Study to Develop Best Practice Recommendations for Combining Seabird Study Data Collected from Different Platforms

June 2022

Scottish Government
Riaghaltas na h-Alba
gov.scot

Study to Develop Best Practice Recommendations for Combining Seabird Study Data Collected from Different Platforms

Jason Matthiopoulos, Mark Trinder and Bob Furness

Published by Marine Scotland Science

Marine Scotland is the directorate of the Scottish Government responsible for the integrated management of Scotland's seas. Marine Scotland Science (formerly Fisheries Research Services) provides expert scientific and technical advice on marine and fisheries issues. These reports are not subject to formal external peer-review.

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

While every effort has been made to make this publication accessible to all, some sections may remain inaccessible due to the nature of the content. If you are unable to access any content you require, please contact ScotMER@gov.scot

MacArthur Green

Study to develop best practice recommendations for combining seabird study data collected from different platforms

Prepared by:	Jason Matthiopoulos, Mark Trinder, Bob Furness
Date:	25th March 2020
Tel:	0141 342 5404
Web:	www.macarthurgreen.com
Email:	mark.trinder@macarthurgreen.com

Address:	93 South Woodside Road | Glasgow | G20 6NT

Contents

Study to Develop Best Practice Recommendations for Combining Seabird Study Data Collected from Different Platforms

Jason Matthiopoulos, Mark Trinder and Bob Furness

1 Executive summary

Existing frameworks for the statistical analysis of spatial survey data offer a clear workflow towards the estimation of absolute and relative abundance of wildlife, in association with present and future environmental profiles (whether naturally or anthropogenically effected). At the same time, more broadly in applied ecology, there is a keen interest in integrated analysis and adaptive resource management. Momentum behind these ideas is encouraging the incorporation of different sources of spatial information onto a single, joint inference framework, so that statistical power can be greatly enhanced, even if the data themselves cannot be directly pooled because of their qualitative differences. The present project used systematic literature review, expert knowledge on survey methodology, bespoke model development and sensitivity analyses on realistic simulation data to derive methodological and quantitative guidelines for best practice in conducting such joint inference for multi-platform seabird survey data. We subdivide our recommendations into six distinct categories.

1.1 Appropriate response and explanatory variables

- Keep the highest-grade form of data. Data collected in aggregated or thresholded form can be analysed but these operations should be avoided on highly resolved data.
- Analyse even low-grade data as if originating from abundance. Use of latent surfaces of abundance allows us to interface lower-grade data with high resolution inference.
- Avoid inflated error structures until the end of modelling. Modelling with covariates will generally explain some of the over-dispersion in the raw data and use of spatially and temporally auto-correlated errors will account for unexplained hot- and cold-spots in distribution.
- Partly missing covariates should not necessarily lead to data censoring. This may prove necessary in the end of the analysis however, it may be worth attempting to reconstruct the covariate either as a separate interpolation step, or as part of an integrated analysis with partially missing data.

1.2 Treatment of survey design attributes and observation errors

- Use distance sampling. Distance sampling techniques facilitate the pooling of surveys with different protocols by reducing them into a common set of detectability characteristics. The extensions of distance sampling that deal with transect design and the incorporation of covariates facilitate error correction.
- Prioritise cross-calibration between surveys. Joint analysis of multiple surveys allows the combination of high detectability and high span. Surveys with known detectability errors should be prized highly because they can be used within a joint analysis to cross-calibrate less detailed surveys that may have happened close in space and time.
- Consider state-space approaches. Rather than correcting the observations for biases, prior to the formal analysis, a statistical observation model is combined with the biological model to effect the necessary correction in an integrated way. Both the biological and the observation models are tuned with regard to each other and uncertainty propagation from the observation model to the final predictions happens automatically.

1.3 Treatment of space time

- Use point process models. Point process models allow space-time to be modelled jointly and continuously, they subsume all other valid approaches to species distribution modelling and are compatible with other features of modelling developed to enhance predictive power.
- Use auto-correlated structures. Spatially and temporally auto-correlated structures can account for missing covariates, they can be used to impute gaps in covariate layers and, most importantly, they are the best way to leverage information sharing between surveys according to their spatiotemporal overlap or proximity.
- Take dynamics into account. If we need to account for multi-survey data that include before-and-after control impact, it is important to account for temporal non-stationarity.

1.4 Accessibility and density dependence

- Use realistic distance measures. If we are concerned that birds avoid flying over land, circumnavigate human structures, or if, due to glide-flight, they rely on prevalent wind direction, it is important that we account for these effects in the measure of distance.

- In the present, use abstracted models for density dependence. Currently, the computational demands of a fully spatially explicit model of intra-colony, inter-colony and interspecific competition are prohibitive. We have provided an illustration (in the project vignette), of how a pragmatic model for these processes can be developed and incorporated into joint modelling.
- In the future, consider spatially explicit models for density dependence. As computational approaches become more widespread in the field of SDMs, it may become possible to model competition in a fully spatially explicit way.

1.5 Inferential Platforms

- Use hierarchical models. These allow us to use features such as cross-calibration of observation models, covariate imputation and latency, and use of spatio-temporal proximity to allow the predictions to borrow strength from multiple surveys.
- Use Bayesian approaches. Computer-intensive Bayesian model-fitting allows state-space and hierarchical structures. More importantly, Bayesian inference permits the elicitation of expert opinion in the form of parameter priors.
- Use Data integration. Under joint inference, multiple data sets are analysed simultaneously to extract maximum power. These approaches are also particularly useful for extending the analyses to non-survey data.
- Fully propagate uncertainty to the final predictions. There is always a limit to how much missing information can be imputed by statistical modelling. The multiple sources of uncertainty along data collection and estimation need to be translated to aggregate measures of precision in the final spatial predictions.

1.6 Computational platforms

- Support open source. As a matter of process, all code developed by government funding should be made available to the scientific community.
- Ensure strong interface with Geographic Information systems. Establishing stable protocols for data formatting, and by using the GIS functionality in platforms like R would allow data processing on a single platform.
- Parameterise non-linear model components with exact methods. The prototype models presented in the jointSurvey library are computationally greedy, but they have the best chance of retrieving the difficult parameters pertaining to density dependence and competition.
- Implement large scale predictions using fast approximate methods. It is imperative to move towards efficient methods, such as INLA.

In addition, we review future extensions of methods that could facilitate integration of single-species, multi-survey data with a variety of other data, including survey data from other species, shore-based vantage point data, citizen science data, telemetry (tracking) data, mark-recapture data and demographic data.

2 Project brief

2.1 Overview and objective

Established bodies of theory and software address different challenges in answering questions of species abundance and distribution from survey data, particularly with regard to observation biases (e.g. distance sampling methodologies and frameworks for partial detectability), linking species distribution to habitat (e.g. resource selection functions) and enhancing the transferability of predictions from these models in space and time (e.g. generalised functional responses in resource selection). Although these methods are still the subject of very active research and development, they offer a clear workflow towards the estimation of absolute and relative abundance of wildlife, in association with present and future environmental profiles (whether naturally or anthropogenically effected).

At the same time, more broadly in applied ecology, there is a keen interest in integrated analysis and adaptive resource management. Momentum behind these ideas is encouraging the incorporation of different sources of spatial information onto a single, joint inference framework, so that statistical power can be greatly enhanced, even if the data themselves cannot be directly pooled because of their qualitative differences.

