**Area-based versus Individual-based perspectives on bird distributions**

Dedicated surveys to sample the abundance and distribution of seabirds are a basic requirement to understand bird densities within coastal and offshore marine environments. Distribution data can be obtained from observations collected during surveys within a predetermined spatio-temporal frame of reference or by sampling locations of individual animals using animal‐borne tracking devices. The former provide an area-based sampling perspective, also referred to as a Eulerian perspective in the oceanographic literature, whereas the latter provides an individual based sampling perspective, also referred to as a Langragian perspective (Phillips et al. 2019).

Area-based or Eulerian survey designs sample across a predetermined survey area using at predetermined sampling locations or transects, which may or may not be replicated through time. The primary objective of this sampling approach is to obtain animal distribution and abundance data in a predefined area and time period. For seabirds at-sea such surveys are generally conducted using ships or aircraft (Tasker et al. 1984; Camphuysen et al. 2004; Buckland et al. 2012), although land-based surveys may be used to cover near-shore environments (Arranz et al. 2014; Smith et al. 2015).

Individual-based or Lagrangian survey designs track the locations of seabirds through space and time using data logging or tracking devices attached to individuals (Burger and Shaffer 2008; Hays et al. 2019). Location data may be received in near real-time from transmitting or satellite-linked devices, or obtained by downloading stored location data using base-stations or after device retrieval. Depending on the mobility of a species, individual-based sampling may increase the spatial extent and resolution of the survey area compared to area-based surveys, or even identify colony location and/or at-sea habitats if no a priori information about species movements are available (e.g. Rayner et al. 2015; Bolton 2021).

Modern tracking methods provide high‐resolution data (generally with sub-km precision) at an individual level, but device cost and the logistical challenges of tag deployment usually limit the number of tags deployed, as well as the number of locations at which tags are deployed. So while the spatial extent of survey coverage achieved by animal-borne devices is determined by individual’s movements, and not a priori defined by the surveyor, achieved coverage of the wider marine environment will generally not be independent of the locations at which tracking devices are deployed. Tag deployment further depends on the accessibility and/or catchability of individuals, therefore tracking data sets often exhibit individual-level heterogeneity, which may be related to sex, age, breeding status and/or colony-affiliation (Gutowsky et al. 2015; Krietsch et al. 2017) and may complicate population‐level inferences.

Due to the differing strengths and weaknesses of the two methods, distributions derived from both may be combined to provide a more complete picture of seabird distribution (e.g. Louzao et al. 2009), but formal analytical frameworks for combining area-based and individual-based data remain an active research topic in ecological statistics (Glennie et al. 2021; Chandler et al. 2022).

Synoptic comparisons of both approaches are rare, but the results of Carroll et al. (2019) and Phillips et al. (2019) suggest that tracking data can yield comparable results to area-based survey data when the tracked individuals are a representative sample of the birds using the area of interest. In particular, single-colony tracking appears to yield comparable results to area-based approaches for relatively isolated colonies, whereas multi-colony tracking or tracking of individuals captured at-sea in the area of interest is recommended to get an unbiased picture of seabird distributions in areas used by individuals from dispersed colonies. Assessments of the comparability of sampling approaches therefore remain critical to the interpretation of differences in seabird distribution, when data have been collected using different methods.

