difference comparison: Before value based on this mean-based comparison. Entire study area is of a single colour because there is no spatial aspect to this assessment.

[4] The effectiveness of the BACI approach at quantifying environmental impacts depends, among other things, on the availability of good control sites which, in practice, may be difficult to find or may not exist. The survey area is sometimes delineated into a potentially impacted zone (near the source of the potential impact) and a buffer zone (typically located further from the potential impact). This buffer zone is created to act as a control site and is assumed to be unaffected by the potential impact. While choosing a buffer zone is more convenient than finding an independent control site, choosing the distance at which to delineate the impact site from the control/buffer zone (or indeed the shape of these buffer zones) is difficult. If these zones are not chosen correctly, subsequent impact-control comparisons are useless.

[5] In the absence of appropriate control sites, Before-After (BA) comparisons are sometimes used to compare average numbers in the same potentially impacted site(s) before and after impact (**?**). The BA comparison assesses whether post-impact numbers are within the normal range of numbers seen pre-impact. Unusually large differences pre- and post-impact provide evidence of an impact having occurred.

[6] This comparison can only be correctly attributed to the impact if these large differences would not have occurred in the absence of an impact, and this is almost impossible to ascertain based on simple comparisons of average numbers pre and post impact. For example, average pre and post numbers can be compared using a simple t-test with associated confidence intervals for the average difference (Figure 2). This analysis is simple and based on a single abundance estimate pre- and post-impact for the whole survey area (Figure 3 and 4) and the investigator may not be able to attribute any overall changes in abundance to the impact. This comparison also fails to detect potential redistribution – shifts in the spatial location of the animals in and around the impact (e.g. Figure 6).

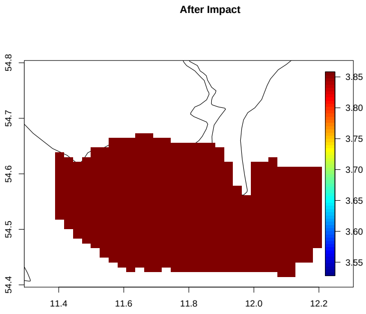

Figure 4: A spatial illustration of a single before/after difference comparison: After value based on this mean-based comparison (single colour for reason described in Figure 1).

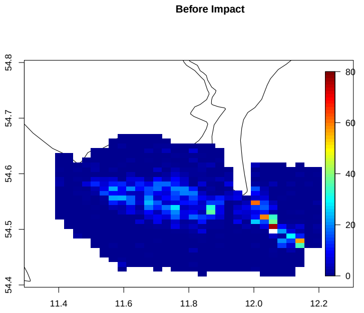

Figure 5: Average across time in numbers of animals in each grid cell before impact.

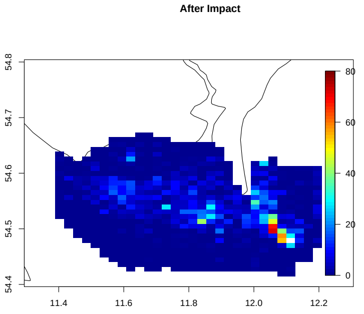

Figure 6: Average across time in numbers of animals in each grid cell after impact. Note that some redistribution into the south-eastern quadrant of the surveyed area occurred after the impact.

[7] In contrast, a "gradient-style" analysis of Before-After data may be considered treating pre-post differences as a function of distance from the impact source (Figure 7). This approach assumes any impact-related differences will decay with distance from the impact source (**?**) and has a one dimensional (Figure 7) and two dimensional representation (Figures 8 and 9). This gradient-based method assumes that any changes pre- and post-impact are a function of distance and thus any impact-related effects are the same in all directions from the point source (given some specified distance).

[8] This simple gradient based method can be problematic since the impact effect is not guaranteed to be symmetrical about the point source. If this is assumed, the gradient-based approach will at best mischaracterise the impact, and might fail to detect an impact altogether. This failure could occur if a decline on one side of the point source is balanced by an increase on the other side. If the impact effect differs with the direction from the point source then the associated analysis must be able to accommodate this feature. Even a direction-based interaction effect inside a gradient design analysis (e.g. permitting the nonlinear gradient to change in each of 4 quadrants: NE, SE, SW, NW) will not capture many types of re-distributions.

[9] To capture an impact effect which differs with the direction from the point source, a surface fitting approach can be used to model the pre (and post) distribution of the focus species. These methods are discussed in section 5.

3 *Survey methods*

[10] The visual observations used as input for these methods are typically collected along transects from aircraft or boats either using observers or imaging devices from the plane (section 3.1). For sites close to land, observers are placed at one or more vantage points and the survey area is scanned by eye or binoculars during survey times. The details of the survey methods employed tend to vary with site but some examples are discussed in section 3.2.

[11] The visual observations collected from boats, visually from aircraft and on-shore vantage points are typically subject to perception bias since animals are increasingly difficult to see as their distance from the observer increases. Raw counts (based on visual observations) generally need to be corrected for imperfect detection. These methods are discussed in section 4.

3.1 *Boat and aerial surveys*

[12] Boat and aerial surveys are most useful when the site of interest is some distance from shore, i.e., sites investigated for offshore wind or wave and tidal installations. The primary goal of these surveys is to cover the area as effectively as possible, given a budget and time scale, and obtain accurate estimates of the number and distribution of focal species.

[13] **?** recommended to survey seabirds at sea from boats using line transects where the perpendicular distances to detections of non-flying birds are recorded in intervals (e.g. 0-50, 50-100, 100-200, 200-300m, >300m when observing from 10m above sea level). As seabirds generally fly faster than the ship's speed and animal movement may lead to biased abundance estimates (see assumptions from section 4.1), the snapshot method was recommended for flying birds with 1- or 5-minute intervals. **?** also suggest that the study area be at least six times the size of the developmental area, including one or more similar sized control areas featuring the same oceanographic characteristics and this control area should be at least 1.5km from the nearest turbine (**?**).

[14] Notably, this advice was issued with off-shore wind sites in mind and the relative size of the developmental and control/buffer areas will likely need modification for the smaller wave/tidal at-sea sites. In particular, it is important to ensure that sufficient sampling

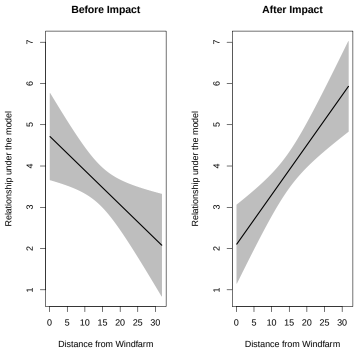

Figure 7: One dimensional representation of changes in animal numbers with distance from the impact source pre- and post-impact.

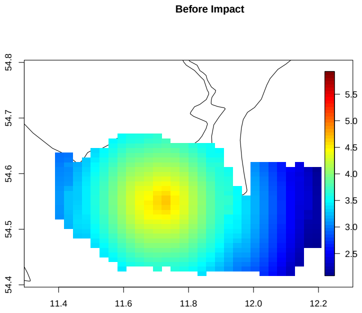

Figure 8: A spatial illustration of the linear distance-from-impact model: Before Impact. Colours represent animal density.