In this project, we used systematic literature review, expert knowledge on survey methodology, bespoke model development and sensitivity analyses on realistic simulation data to **derive methodological and quantitative guidelines for best practice in conducting such joint inference for multi-platform seabird survey data**.

2.2 Adherence to Marine Scotland remit

The aim of this project was to examine how to compare or combine multiple sets of seabird survey data collected from different survey platforms and/or with different temporal/spatial resolution and coverage to adequately characterise seabird distribution and abundances and use this information to develop best practice recommendations to use in assessments.

We considered boat based visual surveys (ESAS methodology), visual aerial surveys, and digital aerial surveys. Regarding data integration we considered different platforms, incompletely overlapping spatial extents, different temporal coverage (e.g. monthly versus seasonal), different spatial coverage, different survey resolution (e.g. swath width and/or transect spacing) and surveys conducted in different years.

By implementing different solutions on synthetic data, we have demonstrated the most appropriate methods and the key considerations for integrating multiple survey datasets. This has involved using existing modelling tools and code and the development of bespoke modelling tools and code.

Solutions are applicable across a wide range of marine bird species, and we have considered four exemplar species (Northern Gannet, Black-legged Kittiwake, Common guillemot, Great black-backed gull).

We have limited our exploration to how best to incorporate survey data from different platforms to characterise seabird distribution and abundance. We have not commented on how previous assessments have been done, but we have outlined those survey design requirements that facilitate data integration. We realise that such recommendations were not required under the project brief but they may usefully inform Scottish Government policy and project level scoping discussions.

2.3 Challenges addressed

The specific challenges that needed to be addressed in deriving guidelines are:

1. **Seabird natural history:** Models for abundance and distribution in seabirds must account for coloniality. Seabird distributions are not just the result of habitat suitability but also of accessibility that varies by colony location, species and season. Difficult questions pertaining to density dependence within colonies or between colonies of conspecifics and hetero-specifics also need to be taken into account.
2. **Survey method characteristics:** Different survey methodologies (boat-based, aerial visual and aerial digital) are affected by different types of biases and imprecisions. These need to be explicitly accounted for.
3. **Effort scale characteristics:** For a fixed amount of effort, any survey will make a decision on the trade-off between spatial/temporal resolution and extent. Different surveys may have entirely different designs, and their overall

effort may also differ. These discrepancies offer challenges, but also opportunities for complementary use of different surveys.

4. **Habitat data:** Similar issues relating to differential resolutions, extents and data absence will permeate the habitat data (e.g. bathymetry, primary production, seabed sediment, any prey survey data etc.). When habitat data are dynamic (e.g. seasonal) these problems are likely to be particularly acute. In analysing spatial data, but particularly when trying to analyse multiple survey platforms in tandem, it is essential to have guidelines for how to deal with missing or incongruent habitat data.
5. **Statistical robustness:** Integrated analyses of multiple data sources aim to enhance statistical power by greatly increasing the effective sample size (but also by using data from different regions, different times and spatial resolutions in a complementary way). Achieving this is the main objective of this project, but it must be done in a way that does not misleadingly increase the apparent precision of the results and model predictions. This could threaten the precautionary approach and have adverse implications for management and policy decisions. Therefore, uncertainty in the observation processes from different surveys and the habitat data must be correctly propagated to the end-results, to give a reliable measure of precision.

2.4 Relevance to Marine Scotland

The intended applications of this work will be in monitoring of marine protected areas, development of marine planning and the licencing workflow for offshore renewables. We identify the following beneficial links to Marine Scotland's key responsibilities:

- **Marine planning** requires good information about spatial and temporal trends in abundance. New human activity needs to be able to avoid critical hotspots of species distribution at valuable habitats and we need to be able to anticipate future crises before they arrive. Spatial and temporal prediction is the core theme of this bid.
- **Integrated planning** is best achieved with integrated data analysis, such as the frameworks outlined here.
- **Fisheries** although the present project focuses on seabirds, the implications for Marine Scotland's broader remit are two-fold. First, it has important implications for the management of fish stocks through estimates of seabird by-catch or predation pressure on managed stocks. Second, methods on multi-survey integration can have direct applicability to taxa beyond seabirds, such as fish.

- **Evidence informing marine development** important implications of conflict, particularly with offshore development and shipping. See marine planning, above.
- **Develop Marine Scotland's organisational skills and competencies p**art of this project is an extensive workshop to Marine Scotland staff (and partners) on the methods of multi-platform analysis.
-

3 The state-of-the-art in seabird distribution modelling

3.1 Overview of species distribution modelling

Statistical analyses of spatial survey data aim to address four questions (Aarts et al. 2008). 1) How many individuals there are within a survey area (abundance estimation), 2) Where they are in space (population distribution), 3) Why they are there (habitat associations), 4) Where else they might be, and where they might go if the environment changes (spatial extrapolation and forecasting).

A puritan taxonomy of SDMs

At first sight, the diversity of methods available for converting spatial data to prediction maps can seem overwhelming. However, there is an emerging hierarchy in the methodological literature that considerably simplifies our effort to outline recommendations for best practice. We can present this as a succession of four branchings, leading up to our recommended approach of inhomogeneous point process models (Figure 1).

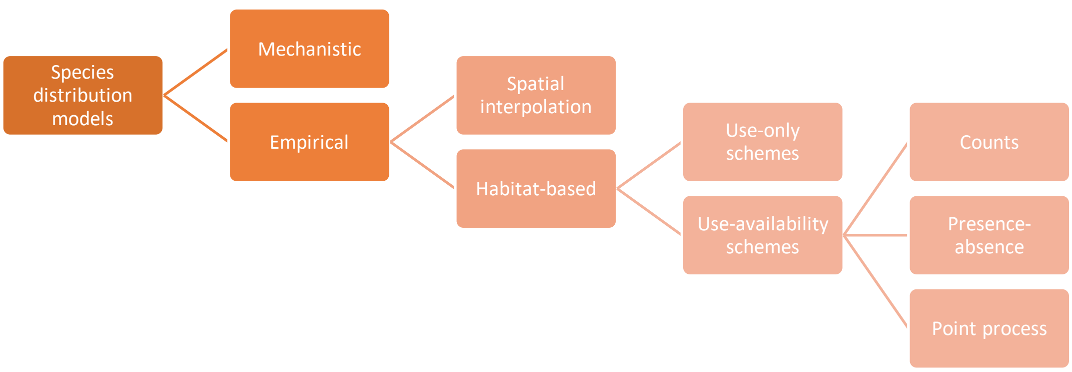

Figure 1: An overview of choices leading up to the recommended approach for species distribution modelling.

Spatial predictions can be generated by building mechanistic models of animal movement and demography from first principles and scaling them up to population distributions (Moorcroft et al. 1999, 2006, 2008, Moorcroft 2012). Arguably, models with a high biological detail (e.g. based on principles of physiological tolerance, movement behaviour and social interactions) bring greater insights and predictive capability (Kearney and Porter 2009, Hefley and Hooten 2016, Robinson et al. 2017). However, mechanistic modelling can be quite demanding technically and is generally more vulnerable to model misspecification and over-parameterisation. Misspecification will result in models that disagree with the data and over-parameterisation will generate models that cannot be sufficiently informed by the available data. For example, the series of papers by Moorcroft et al. cited above, which form the state-of-the-art in mechanistic distribution modelling, require specific assumptions about animal movement and rely on sufficiently mathematical users who can formulate and manipulate partial differential equation models. Alternatively, and more readily, the analysis can be done by mimicking observed patterns of space-use with the aid of empirical, statistical models (Guisan and Zimmermann 2000, Guisan and Thuiller 2005, Guisan et al. 2017). The deciding trade-off between mechanistic and empirical models is one of realism and predictive capacity against the ease of use and robustness to misspecification.