**At-sea survey methodology**

In this section we briefly review existing census techniques for area-based or Eulerian surveys of seabirds at sea. Two primary observation tools are discussed in this study, aerial and ship-based line transect surveys, which both have the potential to provide similar data outputs, i.e. spatially referenced counts of birds. Additional observation methods exist, such as land-based human observers (Arranz et al. 2014; Smith et al. 2015) or terrestrial or marine radars (Lilliendahl et al. 2003; Orben et al. 2019) but are not discussed in detail due to their inability to monitor large offshore areas, and or limitations with regard to species identification. Census techniques generally rely on a common principle: line transect surveys, but different survey methodologies have strength and weaknesses both concerning the nature and quality of the collected data, and in terms of their logistics. Data quality is primarily affected by the level of detectability of individual birds a method affords, and the level of taxonomic identification of detected birds. Logistical characteristics include differing costs, survey speeds, and allowable environmental conditions under which surveys can be undertaken. Aerial surveys are quick, enabling coverage of larger areas per unit time and therefore providing a more synoptic view of seabird distributions than slower shipboard surveys. Bird data obtained during aerial surveys may be combined with remotely sensed environmental parameters in a correlative approach, but in-situ collection of environmental parameters is usually very limited, whereas shipboard surveys can allow the simultaneous collection of in-situ oceanographic data. Both aerial and shipboard surveys generally rely on visual survey methods (human observers, camera systems; Buckland et al. 2012), which restricts them to daylight hours, and are generally limited to relatively benign wind and wave conditions (generally sea states below 3-5). This may fundamentally limit the understanding of realized seabird distributions under the full set of prevailing environmental conditions. Survey platforms further differ in their effect on seabird behaviour. Surface vessels may elicit avoidance or attraction effects on birds (Spear et al. 2004), whereas the degree of avoidance to aircraft is largely governed by their flight height, so that surveys may have little to no impact on bird behaviour when executed at a sufficient aircraft altitude (Buckland et al. 2012).

**Modelling**

High-level summary

The distribution of seabirds at sea has been difficult to study due to the lack of human observers in the marine environment. Besides the use of tracking devices to follow individuals, researchers have used a range of sophisticated modelling techniques to predict where seabirds occur at sea. These species distribution models are widespread to predict the occurrence of plants and animals on land, and have been adapted to the marine environment for at least three decades. This toolkit will not provide any guidance or overview how to implement species distribution models for seabirds, but we will mention these options as potential alternatives to the identification of important sites based on tracking data.

Species distribution models generally work by relating some observations about the presence of a species to environmental variables, and then using maps of these variables and the estimated relationships to project where a species may occur. For seabirds at sea, the observations are generally provided by either direct observations from a vessel (research surveys, opportunistic sightings, museum records etc.), or from tracking data. These observations are then combined with either sampled or random locations where the target species was *not* observed or recorded, and then related to static (depth, latitude, longitude, distance from coast, seamounts etc.) or dynamic (sea temperature, salinity, chlorophyll concentration, wind, wave height etc.) environmental variables. A large range of different model types are available to construct species distribution models, which vary in their data requirements, assumptions, and the way how the relationships between observations and environmental variables are related. Species distribution models generally return spatially explicit predictions of the habitat suitability or probability of occurrence of the target species, which can be shown on a map to visualise the projected distribution of a species. These maps can then be used to inform the identification and protection of important sites at sea in the same way as the maps created in this toolkit. We do not recommend to propose IBAs or KBAs solely based on prediction species distributions, but predicted distributions can be useful for (1) strategically allocating survey effort to areas where there is a high probability of occurrence; (2) delimit site boundaries for sites that have been confirmed as important by other data sources; and (3) extrapolate the number of individuals potentially using a site based on environment-abundance relationships.

Detailed section:

Predictive models are based on the principles that equations and rule-sets can be constructed to represent how a species’ distribution is related to environmental conditions ([Aarts 2012,](https://besjournals.onlinelibrary.wiley.com/doi/10.1111/j.2041-210X.2011.00141.x) [Robinson 2017](https://www.frontiersin.org/articles/10.3389/fmars.2017.00421/full)). Species distribution models (also often referred to as ‘habitat suitability models’ or ‘resource selection functions’) for seabirds have been used for >20 years and were historically based on observations obtained from vessels. More recently, species distribution models are based on the locations that tracked individuals of a species used at sea, which resulted in a much broader distribution of ‘presence observations’ over a wider range of the marine realm than had been possible using vessel-based observation data ([Matthiopolous et al 2022](https://besjournals.onlinelibrary.wiley.com/doi/full/10.1111/1365-2664.14114)). In addition, models are now frequently used to not only predict where a species occurs, but where specific behaviours occur, e.g. where seabirds are foraging rather than just occur in transit ([Boyd et al. 2015](https://besjournals.onlinelibrary.wiley.com/doi/10.1111/1365-2656.12409)).