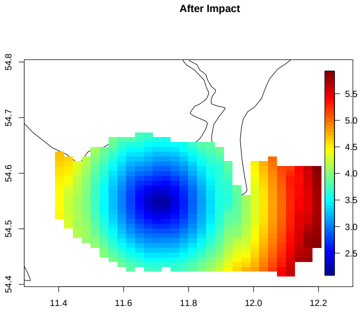

Figure 9: A spatial illustration of the linear distance-from-impact model: After Impact.

effort is allocated to the development area to obtain reliable data for the purposes of monitoring and impact assessment.

[15] **?** also recommended that survey lines should be spaced 0.5-2 nm apart and surveyed in sea state conditions less than 5, whilst ensuring the survey area was surveyed throughout the day to avoid any potentially confounding diurnal rhythms of the animals. For aerial surveys, **?** recommended that parallel transects should be located 2km apart and that all detections should be recorded in distance bands which depend on the type of window the survey plane is equipped with, the flight altitude and the declination in degrees from the horizon used. It was also suggested that surveys be conducted in sea states less than 4. This transect spacing advice more readily relates to the larger off-shore wind sites and may need substantial modification for smaller wave and tidal sites.

[16] Examples of use of the **?** recommendations include proposed development in the the Neart Na Gaoithe study area that encompassed the developmental area (approximately $105 km^2$) and a surrounding buffer zone of 8km width. Line transects were spaced 2km apart and surveys were conducted monthly. Observers scanned one forward quadrant from the ship out to 300m and distance to detections were recorded in predefined distance bands: 0-50, 51-100, 101-200, 201-300, >300m. The count interval for surveys was in 1 minute intervals[1].

[17] The sampling regime for the Beatrice Offshore Wind Farm also used the survey protocol outlined in **?**. This site was sampled using monthly ship-based surveys between October 2009 and September 2011 and included the developmental area (approximately $131.5 km^2$) and a surrounding 4km buffer zone. Non-flying birds were recorded in distance bands: 0-50, 50-100, 100-200 and 200-300m and flying birds were recorded in 1-minute intervals. During aerial surveys, birds were recorded in distance bands: 44-163, 163-282, 282-426, 426-1000m and digital aerial surveys were also conducted using digital video survey methods where transects were spaced 2km apart.

[18] For the Aberdeen Offshore Wind Farm, monthly surveys were undertaken between February 2007 and April 2008 and between August 2010 to July 2011. Non-flying birds were recorded in predetermined distance bands out to 300m in one forward quadrant of the boat. Flying birds were recorded with the snapshot method using 1-minute intervals. The study area comprised the developmental area and a surrounding buffer as well as a control area. The first phase of surveys took a Before-After Control-Impact design approach, while the second phase took a Before-After Gradient design approach. For those bird species with less than 100 detections, no account of imperfect detection was carried out due to small sample

---

[1] 2012. Neart na Gaoithe Offshore Wind Farm Environmental Statement. Chapter 1. Introduction to the Neart na Gaoithe Proposed Offshore Wind Farm Development 1.1. Technical Report. Mainstream Renewable Power.
2012. Neart na Gaoithe Offshore Wind Farm Environmental Statement. Chapter 13 Marine Mammals. Technical Report. Mainstream Renewable Power.
2012. Mainstream Neart na Gaoithe Offshore Wind Farm Ornithology Technical Report. Appendix 12.1 to Neart na Gaoithe Offshore Wind Farm Environmental Statement. Natural Research Projects Limited.

sizes. The authors used simple extrapolations of the overall density estimates across the study area. This implied using correction factors from published literature to adjust the raw counts for imperfect detection.

[19] In the Dutch North Sea, as a part of the Shortlist Masterplan project, nine monthly aerial surveys using line transect methodology were carried out in May-October 2010, as well as January, February and April 2011. Perpendicular distances from the line to the detections were recorded in bands. For the survey design, 35 lines of approximately 75km length and oriented perpendicular to the coast were laid out between the borders to Belgium and Germany. Specifically, 11 ship-based surveys were conducted between April 2010 and February 2011. Marine mammal and seabird observations were made during the "fish eggs and fish larvae" surveys of the Shortlist Masterplan project using line transect methodology. All detections within 300m on one side of the ship were recorded in predetermined distance bands (0-50, 50-100, 100-200, 200-300, >300m) in 5-minute segments. Flying birds were recorded separately as within or outside 300m. As egg and larvae sampling was carried our around the clock, marine mammal and seabird observations had to be done during transits between sampling stations. These transit lines covered roughly the same outlined area as the Shortlist Masterplan aerial surveys, however, the design was non-random and the coverage of the area uneven as they were transiting between predetermined plankton sampling stations.

[20] Concurrently, five bimonthly aerial surveys using strip transect methodology were conducted as part of the Monitoring Waterstaatkundige Toestand des Lands MWTL monitoring program. These surveys were designed to cover the entire exclusive economic zone of the Netherlands, with two lines running parallel to the coast across the entire Dutch coast while other lines covered the areas further offshore in a non-systematic pattern.

[21] The developmental area in the Moray Firth zone was also surveyed in line with **?**. The study area of $522.2km^2$ includes sites for three proposed wind farm sites. The studies conducted in this area included boat-based and aerial line transect surveys between April and October 2004 and 2005 and 28 boat-based marine mammal and seabird surveys between April 2010 and March 2012 conducted by Natural Power Consultants. The line transect surveys followed a survey protocol as recommended by **?** and **?** and were generally conducted in passing mode which implies no deviation off the trackline when a sighting is made (e.g. for the purpose of species identification or improving group size estimates). The survey area included the three proposed sites for development as well as a buffer of 4km around them. Lines were laid out perpendicular to the coast. The 28 boat-based marine mammal and seabird Moray

Firth zone surveys (April 2010-March 2012) involved snapshot counts of flying birds in 1-minute intervals and recording non-flying birds in predetermined distance bands: 0-50, 50-100, 100-200, 200-300, >300 perpendicular to ship in 1-minute sessions.

[22] In the Moray Firth, a total of six digital aerial surveys were also conducted between May and July 2011 to produce additional population estimates and smoothed density surface distribution maps (section 5) for the surveyed area. These were used to put the estimates and distributions from the developmental areas into a wider context and to further address connectivity between the focal species with special protection areas (SPAs).

[23] One consistent feature of these examples is the almost universal use of 2km transect spacing regardless of the size of the survey area. While this survey regime is appropriate in some situations, 2km spacing will likely give poor coverage and replication for small sites (because the transects will be too far apart) and 2km spacing will be impossible to implement for the very large sites due to sampling practicalities and/or prohibitively high cost. Survey design should be based upon allocating effort within the development footprint and the buffer area, rather than based upon transect spacing.

[24] Aerial surveys in these examples do not always seem to use the 2km spacing for their survey design. For instance, the distance between transects was set to 4km in the Moray Firth, but this spacing was specified to change with consecutive surveys to improve spatial coverage. Specifically, aerial surveys were conducted between August and September 2010 in two survey blocks of each $625km^2$ in size with 4km transect spacing. During consecutive surveys, line positions were offset by 1km to increase spatial resolution. In addition, 2 sets of coastal transects were flown, each with one line at 1km and one at 5km from shore parallel to the coast line. Survey blocks were covered 9 times and coastal lines 6 times. Information on sighting conditions was recorded (sea state, glare intensity and using an overall subjective measure with four categories from poor to excellent).

[25] Boat and aerial platforms are ideal for at-sea-surveys however these will not be appropriate for marine renewables sites close to shore. These sites might be better surveyed from shore-based vantage points.