Within this class of empirical models (see review by Matthiopoulos and Aarts 2007), we can distinguish between models that merely reconstruct the spatial density of a population (such as kernel smoothing, additive smoothing, or geo-statistical methods) and regression methods that rely on habitat information as explanatory variables. Spatial density estimation methods rely on geographical proximity and the existence of spatial autocorrelation (Levin 1992) to interpolate between observation points and map density in unobserved space, or alternatively, to smooth a finite data set of synoptic observations into a population-level expectation of usage. Density estimation methods focus on removing spurious variability from the predictions, but aim to stay as close as possible to the observations. Therefore, their ability to describe the available data is often better than that of habitat models (Bahn and Mcgill 2007). Habitat models, on the other hand, are not by default spatial, since they are fitted in environmental (or niche) space (Pearman et al. 2008). Consequently, the greater ability of habitat models to interpolate and extrapolate spatially relies on the quality and relevance of their underpinning covariates. Nevertheless, the need to include such covariates when modelling seabird distributions is not in doubt (Camphuysen et al. 2004). The deciding trade-off between density estimation and habitat models is one of faithfulness to the particular distributional data collected and the ability to extend predictions beyond the spatial and temporal frame of data collection.

Within the class of habitat models, we distinguish between two main categories. The first, are known as profile methods and they argue that knowledge of where, in niche space, a species occurs is sufficient to understand its fundamental niche and map its current and future distribution. Broadly, this category includes methods such as climate envelope models and the use of multivariate statistical methods such as PCA (Robertson et al. 2001) for the analysis of presence only methods (reviewed in Pearce and Boyce 2006). The alternative class of use-availability schemes either contain representative information on the distribution of organisms (i.e. presence and absence), or they supplement presence data with availability data, allowing the models to contrast the habitat choices that organisms made, with the options that they had available to choose from. The broad area of use-availability schemes includes the vast literatures on resource selection functions (Boyce and McDonald 1999, Morris et al. 2016) and maximum entropy approaches (Elith et al. 2011, Merow et al. 2013). Profile methods have been critiqued extensively in the methodological literature (see Pearce and Boyce 2006 for a review), and there is really no sound scientific reason for choosing to use a profile method.

The final decision stage is mostly perceptual, relating to how space is conceptualised for the purposes of modelling the data. For example, space may be thought of as a regular grid (e.g. comprising squares, or other regular forms of tessellation, such as hexagons - see Grecian et al. 2016). In that case, the spatial data take the form of counts and are modelled by appropriate probability models such as the Poisson. Alternatively, space may be thought of as continuous and different spatial locations may be characterised by whether a species was present or absent. In that case, spatially reference data take the form of zeroes and ones and the most appropriate probability model is Bernoulli (Aarts et al. 2008). Yet another approach within the continuous space framework is to imagine that observations of organisms appear at random locations almost like pin-lights that blink in and out at different time frames of observation. This framework, known as the Inhomogeneous Point Process, models the occurrence of events within a unit of time and space as originating from a smooth intensity surface, describing the instantaneous and infinitesimal rate of the Poisson process (Chakraborty et al. 2011, Aarts et al. 2012, Fithian and Hastie 2013, Renner et al. 2015, Fletcher et al. 2019, Miller et al. 2019). It is an elegant approach that makes an implicit comparison between use and availability, captures heterogeneities in the distribution of the population (e.g. due to environmental covariates) but, can with equal ease, use the intensity surface to represent heterogeneities in the distribution of spatial observation effort (so that, regions that receive no observation effort will have a zero intensity when modelling the data). The deciding trade-off between count, presence-absence and point process models is in whether the user feels comfortable in conceptualising infinitesimal quantities and is happy to relinquish the notion of explicitly contrasting use and availability for the advantage of greater spatial and temporal precision in model fitting and prediction.

Hybridisation of SDMs

Different modelling approaches to species distributions are often presented as non-overlapping, much as we did in the preceding section. This often gives the impression that the only way to deal with multiplicity of approaches is to compare their performance and choose the best (e.g. Oppel et al. 2012). However, there is a methodological kinship between many of these approaches that is rarely apparent in the applied literature. Having so far organised the literature as a sequence of three strict dichotomies and a final trichotomy (Figure 1), it is good to take a less clear-cut, but more synthetic and conciliatory view on the above decisions. It is, in fact, the case that a hybrid approach is possible that retains the best elements of all the approaches discussed above.

Specifically, if one starts from a purely empirical model it is possible to move it towards a higher mechanistic content. In the simplest case, this can be done by carefully considering the biological relevance of the set of covariates that are offered to the model (Bell and Schlaepfer 2016). It is also possible to construct more sophisticated covariates using mechanistic models to try and increase the explanatory power of empirical models (Kearney and Porter 2009, Matthiopoulos et al. 2015). More recently, it is becoming possible to fit structurally complex models directly to data either by likelihood approaches, but most often, via Bayesian approaches. These developments have come mostly from the field of integrated population modelling (Matthiopoulos et al. 2014, Zipkin and Saunders 2018, Yen et al. 2019). As well as data-integration, the main benefit of integrated models is their capability to deal with non-linear features, a feature characterising most biologically realistic (i.e. mechanistic) models. We will use this aspect of hybridisation in this report when considering density-dependent effects on seabird distribution.

Also, the separation between spatial and habitat-based models can be made less strict. Geo-statistical models can accept habitat covariates and habitat models can accept spatial autocorrelation structures (Dormann et al. 2007). A considerable advantage of these models is that they optimally account for variability in the data (but see Hodges and Reich 2010). A perceived limitation of this approach is that predictions from hybrid spatial-habitat models are tied to the spatial extent of the data collection. We will discuss these issues at length later in the report and use them to construct spatial-habitat hybrid models.

The separation between use-only (or, profile) models and use-availability models is perhaps the most clear-cut of the branchings in Figure 1. Profile methods are problematic because, by ignoring the availability of different habitats, they interpret the habitat choices of organisms as purely the result of preference, not as a combination of preference and availability (Matthiopoulos 2003a). Profile models also misappropriate the ecological term “niche” because they aspire to define a species’ viable hyper-volume in environmental space, and yet make no explicit connection between habitat data and population trends representing information on viability (Peterson et al. 2011, Matthiopoulos et al. 2015). Therefore, profile methods are fundamentally flawed from an ecological perspective. And yet, despite their limitations, their aspiration is worthwhile. Using habitat models to make sense of population viability should be a key objective in our search for defining critical habitat and in driving conservation efforts. Recent publications (Matthiopoulos et al. 2015, 2019) have shown how this can be achieved in practice by using the more defensible option of use-availability methods as a platform to build upon. We will briefly discuss how these methods could be used to integrate non-spatial data with spatial models.