Species distribution models rely on contrasting environmental conditions between locations where a species or a particular behaviour is observed, and where it is *not* observed to infer relationships between environment and occurrence. However, for highly mobile species like seabirds, defining and interpreting what ‘not observed’ actually means can be very difficult and affect the quality and interpretability of species distribution models. In contrast to sessile plants, it is virtually impossible to objectively determine whether a seabird species is truly ‘absent’ (= never occurs) at a given point in the sea – the fact that no birds of that species were seen during a research cruise at that location, or that none of the tracked individuals visited that location, does not exclude that other individuals of the same species may occur there at other times. Many seabird distribution models therefore rely on ‘presence-only’ data, which are the least robust form of data that cannot differentiate between the limit of a species’ distribution and the limit of the sampling effort.

One of the key decisions when modelling the distribution of seabirds at sea is therefore to select the right contrast and appropriate data to model the contrast ([Matthiopolous et al 2022](https://besjournals.onlinelibrary.wiley.com/doi/full/10.1111/1365-2664.14114)). If survey effort is sufficient to use non-detection data as contrast to locations where a species was observed, then so-called detection – non-detection models can be fitted which generally permit stronger inference about the probability of a species to occur at a certain location. For tracking data, this can be the case if only certain behaviours are modelled (where the occurrence of ‘foraging’ can be contrasted with all other behaviours), but more frequently the observations will be contrasted with a set of random points where no information exists whether the species did or did not occur there. These random points can be created based on virtual tracks that mimic the movements of seabirds ([Zydelis](https://royalsocietypublishing.org/doi/full/10.1098/rspb.2011.0330)). Presence-only data can be combined with other data sources to inform the underlying distribution and the observation process ([Matthiopolous et al 2022](https://besjournals.onlinelibrary.wiley.com/doi/full/10.1111/1365-2664.14114)), but they need to be interpreted with caution given that there are natural limits of inference ([Hastie 2013](https://onlinelibrary.wiley.com/doi/full/10.1111/j.1600-0587.2013.00321.x)).

The spatial and environmental extent and variability at random locations against which observations are contrasted will have a great influence on the model and the inferred relationships between the presence of a species or behaviour and the environment. The scale at which inference is sought (e.g. do we want to know where species X occurs in the world, or which bay they prefer for foraging from colony Y) is critical in guiding the environmental data and the selection of background or non-detection points to ensure that the model can yield information at the desired scale of inference.

Popular and widely used approaches to predict seabird distributions or behaviours at sea are Generalised Linear Models (GLM; [Huettmann 2001](https://www.sciencedirect.com/science/article/pii/S0304380001002782)), Generalised Additive Models (GAM; [Zydelis 2011](https://royalsocietypublishing.org/doi/full/10.1098/rspb.2011.0330); [Warwick-Evans 2022,](https://esajournals.onlinelibrary.wiley.com/doi/full/10.1002/ecs2.4083) [Critchley](https://onlinelibrary.wiley.com/doi/full/10.1111/ecog.04653)), mixed-effects implementations of GLM and GAM that allow for the serial dependence of observations through the incorporation of random effects (GLMM or GAMM; [Waggitt,](https://besjournals.onlinelibrary.wiley.com/doi/full/10.1111/1365-2664.12646) [Weber](https://www.pnas.org/doi/abs/10.1073/pnas.2101325118), [Gilman,](https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0084499) [Chimienti](https://onlinelibrary.wiley.com/doi/full/10.1002/ece3.3551)), Maximum Entropy ([Kruger,](https://zslpublications.onlinelibrary.wiley.com/doi/abs/10.1111/acv.12339) [Lemos et al. 2023,](https://www.sciencedirect.com/science/article/pii/S2351989422003626) [Hodges](https://besjournals.onlinelibrary.wiley.com/doi/full/10.1002/2688-8319.12181)), Boosted Regression Trees ([Evans](https://esajournals.onlinelibrary.wiley.com/doi/full/10.1002/eap.2343), [Humphries](https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0137241), [Torres](https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0120014)), Random Forests ([Huettmann et al. 2011](https://link.springer.com/article/10.1007/s12526-011-0083-2), [Diop,](https://www.int-res.com/abstracts/meps/v607/p221-236/) [Mikami,](https://www.sciencedirect.com/science/article/pii/S0301479722007939) [Boyd](https://besjournals.onlinelibrary.wiley.com/doi/full/10.1111/1365-2664.12832)), point process models ([Wakefield](https://esajournals.onlinelibrary.wiley.com/doi/full/10.1002/eap.1591), [Renner](https://besjournals.onlinelibrary.wiley.com/doi/full/10.1111/2041-210X.12352)), and ensembles of multiple models with average predictions based on the performance of each model ([Fox](https://www.int-res.com/abstracts/meps/v566/p199-216/), [Scales](https://onlinelibrary.wiley.com/doi/full/10.1111/ddi.12389), [Hakkinnen,](https://onlinelibrary.wiley.com/doi/full/10.1002/ece3.8272) [Renner](https://www.int-res.com/abstracts/meps/v484/p259-277/), [Lieske](https://www.sciencedirect.com/science/article/pii/S0006320714000627), [Lavers](https://conbio.onlinelibrary.wiley.com/doi/full/10.1111/cobi.12324), [Oppel](https://www.sciencedirect.com/science/article/pii/S0006320711004319), [Pereira et al. 2018](https://www.sciencedirect.com/science/article/pii/S0964569117306701)). The choice of the algorithm and how the contrast is selected will have a greater effect on the resulting model predictions than the spatial and temporal resolution of the data used to train the model ([Quillfeldt 2017](https://onlinelibrary.wiley.com/doi/full/10.1111/jav.01238)).