3.2 *Vantage point surveys*

[26] Vantage point surveys refer to the case when the observer conducts repeated surveys generally from one or more points providing a good view of the study area. The surveyed area may encompass the entire area that is in view or may be limited to a predefined arc, e.g. 90°. This type of survey can also provide data on the flight path of birds.

[27] Vantage point surveys may be regarded as a special case of point transect surveys where the surveyed point does not consist of a full circle. However, correcting the counts for imperfect detection (section 4) is not possible without information about the observation process consisting of the observed distances from the point to the detections and the distribution patterns of the animals within the search area. For instance, if few animals are seen far from shore it is not possible to determine if these low numbers are due to the observer failing to see animals which are at the surface, or because there are few animals to be found there.

[28] Unfortunately, separate information about the observation process is rarely available for vantage point surveys and this makes it impossible to disentangle the underlying distribution of animals from the imperfect detection process.

[29] In some cases (e.g. due to the scale of the tidal site) it is claimed to be possible to detect birds accurately from the vantage point[2]. Additionally, more than one vantage point has been used to improve spatial coverage of the site (although the two vantage points were not occupied simultaneously). For Kyle Rhea, the vantage point surveys consisted of repeated and alternating short bouts of three activities: snapshot scans of birds and marine mammals, timed marine mammal watches and timed flying bird watches. While detections out to 1km were claimed to be accurate for this site, the estimated (relative) abundance of these animals systematically declined with distance from each vantage point. This decline in animal numbers with distance from both vantage points could genuinely reflect animal distribution from the survey locations or reflect a mix of the imperfect detection process and the underlying animal distributions.

[30] Vantage point surveys were also conducted for assessing migration of wildfowl across the Moray Firth for the Beatrice Offshore Wind Farm and the information required for correcting the observed counts was not recorded. Specifically, for the Beatrice Demonstrator project, vantage point surveys of the demonstrator site were conducted before, during and after construction of two wind turbines adjacent to the Beatrice Alpha Oil Platform. Surveys were conducted of two 90° arcs, one covering the impact area and one a control area[3]. Vantage point data from the Beatrice Demonstrator sites were summarised as the number of observations per hour which were then used to estimate monthly averages and variance in numbers seen per hour. Density of each species was calcu-

[2] The Kyle Rhea Tidal Stream Array. Volume II. Environmental Statement prepared by SeaGeneration (Kyle Rhea) Ltd. (no year shown on report)

[3] Beatrice Windfarm Demonstrator Project. Environmental Statement. Prepared by Talisman Energy (UK) Limited (no year given)

lated by dividing the number of birds by the area of the surveyed arc without taking into account varying detection probabilities with increasing distance from the point.

[31] For some projects related to offshore renewables this issue is noted. One example is the report detailing the analyses of Aberdeen Offshore Wind Farm data[4]. For this study, weekly surveys of two hours each were conducted from four vantage points over the course of two years (April 2006-March 2008). In the report it is noted that decreasing detection probabilities with increasing distance was not incorporated due to the non-uniform distribution of animals with respect to the coast. The practical consequences of jointly modelling animal distribution and the detection process are outlined in section 4.

[32] Regardless of survey method the interest rarely lies in simple summaries of the counts. Analysis of the observed counts is typically required which may involve correcting the observed counts for imperfect detection and/or modelling their distribution across the site during the baseline and post impact phases. These will be discussed in the following sections.

4 *Correcting for imperfect detection*

4.1 *Conventional distance sampling (CDS) methods*

[33] Distance sampling is a commonly used tool for assessing wildlife populations when the interest lies in the number of animals in a defined study area (**?**). It comprises a suite of methods, e.g. line transect surveys or point transect surveys, that share the common underlying concept that not all animals within the search area are detected and that the proportion of missed animals can be estimated by collecting additional information of distances to the detection.

[34] For line transect surveys, lines are laid out in the study area according to a design and an observer (or team of observers) travels down each line while recording all detections of the species of interest during the survey. For each detection the observer records the perpendicular distance from the line to the detection as well as group size. Similarly for point transects, the observer remains at each point for a fixed period of time and records all radial distances to the detections from the point. A detection function is fitted to the distance data which models the decay in detection probabilities with increasing distance.

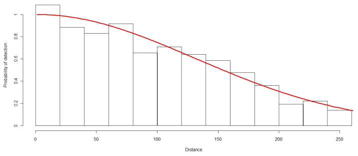

Figure 10: Fitted half normal detection function to underlying histogram of perpendicular sighting distances.

[35] This function may be used to estimate the proportion P_a of

[4] 2013. Aberdeen Offshore Wind Farm Marine Mammals Baseline Addendum. Technical Report. Genesis. Prepared for Aberdeen Offshore Wind Farm Limited
2012. Technip UK Limited âĂŞ Aberdeen Wind Farm Ornithological Baseline and Impact Assessment Addendum. File name: J90008A-Y-RT-24000 G3-Aberdeen Wind Farm Ornithological Baseline and Impact Assessment Addendum. Technical Report. Genesis. Prepared for Technip UK Limited.

animals missed within the search area a. For line transect surveys, the search area is usually defined by the line length and the truncation distance w, i.e. the furthest distance included in the analysis ($a = 2Lw$, where L is the total length of all transects). For point transects the search area around the point is defined by the truncation distance ($a = k\pi w^2$) where k is the number of sampled points. Strip transects are a special case of line transects with small truncation distance and assumed perfect detection throughout the search area, hence, eliminating the need to collect distance data. Strip transect surveys may be useful for monitoring seabirds at sea, e.g. in areas where birds are encountered in high densities and obtaining an unbiased distance measurement for each detection may be infeasible.

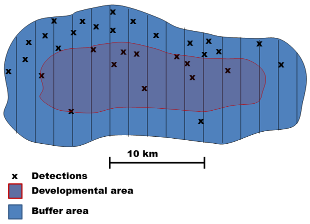

Figure 11: Hypothetical survey design (vertical lines) using systematic parallel placement with random start. Note that transects are placed perpendicular to density contours.

[36] Using conventional distance sampling (CDS) methods, a single estimate of P_a is obtained using a global detection function fitted to all detections made during the survey (unless stratification was performed) **?** proposed to use multiple covariate distance sampling (MCDS) for modelling heterogeneity in detection probabilities. The scale parameter of the half-normal or the hazard-rate function is modelled as a function of covariates that influence detection probabilities. These may include different observation conditions encountered during the surveys (e.g. sea states) or properties of the animals that may render them more detectable in comparison to others (e.g. number of animals within a group). Modelling the detection function with covariates allows us to estimate the average detection probability individually for each detection (P_{a_e} for the eth detection). These P_{a_e} now depend on the observed covariate values for the respective detection.

[37] Given a detection function, we divide the number of detections n by $\hat{P}_a$ to obtain an estimate of the true number of animals $\hat{N}_a$ in the search area. In the case that detections were made of clusters, we multiply $n/\hat{P}_a$ with the expected cluster size (e.g. the mean of the observed cluster sizes) to obtain an estimate of the number of animals $\hat{N}_a$ in the search area.

[38] Using CDS methods, we use a design-based approach to scale up from $\hat{N}_a$ to an estimate of the number of animals in the study area $\hat{N}$, i.e. the population size within the study area. We divide our estimate of the number of animals in the search area by the proportion of the study area that was surveyed: $\hat{N} = \hat{N}_a \times A/a$, where A is the area of the whole study area.