Data types commonly used for SDMs are count, presence-absence and presence only (Hefley and Hooten 2016). Count data can be divided into point counts (e.g. point or line transects) and quadrat counts (comprehensive count in an area) (Hefley and Hooten 2016), although the distinctions between those two can be blurred. Presence-absence (or occupancy) data may either originate from count data that have been converted to binary form, or they may be the result of survey effort units that were terminated as soon as the species was detected once (Hefley and Hooten 2016). Finally, presence only data may include observations from known survey effort units (e.g. telemetry), or alternatively unknown effort surveys (such as museum records, or some citizen science programmes). Several papers (Warton et al. 2010, Aarts et al. 2012, Fithian and Hastie 2013, Hefley and Hooten 2016) have shown that the separation between count, presence absence and point process models is not substantial. Indeed, all of these methods can be thought of and re-formulated as inhomogeneous point processes. Furthermore, widely used spatial modelling packages such as MAXENT, can be thought of as point process models (Fithian and Hastie 2013, Renner and Warton 2013). Conversely, computational methods used for efficiently fitting point process models to data make use of spatial discretisation, similar to grid-based methods, but using more efficient schemes tailored to the data (Lindgren 2015).

During the rest of this report we will move towards recommending an approach that, while based on an empirical, habitat-driven point-process model, is capable of incorporating mechanistic (non-linear) features, using explicitly spatial information and is able to provide useful results for future assessments of population viability. Our approach will be based on the principles previously outlined by work funded by Marine Scotland (Oedekoven et al. 2012a, 2012b) and we will expand on these ideas, exploiting the opportunities offered by multiple surveys. Only around 7% of peer-reviewed marine SDM publications since 1992 have focused on seabirds, a total of 16 papers (Robinson et al. 2017). Although this survey excludes several publication-quality reports on seabird SDMs from the grey literature (Burt et al. 2009, Bradbury et al. 2011, Petersen and Nielsen 2011, Rexstad and Buckland 2012) it is nevertheless an indicator that we need to also look more broadly at the lessons learned from marine survey methods in general, as well as the transferrable elements of the ensuing analyses.

3.2 Overview of marine survey methods

Historically, and until the 1970s, knowledge of the distribution of seabirds around the UK comprised little more than maps of breeding colonies and the expectation that seabird numbers were probably high in the waters around them (Camphuysen et al. 2004). However, the need for risk assessment (and hence, more detailed spatial surveys of at-sea distribution) has been driven by the rapid developments in the energy industry (initially oil extraction, but then marine renewables and more recently, platform decommissioning). Initial efforts focusing on strip-transect designs (Tasker et al. 1984), were followed by stratification of distances from the observer (into detection bands) and by fully developed line-transect designs (Buckland et al. 2001, Thomas et al. 2010). Line transects have been undertaken by boat or aircraft and current best practice for survey design is summarised in (Camphuysen et al. 2004, Certain and Bretagnolle 2008, Oedekoven et al. 2012a, 2012b, Webb and Nehls 2019). Indeed, (Oedekoven et al. 2012b – based on earlier work by Camphuysen et al. 2004) give an excellent overview of the four distinct types of data collection in seabird research: boat surveys, visual aerial surveys, digital aerial surveys and vantage point surveys. More recent projects have also looked at tracking data to describe seabird-at-sea distributions – such as the RSPB FAME/STAR analysis and the Norwegian SEATRACK programme – both of which aim to describe seasonal spatial distribution patterns from tracking data sets, potentially even ignoring all of the other data types. We return briefly to tracking data in Section 7.5,

Line transect methodology has prompted the development of distance sampling methods (Buckland et al. 2008, Thomas et al. 2010) which deal with imperfections of the observation process such as decaying detectability of the subject with increasing distance from the observer (Buckland et al. 2001), the overall detectability at zero distance (Buckland et al. 2001) and the effect of environmental conditions on detectability (Marques and Buckland 2003). Although the original application of these methods was in estimating total population size, they have since been extensively applied to the estimation of relative abundance in order to create species distribution maps, often in response to habitat covariates (Clarke et al. 2003, Hedley and Buckland 2004, Oedekoven et al. 2012a, 2012b, Thiers et al. 2014, Waggitt et al. 2019).

Table 1

The effect of different detectability biases on strip width (effective detection distance) and baseline detection (detection probability at zero distance from the observer).

Source of bias	*Description*	*Strip w.*	*Baseline pr.*
Distance effect	Detection distance is the main variable in line transect analyses.	✓	
Vantage height	Determined either by the observation deck on a ship or the flight altitude of the aircraft.	✓	
Study species morphology	Smaller animals may be harder to see, and body pigmentation may camouflage them against background.	✓	✓
Behaviour of species	Diving birds may be missed, and flocking birds may be over- or under-counted.		✓
Platform effects	Smaller boats may be more vulnerable to movement resulting from wind or waves and noisy engines may conceivably repel animals. Conversely, animals that have developed vessel-following behaviours may be attracted by the presence of a boat.	✓	✓
Visibility conditions	Weather and time of day will affect visibility. More generally, noise will be a problem for methods that use sound attenuation, although this is less important for seabirds. Conditions may change during a single survey or, even, a single transect, affecting the apparent abundance of the species.	✓	✓
Digitisation	Although transfer of visual data to digital form is likely to be supervised by human technicians, processing of digital images may require some level of automation. This could result in over, or under-counting.		✓
Observer experience	Different observers with different levels of skill or training may be introducing biases in their visual detections. These biases may remain consistent within observers, or they may diminish with experience.	✓	✓

A simplistic but instructive statement about distance sampling is that it uses statistical modelling to bring line transect surveys closer to the original assumptions of strip transects. The effective detection distance derived from distance sampling methods allows us to assume that all event occurrences (i.e. birds) within an idealised strip of that width are detected with the same probability while all events outside that strip are missed. New field methods using digital photography from aerial platforms, yielding a relative density of birds per unit area, (Buckland et al. 2012, NaturalEngland 2019) are now the only survey method accepted by the regulator in some countries (e.g. Germany). These developments bring us full-circle to the strip transect as the gold standard. Therefore, counts that have been (or can be interpreted to have been) obtained from strip transects are the necessary starting point for analyses of population size and population distribution.

There are several aspects of the survey specification that can affect detectability but they all reduce to two main effects: the probability of detection at zero distance, and the effective detection distance from the observer. We use Table 1 to enumerate a variety of survey biases and illustrate that their effects can be entirely captured (i.e. corrected for) by modelling these two aspects of strip transects (even if this correction may, in some cases, require additional, calibrating data).

In this report, we, therefore, assume that the data correspond to strip transects, an assumption that would be correct for the raw data in the case of digital aerial surveys, but which implies a distance sampling pre-processing stage for line transect data (both ship-borne and aerial). Particularly in the setting of a multi-survey study (e.g. for the purposes of synthetic data generation in this project), a computational treatment of distinct surveys requires that each survey is concisely and uniquely characterised by a set of parameter values. In this report (see for example Section 3 of the accompanying vignette), the *design specification* of a survey is reduced to six characteristics. We divide those into characteristics of *span* and *detectability*. We outline them below, and also take the opportunity to collect all of the recommendations made by the pivotal report of (Camphuysen et al. 2004) for ship and aircraft survey design under each characteristic.

Span characteristics

Survey extent. Particularly with regard to marine developments, "*it is recommended that a high resolution grid should be deployed, covering an area at least 6x the size of the proposed wind farm area, including at least 1-2 similar sized reference areas (same geographical, oceanographical characteristics), and preferably including nearby coastal waters (for nearshore wind farms only)*" (Camphuysen et al. 2004). As we will discuss in the transferability section of this report (Section 5.5), it is important to conduct a cross-sectional sampling of space, but also to achieve as much spread of environmental conditions in the data, unaffected by expectations of where the species is likely to be.