If tracking data are considered as a presence-only dataset for the purpose of a distribution model, it is important to consider a species’ ecology and the spatial and temporal resolution of the tracking data to avoid the inclusion of sections of the track where the individual is not actively ‘using’ the associated environment, but is just passing through. A range of bespoke analytical techniques are available to identify different behaviours from tracking data (e.g. state-space models, hidden markov models, expectation-maximisation binary clustering etc.), and each of these methods comes with its associated assumptions and limitations which we will not elaborate here. Useful overviews of what methods can be used to identify behaviour are here ([Bennison](https://onlinelibrary.wiley.com/doi/10.1002/ece3.3593), [McClintock](https://besjournals.onlinelibrary.wiley.com/doi/full/10.1111/2041-210X.12995), [Patterson,](https://onlinelibrary.wiley.com/doi/full/10.1002/ece3.4740) [Browning](https://besjournals.onlinelibrary.wiley.com/doi/full/10.1111/2041-210X.12926), [Garriga](https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0151984)).

Species distribution models are generally evaluated with independent test data – observations that were not used to develop the model, but are then used to evaluate how accurate the predictions of the model are. Depending on how far away in space and environmental conditions the test data are from the data that were used to build the model, the resulting predictions generally become less accurate the farther away the test data are from the training data. Seabird species distribution models have so far not transferred very well across different regions ([Torres](https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0151984), [Diop](https://www.int-res.com/abstracts/meps/v607/p221-236/)), indicating that explaining the distribution of a species in one region will not necessarily allow for an accurate prediction where that species may occur in another region even if similar environmental data are available. We therefore do not recommend to use predicted distributions that are based on models which were constructed without any input data from the target species population.

Predicted species distributions can be used in conjunction with other outputs of this toolkit, or across multiple species or seasons to prioritise marine areas based on systematic spatial planning approaches. Popular algorithms such as [Zonation](https://cran.r-project.org/web/packages/zonator/index.html) and [Marxan](https://prioritizr.net/) allow users to overlay predicted distribution maps for several species (or the same species in several seasons) and then determine the areas at sea that are most valuable to protecting all species (or a species in all seasons) based on algorithms that trade off the size of the area that needs to be protected with the minimum amount of habitat that needs to be protected for each species or season ([Dias](https://www.sciencedirect.com/science/article/pii/S0006320716309934), [Heerah](https://www.sciencedirect.com/science/article/pii/S0006320718315994), [Oppel](https://www.sciencedirect.com/science/article/pii/S0006320711004319)). Alternatively, hotspot identification can be achieved mathematically with metrics such as Getis-Ord and Maximum Curvature ([Requena](https://zslpublications.onlinelibrary.wiley.com/doi/full/10.1111/acv.12572), [Cleasby](https://www.sciencedirect.com/science/article/pii/S0006320719310419)). Such spatial prioritisation algorithms can then be combined with IBA criteria to examine whether the priority areas meet international criteria for IBA or KBA.