[39] CDS methods rely on several assumptions for obtaining unbiased estimates of animal abundance in the study area (**??**). These assumptions include:

1. animals on the line or point are detected with certainty,

2. the survey was designed with an element of randomisation independently from the distribution of animals and are a good representation of the study area,
3. the survey represents a snapshot moment of the animals in their habitat and, consequently, animal movement may cause bias

[40] For the design-based approach, the survey needs to be designed with an element of randomisation. This is done either by positioning lines or points randomly into the study area or by placing a systematic grid of lines or points into the study area with a randomly chosen starting point (**?**). If density contours (e.g. due to onshore/offshore density gradients) or linear features (such as shelf breaks) are present in the study area, transects should be orientated perpendicular to these contours or linear features. For studies in which the distance between parallel transects is great zig-zag surveys are often used to maximise on-effort time. The disadvantage in designing a survey in this manner is that in those cases where the developmental area represents a small percentage of the study area, only a small percentage of the lines/points will fall into the developmental area. The importance of adequately sampling the developmental area is an issue for all survey designs.

[41] If the survey design has not been carried out with the appropriate randomisation with various gradients in mind, then the sampled portion of the study area cannot be assumed representative of the entire study area. Then model-based inference is used to scale up the number of animals in the covered region to the number of animals in the study area. This is the only means by which valid inferences to the unsampled portion of the study area can be drawn (section 5).

[42] In the past decade much effort has been expended in improving distance sampling by developing methods which allow the relaxation of some of these assumptions. Some of these developments are discussed in this section.

4.2 Imperfect detection on the line

[43] CDS methods assume that all animals are detected on the line, however this is not always realistic. For seabird surveys, imperfect detection on the line may be caused by ship avoidance or by birds diving in reaction to the approaching ship before they were detected by the observers (e.g. some alcid species). This is also a significant issue for digital aerial survey methods since some seabirds (e.g. shags) may spend as much as 40% of their time underwater at-sea (**?**), but is less of an issue with boat-based surveys due to slower survey speeds. For cetacean surveys, the cause is generally diving animals which are not available to be detected. Imperfect detection on the line results in negatively biased abundance estimates and **?** developed mark-recapture distance sampling (MRDS) methods for line transect surveys where detection on the line is not certain.

[44] MRDS has been used in the renewables industry to account for imperfect detection on the line. For Neart Na Gaoithe project, absolute abundances of porpoises were obtained by estimating the average detection probability using MRDS employing functions provided in the `R` package *mrds*. Visual and acoustic detections were used to set up trials for the respective other detection method. Factors affecting the visual detection probabilities such as wind force and sea state were incorporated in the count model as covariates[5].

4.3 *Examples from industry*

[45] In some cases, no correction is made for imperfect detection and this likely results in estimated numbers which are too small. For instance, the aerial survey data for marine mammals at the Neart Na Gaoithe study area were uncorrected for detection issues[6] and the sightings data were used to build distribution maps. For this site, 24 days of surveys were conducted between May 2009 and March 2010 and five species of cetaceans were recorded, most commonly the harbour porpoise. Distribution maps presented report the unadjusted sightings.

[46] From the aerial survey data at Neart Na Gaoithe, encounter rates of detections and individuals for each species of seabirds and marine mammals (or species group) were estimated by dividing the number of detections or individuals by the transect length. Encounter rate estimates were either global or stratified by summer/winter or inshore/offshore. Fine-scale spatial variation in encounter rate was estimated for the two most abundant cetacean species, the harbour porpoise and white-beaked dolphin. Here the encounter rates were calculated for each transect and each day. These were then combined to estimate the mean encounter rate per transect and standard errors reflect the temporal variation in encounter rate.

[47] Fine scale temporal variation in encounter rate between months was estimated for the two most abundant cetacean species. Encounter rate was estimated for each date. The monthly mean of these estimates was calculated as well as their standard error where the latter represent the degree of spatial variation for the respective month.

[5] Gordon, J. 2012. Marine Mammal Acoustic and Visual Surveys - Analysis of Neart Na Gaoithe data. Technical Report. Marine Ecological Research.

[6] 2012. Grellier and Lacey. Analysis of The Crown Estate aerial survey data for marine mammals for the Forth and Tay Offshore Wind Developers Group region. SMRUL-SGW-2012-015. Unpublished report to The FTOWDG.

[48] Density estimates were obtained by considering the innermost observed distance band (44-163m) as a strip transect with perfect detection. The other distance bands were not considered with the reasoning that the observation protocol for birds resulted in marine mammal data to which no detection function could be fitted. $\mathbf{g}(\mathbf{0})$ was not estimated, hence density estimates were negatively biased.

[49] Distance software is widely used to correct the observed counts for the animals which are missed. For analysis of boat surveys in the Neart Na Gaoithe site, Distance 6.0 was used for estimating densities of non-flying seabirds in the study area using a design-based approach. AIC based model selection was used to choose between the half-normal and hazard-rate detection model and whether or not to include covariate sea state in detection function model (using MCDS methods). These were fitted for individual species or species groups for those cases that the species did not include the recommended 60 detections for a given species. As detection on the line was likely not perfect for some species ($\mathbf{g}(\mathbf{0}) < \mathbf{1}$), the authors commented that presented density estimates might be negatively biased in these cases due to availability bias. These likely included auks, in particular the smaller species. Population size or density estimates were not adjusted.

[50] The average density of flying seabirds was estimated using the snapshot method. Total population estimates for each month were based on density estimates of both flying and non-flying birds. These were given for the study area as well as several sub-zones: the developmental area (DA) only and four sub-zones which included the DA plus increasing numbers of 1km wide buffer zones around it. Population sizes for each sub-zone were calculated as a proportion of the population of the entire area based on the ratios of observed numbers of the different sub-zones.

[51] This method was described as the Before-After-Gradient design which assumes that potential impacts decline with increasing distance from the source and that any impacts due to displacement and habitat loss will be detected on the basis of changes in the distribution of seabirds and marine mammals in these waters.[7]

[52] MCDS (combined with density surface modelling, discussed later) was also used for the Aberdeen Offshore Wind Farm data. Seasonal and monthly estimates of density and abundance were obtained for each stratum (region of the study area) using design-based inference and MCDS. Birds encountered in flight were added to the estimated abundances of non-flying birds during the line transect surveys, where the latter was adjusted for imperfect detection while the former was not. As no estimate of $\mathbf{g}(\mathbf{0})$ was available from the present surveys, imperfect detection on the line for harbour porpoises was assessed by comparing estimates for varying

[7] 2012. Fijn, R. C., Collier, M. P., Jonkvorst, R. J., Japink, M. and Poot, M. J. M. Population, density and collision rate estimates of seabirds at Neart na Gaoithe. Distance analysis of ship-based survey data and collision rate modelling using the extended Band model. Bureau Waardenburg bv. Commissioned by Mainstream Renewable Power.

values of $g(0)$.