Spacing of successive count locations. The spacing of locations will be determined by the combination of platform speed and sampling intervals. For ship surveys, "*time intervals are recommended to be one or five minute intervals (range 1-10 m, longer time intervals are acceptable when less resolution of data is required; short intervals are preferred in small study areas), with mid-positions (Latitude, Longitude) to be recorded or calculated for each interval. Preferred ship's speed should be ten knots (range 5-15 knots)*" (Camphuysen et al. 2004). For visual aerial surveys, "*speed preferably 185 km h-1 at 80 m altitude. GPS positions should be recorded at least every 5 seconds (computer logs flight track)*". Currently implemented digital aerial surveys are based on either still images recorded at intervals pre-determined to achieve a specific percentage coverage or continuous video recording, with image width set to achieve the target coverage.

Spacing between transects. Ship-based, "*survey grid lines are recommended to be at least 0.5nmi apart, maximum 2nmi apart*". Aerial "*transects should be a minimum of 2 km apart to avoid double-counting whilst allowing the densest coverage feasible*" (Camphuysen et al. 2004). Currently implemented surveys aim for distances of 2 km (ship), 2-3 km (aerial visual) and typically 2.5 km (aerial digital), although the latter varies depending on the target coverage percentage.

Transect orientation. Three main considerations enter the determination of the orientation of transects: statistical effort (e.g. avoiding overlaps), logistical constraints (e.g. need to avoid off-effort segments), geomorphological constraints (e.g. coastline and shallow water avoidance in ship surveys). No direct recommendations on this issue were provided specifically for seabirds by Camphuysen et al. (2004) but a general, concise discussion of automated transect design pertinent to this issue can be found in (Strindberg and Buckland 2004). Digital aerial methods, particularly those conducted farther offshore where variations in relation to distance to coast are expected to be reduced may also be oriented to minimise glare from the sea surface.

Detectability characteristics

Effective detection distance (strip width). Ship-based, "*line-transect methodology is recommended with a strip width of 300 m maximum. Subdivision of survey bands is recommended to allow corrections for missed individuals at greater distances away from the observation platform. Preferred ship type is a motor vessel with forward viewing height possibilities at 10 m above sea level (range 5-25 m), not being a commercial or frequently active fishing vessel… Preferred ship-size: stable platform, at least 20 m total length, max. 100 m total length*". In practice, current visual surveys use banding (ship bands 0-50,50-100,100-200,200-300, aerial visual bands are variable but often used are 44-163,163-282,282-426,426-1000. Note that unless 'bubble' windows are available 0-44 m is missed due to restricted view) and truncation distances (ship, 300 m one-sided survey, aerial visual 1 km two-sided) and aerial digital surveys use a strip width that will depend on flight height (300 m for still and 500 m for video). Other recommendations pertain to potential effects on the decay of detectability with distance. For example, it was suggested that, "the grid should be surveyed such that time of day is equally distributed over the entire area (changing start and end time over the area to fully comprehend effects of diurnal rhythms in the area)", also, to "*use an inclinometer to measure declination from the horizon*" and to avoid "*observations in sea states above 3 (small waves with few whitecaps)*" .

Baseline detection probability (detection at zero distance). Similar considerations relate to the baseline detection probability. For ship-based surveys, "*no observations in sea state 5 or more to be used in data analysis for seabirds. Bird detection by naked eye as a default, except in areas with wintering divers Gaviidae. Scanning ahead with binoculars is necessary, for example to detect flushed divers. Two competent observers are required per observation platform equipped with rangefinders, GPS and data sheets; no immediate computerising of data during surveys to maximise attention on the actual detection, identification and recording. Observers should have adequate identification skills (i.e. all relevant scarce and common marine species well known, some knowledge of rarities, full understanding of plumages and moults). Observers must be trained by experienced offshore ornithologists under contrasting situations and in different seasons*". Correspondingly, for visual aerial surveys, it is recommended using, "*high-wing aircraft with excellent all-round visibility for observers. Two trained observers, one covering each side of the aircraft, with all observations recorded continuously on Dictaphone. No observations in sea states above three (small waves with few whitecaps)*" (Camphuysen et al. 2004).

Survey particulars for exemplar species

In general, boat platforms are slower than aerial ones. Visual aerial surveys are fast but are characterised by lower detection rates than digital ones and fly lower so can cause flushing or avoidance in susceptible species. Digital aerial surveys are fast and fly higher, hence having lower risk of response behaviour and yield high rates of detection and species identification. Nevertheless, the relative accuracy of visual v digital methods in marine surveys is situation- and species-dependent (Furness 2016). More specifically, regarding the relationship of the four exemplar species to each of the three platforms: Gannets and great black-backed gulls are likely to be attracted to boats, kittiwakes probably not, whereas guillemots will avoid boats and be hard to spot (on the sea) as well as having a proportion of the birds underwater when the survey passed by. Visual detection from the air is likely to be good for gannet and great black-backed gulls but lower for guillemots. Species identification may be problematic for kittiwake, as a small gull species. Availability bias (birds underwater) may be particularly problematic for guillemots and sun glare may also be an issue for all of them. In digital aerial surveys, detection should be good for all four species although some kittiwakes may be classed as small gull species and great black-backed gulls as large gulls.

4 Approaches to multi-survey modelling

There is an appetite more broadly in applied ecology for integrated analyses and adaptive resource management. Such ambitions (and, indeed, the terms “data integration” or “data pooling”) are motivated by the statistical community (see Section 7.1 on extensions) but are also expressed by more descriptive papers (e.g. Perrow et al. 2015), indicating that there is an increasing dissatisfaction with piece-wise comparisons between surveys and species. Momentum behind these ideas is encouraging the incorporation of different sources of spatial information onto a single, joint inference framework, greatly enhancing statistical power, even if the data themselves cannot be directly pooled because of their qualitative differences. For a fixed amount of effort, any survey will make a decision on the trade-off between spatial/temporal resolution and extent. Different surveys may have entirely different designs, and their overall effort may also differ. These discrepancies offer challenges, but also opportunities for complementary use of different surveys.

4.1 Indiscriminate pooling

The most naïve but (at 73% of reviewed papers by Fletcher et al. 2019) also, the most prevalent approach to dealing with multi-survey data is to pool them without considering their particular observation biases and imprecisions. The (rather wishful) expectation is that somehow these errors will cancel each-other out to give unbiased estimates of distribution and habitat preferences.

4.2 Ad-hoc Comparison

The most obvious approach to using data from multiple sources is to analyse each source individually and then compare the outputs of such analyses. In some cases, the comparison takes the form of validation and calibration (Munson et al. 2010), but this has assumed the existence of a gold standard (i.e. a high-resolution, precise and accurate data set), which may not necessarily be available, particularly in the marine environment. Nevertheless, the core idea of calibrating one data set based on another (as featured in Munson et al. 2010 and elsewhere) need not require that either data set is perfect. Indeed, imperfect observations "borrowing strength" from each other has been widely applied elsewhere in spatial survey design (see double-observer methods in Buckland et al. 2010). This is a concept that we will rely on in Section 4.3 and beyond.