Random citations (I inserted hyperlinks instead above for many citations):

(Cleasby *et al.* 2020)

(Wakefield *et al.* 2017, Matthiopoulos *et al.* 2022)

(Torres *et al.* 2015)

(Oppel *et al.* 2012, Häkkinen *et al.* 2021)

(Engler *et al.* 2017, Devarajan *et al.* 2020)

(Zipkin *et al.* 2015)

(Boyd *et al.* 2015)

(Lieske *et al.* 2014, Smith *et al.* 2014)

Boyd, C., Castillo, R., Hunt, G. L., Punt, A. E., VanBlaricom, G. R., Weimerskirch, H., and Bertrand, S. (2015) Predictive modelling of habitat selection by marine predators with respect to the abundance and depth distribution of pelagic prey. *J. Anim. Ecol.* 84:1575-1588.

Cleasby, I. R., Owen, E., Wilson, L., Wakefield, E. D., O'Connell, P., and Bolton, M. (2020) Identifying important at-sea areas for seabirds using species distribution models and hotspot mapping. *Biol. Conserv.* 241:108375.

Devarajan, K., Morelli, T. L., and Tenan, S. (2020) Multi-species occupancy models: review, roadmap, and recommendations. *Ecography* 43:1612-1624.

Engler, J. O., Stiels, D., Schidelko, K., Strubbe, D., Quillfeldt, P., and Brambilla, M. (2017) Avian SDMs: current state, challenges, and opportunities. *J. Avian Biol.* 48:1483-1504.

Häkkinen, H., Petrovan, S. O., Sutherland, W. J., and Pettorelli, N. (2021) Terrestrial or marine species distribution model: Why not both? A case study with seabirds. *Ecology and Evolution* 11:16634-16646.

Lieske, D. J., Fifield, D. A., and Gjerdrum, C. (2014) Maps, models, and marine vulnerability: Assessing the community distribution of seabirds at-sea. *Biol. Conserv.* 172:15-28.

Matthiopoulos, J., Wakefield, E., Jeglinski, J. W. E., Furness, R. W., Trinder, M., Tyler, G., Mccluskie, A., Allen, S., Braithwaite, J., and Evans, T. (2022) Integrated modelling of seabird-habitat associations from multi-platform data: A review. *J. Appl. Ecol.* 59:909-920.

Oppel, S., Meirinho, A., Ramírez, I., Gardner, B., O’Connell, A., Miller, P. I., and Louzao, M. (2012) Comparison of five modelling techniques to predict the spatial distribution and abundance of seabirds. *Biol. Conserv.* 156:94-104.

Smith, M. A., Walker, N. J., Free, C. M., Kirchhoff, M. J., Drew, G. S., Warnock, N., and Stenhouse, I. J. (2014) Identifying marine Important Bird Areas using at-sea survey data. *Biol. Conserv.* 172:180-189.

Torres, L. G., Sutton, P. J. H., Thompson, D. R., Delord, K., Weimerskirch, H., Sagar, P. M., Sommer, E., Dilley, B. J., Ryan, P. G., and Phillips, R. A. (2015) Poor transferability of species distribution models for a pelagic predator, the Grey Petrel, indicates contrasting habitat preferences across ocean basins. *PLoS ONE* 10:e0120014.

Wakefield, E. D., Owen, E., Baer, J., Carroll, M. J., Daunt, F., Dodd, S. G., Green, J. A., Guilford, T., Mavor, R. A., Miller, P. I., Newell, M. A., Newton, S. F., Robertson, G. S., Shoji, A., Soanes, L. M., Votier, S. C., Wanless, S., and Bolton, M. (2017) Breeding density, fine-scale tracking and large-scale modeling reveal the regional distribution of four seabird species. *Ecol. Appl.* 27:2074–2091.

Zipkin, E. F., Kinlan, B. P., Sussman, A., Rypkema, D., Wimer, M., and O'Connell, A. F. (2015) Statistical guidelines for assessing marine avian hotspots and coldspots: A case study on wind energy development in the U.S. Atlantic Ocean. *Biol. Conserv.* 191:216-223.