[53] For the Moray Firth study, imperfect detection of harbour porpoises on the line was corrected using an estimate of $\mathbf{g(0)}$ for harbour porpoises from the SCANS-II aerial survey in the North Sea. Model selection for the detection function was done using MCDS methods with the potential covariates sea state, glare intensity, observer, and the overall subjective measure of sighting conditions. The best detection model by AIC for harbour porpoises included the latter two covariates. Due to insufficient sightings, a detections of all dolphin species were modelled together in a single detection function. Densities of porpoises and dolphins were estimated for both the entire survey area and the individual sub-areas separately.[8]

[54] Distance sampling was also the method of choice for Beatrice Offshore Wind Farm. For this site, ship-based survey data were analysed using design-based inference. Sightings of non-flying birds were analysed as line transect data (taking into account imperfect detection within the search area), while data of flying birds were considered strip transects (perfect detection within the search area). Similarly for aerial surveys, visual data were considered line transects while digital surveys were considered strip transects. Migration data from vantage point surveys were categorised according to direction of flight relative to the turbines. Displacement was quantified using mechanistic models to predict the impact on the seabirds using the wind farm site in the breeding season[9].

[55] Distance sampling was also used to correct observed counts for imperfect detection from the Moray Firth Development Zone. Density and abundance estimates were obtained for the development area as well as the buffer using CDS methods, using a global detection function per species with at least 60 detections. The other species were considered to be of minor concern. For those species with 60 or more detections per monthly survey, the detection function was stratified by month. Distance data were truncated at 300m and sea state and/or cluster size were used for the detection function modelling using MCDS methods.

[56] Distance sampling was used to correct for imperfect detection for the Nysted Wind Farm (in Denmark). **?** analysed the detections of long-tailed ducks made during aerial surveys at the Nysted Offshore Wind Farm (NOWF) based on a study area covering the NOWF and a large reference area surrounding it. Aerial surveys were conducted 3-4 times each winter of 2000-2005 and 2007. Line-transects were spaced 2km apart. Observers recorded all birds in three distance bands: 44-163, 163-432, 432-1000m. Total abundance of long-tailed ducks within the study area was estimated using conventional distance sampling methods. For fine scale abundance

[8] Moray Offshore Renewables. ES Technical Appendices 4.4 A Marine Mammals Baseline and 7.3 A Marine Mammals Environmental Impact Assessment. Natural Power on behalf of Moray Offshore Renewables Ltd.

[9] 2012. Beatrice offshore wind farm ornithological technical report. Final Version. RPS. This is Annex 13A to Beatrice Offshore Wind Farm Environmental Statement. Arcus Renewable Energy Consulting Ltd.
2006. Beatrice Wind Farm Demonstrator Project. Environmental Statement. Technical Report. Talisman Energy (UK) Limited.

estimates over time, line-transect data were divided into 500m segments, MCDS methods were used to model the detection function and obtain density estimates at the segments (**?**).

[57] For analyses of the Shortlist Masterplan data in the Dutch North Sea, species-specific detection functions were fitted to obtain estimates of the effective area. To correct for imperfect detection on the line resulting from the restricted view under the plane and/or from birds responding to the plane, the authors pooled observations from the first two distance bands for the aerial survey data. The effective strip width estimates from the aerial surveys of the Shortlist Masterplan project were then used to obtain density estimates for 1-minute segments along the lines. For further details, see footnotes in Section 3.1. Correction for imperfect detection under the plane was also made for porpoises in the Moray Firth aerial surveys[10,11] where data were left truncated to account for animals missed.

4.4 *Correcting observations from vantage point surveys*

[58] For vantage point surveys, some of the assumptions of distance sampling methods fail. In particular, the assumption of uniform distribution of animals in the vicinity of the point generally does not hold (e.g. **?**). If an observer conducts cetacean surveys from a land-based vantage point (such as the top of a cliff), it is easy to see that the density of cetaceans might increase with the distance from the shore (due to physical restrictions of shallow depths near shore or specific habitat preferences within the search area). The ability of the observer to detect the animals from the vantage point will likely decrease with distance. In the absence of information about the detection process models fitted to vantage point sightings are a mixture of the detection process and the underlying distribution of animals in the area.

[59] **?** recently developed the R package *nupoint* that allows estimation of density for such data where, in addition to radial distances, angles are taken for each detection. These data allows the estimation of both a detection function and the gradient in animal densities under certain conditions.

5 *Analysis of adjusted counts*

[60] After adjusting counts made either by boat or plane the second stage of the analysis involves modelling the adjusted counts at the lines/points as a function of covariates using a spatial model (e.g. a Generalised Additive Model; GAM). The benefit of this approach is that good results do not require random placement of samplers

[10] Moray Offshore Renewables Ltd Environmental Statement Technical Appendix 4.4: A Marine Mammals Baseline Telford, Stevenson and MacColl Offshore Wind Farms and Transmission Infrastructure. This document was produced by Natural Power on behalf of Moray Offshore Renewables Ltd

[11] Technical Report on pre-consent marine mammal data gathering at the MORL and BOWL wind farm sites. Paul Thompson & Kate Brookes, University of Aberdeen, Institute of Biological & Environmental Sciences,Lighthouse Field Station, Cromarty, Ross-shire IV11 8YJ. Report to MORL & BOWL, 17th June 2011.

in the study area (as the design-based approach from CDS methods described above) and allows prediction of abundance not only in the study area but also in small subareas. A further benefit of this approach is that it allows relationships between animal densities and covariates to be identified which may be of particular interest in the case that an environmental impact has occurred in a particular part of the study area. In this context, **?** fitted generalised linear models (GLM) and **?** fitted generalised linear mixed models (GLMM) to data generated from designed distance sampling experiments where particular interest was in determining the effects of a treatment applied to certain parts of the environment. The latter study included site random effects to accommodate correlations in repeat measurements at the different survey sites (which will be discussed in more detail later).

[61] However, in contrast to the design-based approach from CDS methods, using a covariate model to estimate animal numbers depends on identifying the relationship between the estimated number of animals along the lines/points and the covariates correctly. The covariates may be oceanographic data that were collected concurrently with the observations or remotely sensed data (e.g. sea-surface temperature or chlorophyll concentrations of the surface waters) available online. Other types of data that may be useful for this purpose are information on water depth, sediment type of the ocean floor and spatial co-ordinates. In the case that the survey consists of line transects, each individual line must be segmented for allocation of covariate values. An easy example to illustrate the need for this is the case where one of the covariates included in the analysis is distance from land and transects were perpendicular to the coastline.

[62] It is worth noting that while a model-based approach for analysing distance sampling data does not rely on randomised survey design, the lines or points still need to provide good coverage of the study area and an assumption has to be made that the patterns by which animals distribute themselves in the study area are observable in some form, e.g. by collecting oceanographic data during the survey. Coverage of the range of covariate values used as predictors is of concern because of dangers of extrapolation beyond the range of observed relationship between sightings and covariates.

5.1 General properties of monitoring and assessment data

[63] While the details of the data arising from marine renewables sites will be site specific, there are some common features of data collected for these purposes, either as a part of at-sea-surveys or collected from shore-based vantage points.

[64] For example, the (corrected) counts typically exhibit nonlinear relationships with environmental covariates and the distribution of the focal animals can be patchy and highly variable in number and across the site. It is also often the case that the species distribution can be patchy and highly heterogeneous in some areas of the survey area (e.g. near the coast) and relatively homogeneous in other areas (e.g. in open water). In some cases, survey sites also contain land forms and/or coastlines (e.g. exclusion zones) which animals must avoid. Because data are collected over time, consecutive counts collected along the track lines (or from vantage points) are likely to be more similar than counts measured across months or years, and thus are correlated in time. Due to some (or all) of these characteristics, the modelling approach(es) used may need to vary from site to site to obtain reliable predictions and measures of uncertainty.