However, staying, for now, with the idea of map comparisons from different (imperfect) data sets, these are carried out either visually (e.g. Bradbury et al. 2014, Perrow et al. 2015), or via some ad-hoc quantitative method (Sardà-Palomera et al. 2012, Sansom et al. 2018). For example, (Sansom et al. 2018) used four distinct analyses carried out on data (both survey and telemetry) from four UK seabird species. Using as their starting point the utilisation maps generated from each analysis, on each species, they performed all possible pairwise comparisons. They focused on overlap between each pair of maps measured both as the extent (area) and density (utilisation) shared by them at their core areas (defined using varying density contours). This allowed them to discuss patterns of similarity in these estimated snapshots of distribution. However, they were not in a position to draw combined inferences about parameter values relating the patterns of utilisation to their underlying covariates. Further, they were not in a position to share statistical power between surveys conducted on the same species, possibly at similar times or regions.

Overall, therefore, such ad-hoc comparisons are biologically valuable because they inform intuition and motivate scientific hypotheses. However, methodologically, they are of limited utility because they do not facilitate the flow of information between data sets.

4.3 Post-hoc combination

An improved approach which allows information from one data set to flow into another (but not the other way around) is a sequential analysis, which completely deals with one data set first and then somehow incorporates the second data set as a second stage of fitting. Such approaches are not wide-spread and they seem to be specific to the particular analyses at-hand (Yamamoto et al. 2015). However, particularly in the context of Bayesian updating, where sequential analyses are possible, it is plausible to think of methods that use the results of one analysis (based on a single data set) to specify priors for the analysis of the next data set (Matthiopoulos 2003b, Talluto et al. 2016). Such ideas have been proposed, but not realised in spatial ecology, mainly because they require assigning parametric probability distributions (the priors) to space as a whole.

→ An alternative idea, ensemble forecasting, examines a large (infinite, even) models of a system (Araújo and New 2007). Instead of picking the best model from the ensemble, assuming that each model carries some independent information, the combination of forecasts from different models is characterised by a lower mean error than individual forecasts. This idea generalises on the field of model averaging (Dormann et al. 2018) because ensembles can be created by examining different models, different parameterisations of the same models, different initial or boundary conditions and different stochastic realisations from each model (Fig. 1 in Araújo and New 2007). Post-hoc combinations from an ensemble can be unweighted (i.e. combination by committee) or weighted according to some measure of quality (e.g. based on assessments of data precision). Established model averaging methods adopted in ecology have previously used weights derived from information criteria (Burnham and Anderson 2004, Burnham et al. 2011), hence rewarding parsimony in the weighting.

→ The above post-hoc approaches seem to fall naturally into categories of parallel and sequential model fitting. The "wisdom of crowds", a pervasive idea represented here by ensemble modelling, achieves a pooling of predictions from a collection of *parallel* models. This leads to robust predictions, but the models are not allowed to inform each other. Less developed, but perhaps more powerful ideas about *sequential* model-fitting (Matthiopoulos 2003b, Yamamoto et al. 2015) allow later models to be informed by earlier ones, but the information flow is unidirectional.

Summarising some of the above ideas in their recent review on data integration, (Fletcher et al. 2019) identify three distinct cases of ad-hoc combination. 1) Ensemble modelling of independent models, 2) Use of the maps produced by one model as a covariate participating in the linear predictor of the other model and 3) In a Bayesian context, using one model to generate informative priors for the parameters or the predicted distributions of the other model. Together, these categories take up about 20% of the data-combination literature.

4.4 Spatial data integration

Integrated analyses of multiple data sources aim to enhance statistical power by greatly increasing the effective sample size of the data set but, also, by using data from different regions, different times and spatial resolutions in a complementary way. Developing the fundamentally useful idea of calibration (see Section 4.3), into the more general concept of integration, several papers (Fletcher et al. 2016, Pacifici et al. 2017, Koshkina et al. 2017, Peel et al. 2019) examined whether using presence-only (opportunistic) data in combination with the gold standard of presence-absence (survey) data could improve the descriptive and predictive ability of species distribution models. This is indeed likely to be the case, but the statistical method for achieving it must first be considered, so that the inferential platform, built from the perspective of calibration, can be used for integrated analyses that do not necessarily contain a gold standard.
Fletcher et al. (2016) pointed out that the first methodological decision in data integration is whether space should be treated as a nested hierarchy of grid resolutions (e.g. Keil et al. 2013, 2014) or as a continuous plane of coordinates. The former approach is possible by conditioning higher resolution grid cells on the observed/estimated contents of lower resolution grids, but this runs the risks of data mismatches between scales. The alternative, of treating space as continuous is represented by the Inhomogeneous Poisson Process approach, discussed in Section 3.1. The key conceptual advantage of IPPs is that they acknowledge that spatial processes occur at individual points in space and may remain unobserved (see, thinned IPPs), be reported with some spatial error, or be aggregated into counts at coarser spatial resolutions. Hence, the IPP paradigm recognises that the data will have an underlying common scale, even if they are reported at coarser resolutions. The underlying IPP is considered *latent* or *unobserved*. Different data can then be considered to originate from it, subject to the span and detectability limitations of the particular survey scheme (see Section 3.2). This allows us to write a *joint likelihood* for multiple data sets, conditional on the latent IPP. The distinction between a latent biological process and the different data-collection processes that

can be used to observe it gives us the ability to think more mechanistically about the origin of the data (Hefley and Hooten 2016, Fletcher et al. 2019).

Data integration must also be done in a way that does not misleadingly increase the apparent precision of the results and model predictions (Miller et al. 2019). For example, un-modelled spatial and temporal autocorrelation in the data (see Section 5.4) may artificially inflate the apparent sample size of the data, despite the prudency recommendations made for spacing out observations and transects (see Section 3.2). These concerns about pseudo-replication apply particularly for multi-survey analyses because different surveys may have overlapped in space or in time. Alternatively, uncertainty contained in the pre-analysis of transects (e.g. uncertainty in the detection function see Section 5.1), if not propagated to the final results, may under-represent the uncertainty in distribution. All of these mechanisms could threaten the precautionary approach and have adverse implications for management and policy decisions. Therefore, uncertainty in the observation processes from different surveys and the habitat data must be correctly propagated to the end-results, to give a reliable measure of precision. The current situation in the literature is far from ideal, given that most published marine SDM studies (94%) have failed to report the amount of uncertainty derived from data deficiencies and model parameters (Robinson et al. 2017).

Fletcher et al. (2016), Pacifici et al. (2017) and Peel et al. (2019) found that the combination of the data gave better explanatory and predictive performance than either of the two data sets on their own. Crucially, the use of opportunistic data improved the performance of the model based on formal survey data. The authors attributed these improvements to the sheer sample size of opportunistic data and their broader extent compared to survey data, both spatially, but also in terms of environmental variables.

A central theme in integrated SDMs is the idea of *complementarity* in achieving spatial breadth and depth. In most situations of data-collection, logistic and budgetary constraints mean that we need to settle on trade-offs between the resolution and the extent of surveys. For example, given a fixed amount of ship-time, covering a greater area at sea necessarily means using sparser transects (i.e. either increasing the distance between successive observation points, or increasing the spacing between transect lines). Similar trade-offs exist between different types of data. For example, opportunistic data tend to have greater sample sizes but lower accuracy and precision, compared to formal survey data. Several authors (Pacifici et al. 2017, Nelli et al. 2019) have now pointed out that, by integrating different surveys and different data types into one analysis we do not merely achieve an increase in sample size, but a complementary use of the different spatial extents and resolutions that characterise these data. Complementarity means that detailed features of species distributions can be embedded in big-picture data, even where such details have not been directly observed.