Arranz, Patricia, David L Borchers, Natacha Aguilar de Soto, Mark P Johnson, and Martin J Cox. 2014. “A New Method to Study Inshore Whale Cue Distribution from Land-Based Observations.” Marine Mammal Science 30 (2): 810–18.

Bolton, Mark. 2021. “GPS Tracking Reveals Highly Consistent Use of Restricted Foraging Areas by European Storm-Petrels Hydrobates Pelagicus Breeding at the Largest UK Colony: Implications for Conservation Management.” Bird Conservation International 31 (1): 35–52.

Buckland, Stephen T, M Louise Burt, Eric A Rexstad, Matt Mellor, Adrian E Williams, and Rebecca Woodward. 2012. “Aerial Surveys of Seabirds: The Advent of Digital Methods.” Journal of Applied Ecology 49 (4): 960–67.

Burger, Alan E, and Scott A Shaffer. 2008. “Perspectives in Ornithology Application of Tracking and Data-Logging Technology in Research and Conservation of Seabirds.” The Auk 125 (2): 253–64.

Camphuysen, CJ, AD Fox, MF Leopold, and Ib Krag Petersen. 2004. “Towards Standardised Seabirds at Sea Census Techniques in Connection with Environmental Impact Assessments for Offshore Wind Farms in the UK: A Comparison of Ship and Aerial Sampling Methods for Marine Birds and Their Applicability to Offshore Wind Farm Assessments.” Report commissioned by COWRIE Ltd., London. www. offshorewindfarms. co. uk.

Carroll, Matthew J., Ewan D. Wakefield, Emily S. Scragg, Ellie Owen, Simon Pinder, Mark Bolton, James J. Waggitt, and Peter G. H. Evans. 2019. “Matches and Mismatches Between Seabird Distributions Estimated from at-Sea Surveys and Concurrent Individual-Level Tracking.” Frontiers in Ecology and Evolution 7: 333. https://doi.org/10.3389/fevo.2019.00333.

Chandler, Richard B, Daniel A Crawford, Elina P Garrison, Karl V Miller, and Michael J Cherry. 2022. “Modeling Abundance, Distribution, Movement and Space Use with Camera and Telemetry Data.” Ecology 103 (10): e3583.

Glennie, Richard, Stephen Terrence Buckland, Roland Langrock, Tim Gerrodette, LT Ballance, SJ Chivers, and MD Scott. 2021. “Incorporating Animal Movement into Distance Sampling.” Journal of the American Statistical Association 116 (533): 107–15.

Gutowsky, Sarah E, Marty L Leonard, Melinda G Conners, Scott A Shaffer, and Ian D Jonsen. 2015. “Individual-Level Variation and Higher-Level Interpretations of Space Use in Wide-Ranging Species: An Albatross Case Study of Sampling Effects.” Frontiers in Marine Science 2: 93.

Hays, Graeme C, Helen Bailey, Steven J Bograd, W Don Bowen, Claudio Campagna, Ruth H Carmichael, Paolo Casale, et al. 2019. “Translating Marine Animal Tracking Data into Conservation Policy and Management.” Trends in Ecology & Evolution 34 (5): 459–73.

Krietsch, Johannes, Steffen Hahn, Matthias Kopp, Richard A Phillips, Hans-Ulrich Peter, and Simeon Lisovski. 2017. “Consistent Variation in Individual Migration Strategies of Brown Skuas.” Marine Ecology Progress Series 578: 213–25.

Lilliendahl, Kristjan, Jon Solmundsson, Gudmundur A Gudmundsson, and Lorna Taylor. 2003. “Can Surveillance Radar Be Used to Monitor the Foraging Distribution of Colonially Breeding Alcids?” The Condor 105 (1): 145–50.

Louzao, M, J Bécares, B Rodrı́guez, K David Hyrenbach, A Ruiz, and JM Arcos. 2009. “Combining Vessel-Based Surveys and Tracking Data to Identify Key Marine Areas for Seabirds.” Marine Ecology Progress Series 391: 183–97.

Orben, Rachael A, Abram B Fleishman, Abraham L Borker, William Bridgeland, Amanda J Gladics, Jessica Porquez, Peter Sanzenbacher, et al. 2019. “Comparing Imaging, Acoustics, and Radar to Monitor Leach’s Storm-Petrel Colonies.” PeerJ 7: e6721.