5.2 Surface fitting methods

[65] A variety of approaches have been used to model monitoring and impact assessment data, however some studies relating to offshore renewables no modelling of the count data is attempted – presumably due to the low numbers of animals seen. For example, the report illustrating marine mammal data recorded from six aerial surveys of the Aberdeen Offshore Wind Farm (December 2004 – April 2007) only show raw detections[12].

[66] The basic premise of the analysis methods used and the ability of each method to address commonly encountered features of this type of data will be discussed in turn.

[12] 2013. Aberdeen Offshore Wind Farm Marine Mammals Baseline Addendum. Technical Report. Genesis. Prepared for Aberdeen Offshore Wind Farm Limited

2012. Technip UK Limited âĂŞ Aberdeen Wind Farm Ornithological Baseline and Impact Assessment Addendum. File name: J90008A-Y-RT-24000 G3-Aberdeen Wind Farm Ornithological Baseline and Impact Assessment Addendum. Technical Report. Genesis. Prepared for Technip UK Limited.

5.2.1 Generalised Additive Models (GAMs)

[67] A GAM is similar to a generalized linear model in that the relationship between the response and explanatory variables (covariates) is defined via a link function and may be implemented for a range of error distributions. Further, since the input data are typically (corrected or uncorrected) counts, either a Poisson or Quasi-Poisson distribution is assumed. For a GAM, however, the linear predictor is given by the sum of smooth functions of covariates although it may also contain linear terms and/or factor covariates. Several underlying 'basis' functions (which underpin the smoothing method) for the smooth terms have been investigated and are available in different `R` packages (**?**). One of the main differences in the smoothers that underpin these methods lies in the nature of the spatial smoothers employed. For data from offshore renewable studies, these are often used to model the spatial relationship between the number of birds/marine mammals and latitude and longitude (see Figs. 12 and 13).

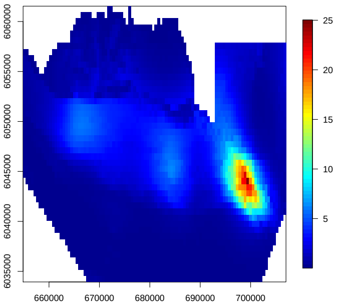

Figure 12: An example plot showing the predicted animal distribution based on a GAM analysis after the installation of a windfarm.

[68] For the *gam* function of the *mgcv* package (**?**), thin plate regression splines (TPRS) are the default smooths for the two-dimensional terms (e.g. `s(Latitude, Longitude)`), although other bases are available for this purpose (e.g. tensor products). These smoothers use Euclidean (straight-line) distances to describe the closeness of points across the survey area and it is this connectivity that dictates, in part, the fitted smooth relationship between the response (counts) and the covariates (latitude and longitude). GAMs have a single smoothing parameter which applies across the fitted surface. This results in surfaces which are equally flexible across the survey area - around the coast and out at sea. This method has been shown to work well for a wide range of smoothing problems (**?**) and is widely used to model seabird and marine mammal distributions.

[69] For example, detection of seabirds from boat surveys were modelled using GAMs in Neart Na Gaoithe[13]. Distribution maps were created using a two-dimensional smooth function inside a GAM with eastings and northings as covariates using line transects divided into 2km segments. GAMs were also the analysis method of choice for marine mammals in Neart Na Gaoithe. Specifically, presence-absence models (using binary input data) were constructed combining visual and acoustic data. Acoustic detections were always counted as '1' and for visual detections, if the group size was larger than one, then these were 'smeared' across adjacent segments to reflect the observed group size. This approach was used to allow groups greater than one to contribute to the observed number of animals while still maintaining the same binary modelling approach as used for the acoustic data. This is a non-standard approach to combining acoustic and visual data and the

[13] Barton, C., Jackson, D., Bloor, P. and Crutchfield, Z. Using a before-after-gradient design to determine post-construction effects of an offshore wind farm on birds. Preliminary results. Cork Ecology, Natural Research Projects Ltd, Pelagica, Mainstream Renewable Power.

reliability of the results is unclear.

[70] GAMs were also the surface fitting model of choice for the Moray Firth Development Zone. For some bird species (fulmar, gannet, kittiwake, guillemot, razorbill, puffin) GAMs were used to produce density surface maps based on transects divided into 600m segments. In this example, depth, sediment type, month, latitude and longitude were all used as model covariates. Additionally, GAMs were used to produce distribution maps of animal abundance throughout the Moray Firth survey area using latitude, longitude, depth, distances to shore and to each of the three nearby SPAs.

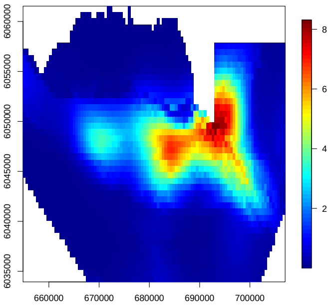

Figure 13: An example plot showing the predicted animal distribution based on a GAM analysis after the installation of two windfarms.

[71] While useful and widely applied, GAMs assume that the patterns in the input data are correctly modelled using the model for the mean (containing a chosen set of model covariates) and the noise component is therefore patternless and independent. This may be unreasonable in many cases. The data collected along transects are likely to be correlated in time and/or space. While the available covariates associated with these segments may also be correlated in time/space , a fitted model including these covariates is still unlikely to explain this correlation in full. For instance, some covariates that contribute to the correlation in animal numbers along the segments of a given transect may be absent from the model and thus, the pattern in the count data that originated from these omitted relationships is necessarily allocated to the errors.

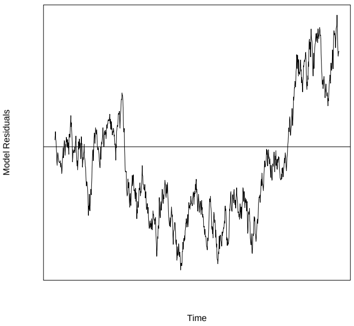

Figure 14: An example plot showing patterns in model residuals.

[72] This results in correlation patterns in the model residuals (Figure 14) which is a violation of an important model assumption: independence of residuals. Specifically, residuals tend to be positively correlated (evidence by long uninterrupted sequences of negative and positive residuals) that tends to result in estimates of precision (and associated p-values) that are too small. This can mean that non-significant covariates are falsely reported to be statistically significant, which is of particular concern if the change in average numbers before and after impact (via an "impact" term) is explicitly tested in a model. Here, natural fluctuations in animal numbers might be mistaken for a genuine change in average numbers before and after impact, and potentially attributed to the "impact." There are simple graphical and numerical methods to test for non-randomness in model residuals (e.g. auto-correlation function (acf) plots and the Runs test, **?**) however, these do not currently appear to be standard diagnostic tools for models of spatial impact assessment studies.

[73] The concerns about using GAMs for modelling correlated data are sometimes recognized in the associated reports. For example, the authors of the GAM based analysis in the Moray Firth development zone noted the failing of this approach to incorporate autocorrelation in the line transect data and the potential problems resulting from this. However, in this case no alternative methods which address the residual correlation were presented.