Spatial data integration need not only be used with data that inform the same latent surface. For example, hurdle approaches that combine abundance conditional on occupancy (or occupancy conditional on abundance) have traditionally been implemented as two-stage analyses (Waggitt et al. 2019). However, these can be easily implemented as integrated analyses, as in Clark et al. (2019).

5 Challenges and opportunities in multi-survey modelling for seabirds

Some of the challenges faced in seabird modelling are common to all SDMs. Errors in the observation of usage data (Section 5.1) and imprecise, or missing, environmental data (Section 5.2) can plague any analysis. Equally common to all analyses, although less frequently discussed, is the issue of model transferability (Section 5.5), the task of using insights gleaned from one place at one time, in other spatiotemporal frames. Beyond these common problems, the natural history of seabirds poses unique challenges to distribution modelling that are accentuated when trying to analyse multi-survey data. The high mobility of individuals, the stochastic nature of their local spatial distribution with ephemeral aggregations of large numbers at temporally transient foraging opportunities, combined with their association with breeding colonies for parts of the year gives rise to complex processes of density dependence between individuals, colonies and species (Section 5.3). These processes may manifest spatially, and the detail used to describe and estimate them from spatial data generates much of the complexity (non-linearity) in seabird SDMs. Model complexity is compounded by the need to deal with (and take advantage of) spatial and temporal autocorrelation in multi-survey data (Section 5.4). Therefore, the key issue in deriving appropriate frameworks for seabird multi-survey data is navigating the sources of computational complexity to achieve a workflow that makes the most of the data but reaches results in plausible times (Section 5.6).

5.1 Imperfect observation and multi-survey modelling

Different survey methodologies (boat-based, visual aerial and digital aerial) are affected by different types of biases and imprecisions. The behaviour of seabirds can amplify the differences between survey methods. For example, the strong attraction of scavenging seabirds to boats at some times of year but not others, and in some regions but possibly not in others can caused strong fluctuations to detectability (potentially even boosting it above 100%). These effects need to be explicitly accounted for during analysis. In general, observation error types found in survey data comprise false negatives, false positives, effort imbalances and location error (Miller et al. 2011, 2019, Hefley and Hooten 2016).

False negatives and false positives

False negatives involve errors of omission and misidentification, whereas false positives are predominantly due to double-counting or misidentifications of one species as another. There is some debate about the relative importance of these two types of error. For example, (Tyre et al. 2003) argue that false negatives will be less frequent for two reasons: First, misidentifications are conditional on a detection happening, and their frequency can therefore be reduced by observer training and improvement in survey protocols. Second, it is common practice in several designs to not record a detection (or to exclude it from the analysis) if there is any doubt about its identity – hence converting a potential false positive into a potential false negative. However, it could be argued that observer training could be as beneficial for detection as it is for identification. Equally, detections can be recorded and analysed in conjunction with a degree of certainty for their correct identification (Miller et al. 2011). Therefore, it is perhaps ideal to proceed with the assumption that analysis frameworks will need to deal with both false positives and negatives (Miller et al. 2011).

In the setting of transect surveys, both false negatives and false positives relate clearly to the baseline detection probability (the intercept of the model). Most obviously, false negatives give the impression that an organism is less prevalent than it actually is. But even false positives can create consistent biases. If two species are easily mistaken for each other, then the bias will be positive or negative depending on whether the true prevalence of the focal species is respectively smaller or larger than the prevalence of the non-focal species. In logistic habitat models, focusing purely on occupancy, they have an even greater effect on the slope of the relationship with covariates (Tyre et al. 2003). In such studies, it is suggested that at least three repeat visits are required to correctly estimate the probability of omission. Furthermore, a statistical trade-off between survey extent and survey overlap is expressed. Simulated results indicate that when the rate of false negatives is low (e.g. <50%), it may be better to increase the extent of the surveys rather than the number of visits. As false-negative rates increase, the variance of parameter estimates is reduced more by increasing the number of visits, especially when the

overall extent is low. There are several reasons for being careful about taking these quantitative recommendations immediately on-board in developing best practice for seabird surveys. First, the probability of detection will generally not be the same in different seabird surveys. Second, these recommendations were based on the assumption of near homogeneity in distribution. Third, occupancy data analysed by logit or probit models are considerably more vulnerable to observation errors and therefore more limited in their inferential abilities than models of abundance. Nevertheless, similar effects have been recovered in broader classes of models, including the Poisson point process (Lahoz-Monfort et al. 2014).
Discussion of errors with higher relevance to seabird surveys can be found in commissioned reports. For example (Camphuysen et al. 1995), carried out an inter-calibration exercise for 20 observers on ten ships for ship-based ESAS surveys in North Sea. The authors reported major differences in detections from different observer teams. For example, team A reported about 6-10 times more kittiwakes than teams B, C and F in the same region and time period. Anecdotally of relevance is the misidentification of guillemot and razorbill with high variance in the ratio reported by different observers. Equally important is the practice of assigning detections to group rather than identifying them to species. For example, divers are often recorded as 'diver species'. These can be apportioned to species later, according to the ratio of red-throated, black-throated and great northern in the sample that were identified to species. However, great northern is easy to ID to species whereas many black-throated divers are difficult to separate from red-throated divers. So, apportioning 'divers' to species based on proportions ID to species tends to result in overestimating numbers of great northerns.

These are real problems where understanding the natural history might help resolve some of the biases, whereas others (especially observer effects) can't easily be taken into account (ESAS has too many individual observers for example so quantifying observer effects would not be statistically feasible).

Location errors

Errors in the detection, identification and location of individuals may vary temporally (see example of diurnal and seasonal patterns along distance sampling transects in (Furnas et al. 2019)), according to environmental conditions (such as weather and ambient light) or according to geomorphology (Frair et al. 2010). Note that, contrary to some practices (e.g. Oppel et al. 2012) when the detection probability varies with environmental conditions, the counts cannot be used as a relative index of abundance without a correction to the effective detection distance and the baseline probability of detection, or much better, incorporation of these influences as of

detection in the distance sampling analysis (Marques and Buckland 2003, Buckland et al. 2008). In general, both the probability of detection and the abundance of a species will depend on environmental covariates. For example geomorphology (e.g. proximity to land) will affect both, bathymetry will affect only abundance and ambient light conditions may affect only detection. This gives rise to potential problems of identifiability (our ability to pinpoint estimates for coefficients pertaining to both detection and species abundance). The statistical requirements for achieving identifiability have been examined by (Lele et al. 2012). For occupancy models, these authors concluded that if the intersection of the two sets of covariates contained a non-categorical variable, and if there was at least one variable that belonged to one set but not the other, estimation was identifiable.
This debate about detection errors is particularly pertinent to the analysis of multi-survey data. For many situations, it is impossible or logistically unlikely that more than one visit to any location will take place, and frameworks must be developed to ensure that detection errors can be ameliorated by use of the habitat similarity in non-overlapping visits (Lele et al. 2012) . However, in the presence of overlapping surveys, progress can be made more directly to account for detection errors as part of inference (Bolker 2008). In repeated visits, the target species is recorded as being detected or not detected at each visit. At locations where the species is present, detection error will occasionally result in species not being detected even though it is present at the site. Assuming the true abundance does not change over the duration of repeated visits fluctuations in observed abundance at a particular site can be attributed to detection error and thus facilitate the estimation of the baseline detection probability in a survey. Although not widely available for abundance data, analytical expressions have emerged that describe how repeat visits increase the statistical power in the estimation of detection probabilities (Guillera-Arroita and Lahoz-Monfort 2012).