[74] GAMs can also be sub-optimal for modelling survey data with complex topography and internal exclusion zones (e.g. islands and coastline). Land forms necessarily pose an obstacle for marine mammals and must be travelled around while some seabirds will also avoid flying over land (e.g. terns and auks). In smoother based methods, defining the distance appropriately between points can be key for realistic surfaces. Using Euclidean distance based smoothing methods it is possible to predict large animal numbers on one side of an island where no animals were ever seen, simply because high numbers were seen on the other side. In this case, the large numbers effectively "leak" across the island due to the apparent, yet mistaken, closeness in space.

[75] Methods for smoothing over complex regions which respect boundaries (such as the outline of an island that cannot be traversed by a whale) have been proposed when fitting two dimensional smooths. Soap film smoothing (SOAP) (**?**) is one such package and can be implemented with the *gam* function of the *mgcv* package by setting the argument *bs* = *"so"* within the definition of the smooth terms *s* or with the *soap* package. This is a relatively new package which requires considerable user input regarding aspects of model selection. There are other modelling alternatives for survey areas with complex topography, and one of these (Complex Region Spatial Smoother; CReSS) which also has different smoothing properties will also be mentioned here.

[76] GAMs also typically assume that model flexibility is the same across the surface/survey area. Specifically, the `gam` function has a global smoothing parameter that applies across the surface, and this is often estimated from the data using Generalized Cross Validation (GCV). While simple, this can make it difficult to model surfaces which are rapidly changing in some areas (e.g. near the coast) and yet are relatively smooth in others (e.g. in open ocean/deep water). There are other methods employed for monitoring and impact assessment data which allow an uneven smoothness across the surface and two of these "spatially-adaptive" methods are discussed here: Kriging and CReSS.

5.2.2 Complex Region Spatial Smoother (CReSS)

[77] CReSS is a smoothing method which is based on geodesic distances (as opposed to Euclidean distances) which more closely approximate "as-the-animal-swims." This method explicitly constrains connections between points to be within the domain (which one may roughly think of as the path the whale would have to take around the island), allowing more accurate distances to be determined, even for sparse data sets.

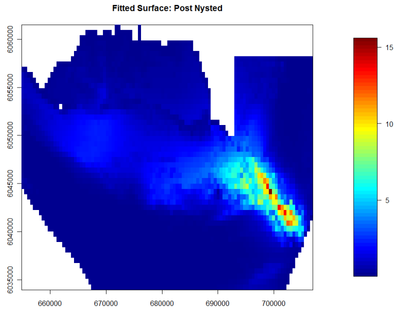

Figure 15: An example plot showing the predicted animal distribution based on a CReSS analysis after the installation of a windfarm.

[78] CReSS has the added advantage of employing locally varying radial basis functions in a GAM framework allowing some areas of the surface to be more flexible than others (**?**). This is achieved by allowing model flexibility to be targetted across the surface (using a model selection algorithm, spatially adaptive local smoothing algorithm SALSA), resulting in a surface which is "spatially adaptive." SALSA was developed for one dimensional smoothing applications (**?**) but has recently been extended to operate with two dimensional smoothers. This method allocates even flexibility across the surface as a starting point (similar to a GAM) but can target (and reallocate) model flexibility if this results in improved model fit (as determined using objective fit criteria).

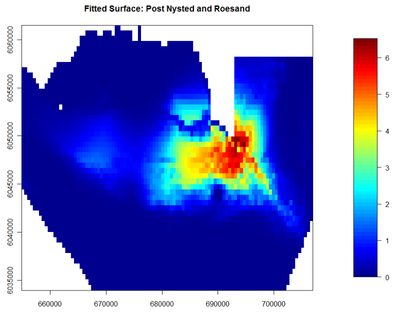

Figure 16: An example plot showing the predicted animal distribution based on a CReSS analysis after the installation of two windfarms.

[79] CReSS is a method for modelling one and two dimensional covariates in a flexible way and thus, can be fitted using a variety of modelling frameworks. This means that estimation and inference procedures which also permit auto-correlated residuals can be employed. Generalized Estimating Equations (GEEs) can be used to generate fitted values (Figs. 15 and 16) and associated measures of precision when faced with autocorrelation in model residuals along transects (page 23).

[80] CReSS-based analysis (with SALSA-based model selection) coupled with GEEs (to account for correlated residuals) has been used to model Kyle Rhea tidal data[14], tern distribution for many sites across the UK: the Orkney Isles, the Pentland Firth, the Solent and in and around the Isle of Wight. CReSS/SALSA has also be used to model cetacean distribution for the Joint Cetacean Protocol project (JCP[15]) which covers approximately one million square kilometres over several years.

[81] CReSS (based on a quasi-Poisson error distribution) was also used to model local densities for the Nysted Off-Shore Wind Farm.[16] This model included a one-dimensional smooth term for depth and a two-dimensional term for the spatial coordinates. These smooth terms were allowed to differ in shape and magnitude before and after construction by including construction-based interaction terms. The log of the search area was included as an

[14] The Kyle Rhea Tidal Stream Array. Volume III. Appendices prepared by SeaGeneration (Kyle Rhea) Ltd. (no year shown on report). Statistical analysis conducted by DMP Statistical Solutions UK Ltd.

[15] http://jncc.defra.gov.uk/page-5657

[16] Petersen, I. K., Mackenzie, M. L., Rexstad, E., Wisz, M. S., and Fox, A. D. 2011. Comparing pre- and post-construction distributions of long-tailed ducks *Clangula hyemalis* in and around the Nysted offshore wind farm, Denmark: a quasi-designed experiment accounting for imperfect detection, local surface features and autocorrelation. Technical report, Centre for Research into Ecological and Environmental Modelling. Technical report 2011-01.

offset. The smooth terms used were cubic B-splines for depth and exponential radial basis functions with a fixed number of knots for the two-dimensional smooth.

5.2.3 Generalised estimating equations (GEEs)

[82] GEEs (**?**) allow modelling of correlated responses for different types of input data (e.g. binary and count data). GEEs (as they are typically implemented) are population averaged models and model how the average response in the population changes with modelled covariates. However, GEEs extend the GLM framework by allowing the noise component to be correlated within panels/subjects. Estimates of precision (such as standard errors and significance tests) can either be based on a chosen correlation structure within-subjects, or "so-called" robust empirical, estimates of variance can be used instead which use the residuals directly within-subjects to generate estimates of precision. Using the latter approach to account for residual correlation allows for unbiased variance estimation despite possible misspecification of the correlation structure (**?**).

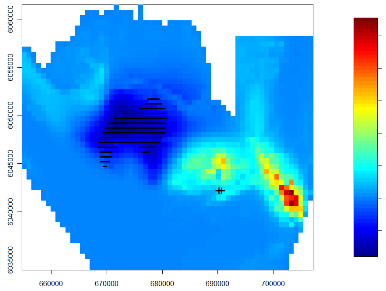

Figure 17: An example plot showing statistically significant pre-post declines (-) and increases (+) based on GEE-based confidence intervals for pre-post differences.

[83] GEEs may be fitted in R using the *geeglm* function of the *geepack* package (**?**) or inside SAS using the GENMOD procedure (**?**). GEEs have been used as inferential tools for predicting cetacean distributions (**??**) and are being used for many types of surveys. Recent examples include using a spatially adaptive model for the mean coupled with GEEs for the JCP analyses, and the Nysted Wind Farm (**?**).

[84] In the Nysted case, 95% confidence intervals for model-based differences before and after construction were produced using GEEs where the "panels" were set to be transect-days, i.e. residuals within transects on a given day were allowed to be correlated, while those from different transects on the same day or those from the same transects on different days were considered independent. Uncertainties from the two modelling stages (fitting a detection function and the CReSS-GEE model) were combined by fitting both stages to resampled data during 500 bootstrap iterations (non-parametric bootstrapping). Parametric bootstrapping (based on GEE estimates of precision) was conducted using each of the 500 CReSS-GEE based estimates and standard errors from the non-parametric bootstrapping for 1000 parametric bootstrap realisations. For each set of predictions, the estimated difference in long-tailed duck abundances before and after construction across the grid was found in order to generate 95% confidence intervals (Fig. 17).

[85] High quality model predictions based on GEE analysis still require a realistic model for the mean. Since covariate relationships are often nonlinear (even on the link scale) using off-the-shelf functions still requires some kind of manual implementation of smoother based functions for one-dimensional and two dimensional smooths. This may be preventing their widespread use. GEEs can also give unreliable estimates of variance if the number of correlated blocks are small - however this is not typically the case for survey data from offshore renewables projects.

[86] GEEs are population-averaged models which accommodate the correlation within blocks explicitly, however there are other ways to adjust geo-referenced confidence intervals for this correlation. Computationally intensive methods are also available for this purpose, such as the non-parametric boostrap and parametric moving-block bootstrap methods which are based on resampling independent blocks (e.g. transects); these were methods of choice for the Aberdeen Offshore Wind Farm and for the Moray Firth development Zone, [17] for example. Mixed-effects models (e.g. GAMMs) are a further way to model the correlation in the data. These "conditional" models effectively induce the (necessarily positive) correlation within blocks by allowing one or more coefficients in the model to vary across correlated blocks.

5.2.4 Generalised additive mixed models (GAMMs)

[87] GAMMs extend GAMs by including the same set of potential functions for describing the response-covariate relationship (smooths, linear and factor variables) but also include "random-effect" terms to model correlation between observations within panels/subjects. These random effects, for which normality is generally assumed, may consist of a random intercept and/or random slopes. If repeat surveys were taken on the same transect, a random intercept may be fitted to accommodate correlations of measurements within transects. For a generalised additive mixed model using the *gamm* function of the *mgcv* package, the smooth functions are specified in the same manner as in a GAM, however, the components defining the flexibility of the smooths are treated as random effects (**?**).

[88] GAMMs have been used to model monitoring and assessment data. For the Aberdeen offshore wind farm, survey effort was split into 1km segments for modelling purposes and both GAMs and GAMMs were fitting based on depth, season, distances from the coast/harbour as candidate variables.

[89] GAMMs were also used to model data from the Moray Firth. Habitat association modelling was performed for harbour porpoises where the habitat data were split into 4×4 km^2 grid cells. The covariates considered for the analyses were depth, sediment type,

[17] Moray Offshore Renewables. Environmental Statement. Technical Appendix 4.5 B - Aerial Ornithology Surveys for the Moray Firth Zone, Summer 2011. Natural Power on behalf of Moray Offshore Renewables Ltd.

slope and distance to coast, however only depth and sediment type (the proportion of the cell in sand and gravelly sand sediments) were selected. The survey data (boat and aerial surveys combined) was also divided into 4×4 km^2 grid cells. If repeat surveys covered the same grid cells, these were considered separately in the analysis. If cells included only 1km of effort, these were not included. A total of 429 cell observations were analysed representing repeat observations from 241 unique cells. Survey type was also considered as a candidate covariate to account for potential differences in sighting rates, but its inclusion increased the objective fit score (AIC) and so it was not retained in the final model.

[90] Generalised additive mixed models were fitted using the *mgcv* package in R. A random effect for cell ID was included to account for correlated measurements of the same grid cells. The log of effort expended within the grid cell was included as an offset. A negative-binomial error structure was assumed as initial models were found to be overdispersed. Model selection was done using AIC. Over 1000 porpoise sightings were used in the habitat association models. Distribution maps and error maps were produced for the larger Moray Firth region.

[91] The quality of GAMM-based predictions (for non-normal data) rely on the validity of model assumptions and these can be particularly difficult to check. Specifically, both model coefficients and associated measures of precision can be biased if the model is misspecified. In particular, GAMMs require the specification of random effects which describe how baseline levels and/or covariate relationships vary across correlated blocks/individuals in the population. Incorrect specification of how these terms vary across correlated blocks can result in biased model predictions and unrealistic measures of uncertainty.

[92] GAMMs are also so-called "conditional models" which effectively model the average correlated block – blocks which are defined by the user. Typically, baseline monitoring and impact analyses are concerned with predicting average numbers in the population which is the focus of so-called "population averaged" (or marginal) models. Marginal results can be obtained using GAMMs, but the accuracy of these population averaged results depends on the quality of model assumptions.

5.2.5 Kriging

[93] Kriging is an interpolation technique that fits smooth surfaces throughout irregularly spaced data. To make predictions in the locations where no observations were made, the information from the observed data points is used taking their closeness into account. In contrast to GLMs, GAMs and GEEs, ordinary kriging does not rely on explanatory covariates, however, in some cases kriging models may be improved by incorporating covariates.

[94] GAM methods can generalise to be kriging methods under certain conditions (and vice versa) but the fundamental difference in methods is that kriging models the current process (what is there at the time) and GAM methods model the underlying process (the population mean). Furthermore, kriging was designed for exact measurements taken at precise locations, such as heavy metal concentrations at well sites. Wildlife survey data are inexact and tend to cover an area rather than a precise location.

[95] Regression-kriging is a technique that combines a deterministic (regression) model to estimate the global trend, with a stochastic component (**?**). The overall predictor is the sum of the two components. The regression model may contain covariates and associated coefficients. The stochastic component represents the interpolated residuals which are determined by kriging weights and the residuals of the observed data points from the global trend. This method accommodates spatial autocorrelation; in case no spatial autocorrelation is present, regression-kriging is equivalent to pure regression. The only way to appropriately model the variation in this combined process is by stochastic simulation. To the best of our knowledge there is no simple one-step package to achieve this within `R` or `SAS`.

[96] Regression kriging appears in the literature. In the Dutch North Sea (as a part of the Shortlist Masterplan Project) estimated bird densities were extrapolated into the areas of no survey effort between the lines using regression-kriging to account for spatial autocorrelation in the data. GLMs were constructed using depth and distance from the coast as covariates. Predictions throughout the study area were made from with the GLM for which the predictions were adjusted by kriging the residuals. For the latter, normalised model residuals were tested for spatial autocorrelation though a variogram analysis. Block mean kriging predictions were estimated for 5 km × 5 km squares. Similarly for the MWTL surveys, the covered area was used to obtain density estimates for the 1-minute segments and the equivalent regression-kriging method applied[18]

6 *References*

---

[18] Poot, M. J. M. and Fijn, R. C. and Jonkvorst, R. J. and Heunks, C. and Collier, M. P. and de Jong, J. and van Horssen, P. W. 2011. Aerial surveys of seabirds in the Dutch North Sea May 2010 âĂŞ April 2011. Seabird distribution in relation to future offshore wind farms. Technical Report. Bureau Waardenburg bv. Commissioned by IMARES.