Variable observation effort

Heterogeneity in observation effort is a crucial driver of the point patterns presented by the raw data (Manly 2003, Miller et al. 2019). In particular, (Chakraborty et al. 2011) propose a distinction between the *potential* intensity surface (representing the biological process generating occurrences of a species in space) and the *realised* intensity surface (which is curtailed by the distribution of observation effort). This is more widely known as a thinned Poisson point process because it incorporates an observation model into the intensity function (Hefley and Hooten 2016), so that the point patterns recorded are sparser than what would be seen if observation effort saturated the whole of space. Such combinations of the observation model with the underlying biological model offer the potential of fully integrated inference, the idea of conducting the estimation of detection functions simultaneously with the habitat analysis. This would allow the forward and back propagation of errors as discussed in the previous sections, so that the estimation of detection distances would benefit from the information on habitat. A key assumption in distance sampling is that the distribution of the species in the detectible vicinity of the transect is uniform. Violations of this assumption could be prevented by modelling the distribution of the species (as driven by covariates) jointly with the detection function associated with the transect (along with its own covariates). This approach was taken by (Nelli et al. 2019) in their study of malaria incidence patterns. In that paper, a modification of point transect distance sampling (appropriately tailored to epidemiological data) was embedded in a model for the covariates of incidence, hence improving inference for both the observation and biological model.

5.2 Imperfect habitat data and multi-survey modelling

Imperfections in the covariate data can arise in different forms. The resolution of different variables or different regions of the same variable may be mismatched. Covariate data may be partly, or wholly missing. Alternatively, the values of known covariates may be measured with uncertainty, or be temporally fluctuating. The idea of data integration from multiple surveys and data types has a crucial role to play in using such imperfect data.

Mismatched scales

Mismatches between explanatory variables are informally addressed by some process of alignment, whereby a common reference grid is applied to all the covariates, usually involving a process of linear interpolation to shift the centre-points of existing cells to the new grid nodes, and also a process of coarsening of the resolution of maps to the lowest resolution available in the data (Kent et al. 2006). We will call this the *lowest common denominator* approach to mismatched scales. None of these steps are necessary if the data generating mechanism is modelled as a point process. In this case, covariate values are extracted from the available maps, in their native resolution, at the location of each point in the response data. Without a doubt, different scales will have characteristic impacts on the results of the SDM (Levin 1992, Paton and Matthiopoulos 2018, Pacifici et al. 2019), but at least in this way, explanatory data are used at the finest resolution available and in a minimally processed form.

Mismatches in different geographic regions for the same explanatory variable are a problem that requires treatment because it affects the consistency with which a variable is allowed to affect the response in the model. The key decision is between downgrading the high-quality areas (an easy but somewhat destructive solution), or upgrading the low-quality areas. Upgrading could be approached by geo-statistical interpolation, as a pre-modelling stage of the analysis. Specifically, a method like kriging could be used for the coarse resolution regions (e.g. (Monestiez et al. 2006)). Kriging relies on an estimated object (the semi-variogram) which captures the characteristic spatial autocorrelation in a variable. By estimating the semi-variogram in the fine resolution regions and applying it across space, the coarse scale region could be resolved. This will crucially depend on whether the coarse scale readings are localised (albeit sparse) measurements, or averages over coarse scale cells. The latter scenario would preclude the geo-statistical interpolation approach.

A much more general and flexible approach to scale mismatches involves the use of hierarchical Bayesian approaches to resolve coarser regions into finer scales (Keil et al. 2013). Such approaches could also be used to increase precision when multiple layers of information exist at different scales for the same environmental variable.

Partially missing covariates

There will be situations where the desired spatial extent of the region to be used for modelling is not spanned by available covariate layers. Similarly, cloud cover or other obscuring influences will result in gaps in remotely sensed layers. Once again, the most direct approach (the lowest common denominator approach for partially missing covariates) is to retain a minimal data set that comprises only cells with complete covariate entries. This invariably leads to heavy censoring of the data either via cells being dropped, or via covariates being judiciously removed to try and retain more cells. It is possible to formalise this process of censoring by starting with a reduced set of cells and a full set of covariates, begin model fitting and drop some variables through formal model selection. As the set of covariates in the model is reduced, the hope is that more cells will be able to be once again included in the analysis. This pragmatic approach is not ideal because it involves information loss through censoring, the degree of which is not entirely in the analyst’s control. Instead, interpolation methods (i.e. density estimation methods, see Section 3.1 or geo-statistical methods, see previous subsection) can be used to reconstruct the expanses of missing data.

A particularly relevant problem in this category arises from Camphuysen et al. (2004) recommendation that to enhance the cost-effectiveness of ship-based surveys, vessels should be equipped with an Aquaflow (logging surface water characteristics including temperature, fluorescence (chlorophyll), and salinity information simultaneously with species abundance). The idea of contemporaneous recording of environmental variables at high resolution is certainly appealing and can lead to models of high explanatory power. However, this approach is limited when spatial predictions are required for areas not visited by the boats, because high-resolution environmental data are unlikely to be available for the whole of space. Interpolation of the environmental variables between transect lines can solve this problem, assuming that the spatial autocorrelation in the relevant environmental data does not decay rapidly compared to the separating distance between the transects. If that is not the case, then it may be necessary to combine the measurements of variables along the transects with synoptic raster data at coarser resolutions. That brings the problem back to the class of scale mismatches (see previous subsection), whereby integration of environmental data that differ in resolution and extent is used to reconstruct uninterrupted layers of explanatory variables. A key consideration in this is the ability to predict from such models. If the Aquaflow variables are dynamic, then these data will not be available outside the temporal window of measurement. In such cases, forecasts will be impossible (see subsection on Variability and Errors in Measurements, below).

Once again, all of these processing steps can be incorporated into the main inferential framework via hierarchical Bayesian approaches to resolve scale mismatches and data absence (Keil et al. 2013, Nelli et al. 2019).

Unknown covariates

Missing covariates is a widely recognised and difficult to diagnose source of estimation bias (Barry and Elith 2006, Fieberg et al. 2018). Although it may not be possible to reconstruct multiple missing covariates (such a data vacuum is beyond the reach of even the best statistical model), it is possible to capture the collective effect of missing covariates in the residual autocorrelation of the fitted model as in (Beale et al. 2014, Nelli et al. 2019), and also see discussion in Section 5.4.

Variability and errors in measurements

Dynamic explanatory variables may follow seasonal (e.g. monthly average temperature), diurnal (e.g. tide) or less predictable (e.g. weather) fluctuations. Although there is certainly a biological appetite to include such variables to explain the distribution of a species, data availability can be a constraining factor in this. At the model fitting stage, inclusion of dynamic variables means that each observation of species abundance must be synchronised with contemporaneous environmental data. This can be difficult because data absence becomes more pronounced as data

are subdivided into smaller time frames. However, problems are more pronounced at the prediction stage. Forecasting species distribution into the future requires us to have complete layers of explanatory maps, which must also be forecasted, if they refer to dynamic variables. This calls for some consideration on which variables to include in a dynamic format. If the spatio-temporal availability of data for an explanatory variable (either from historical data or from available forecasts) is low, then even if there is biological confidence about their importance, it may still be necessary to exclude them from the analysis.
Explanatory variable data may contain measurement biases and imprecisions which may also be spatially correlated (Barry and Elith 2006). Known, consistent biases are often corrected at the stage of pre-processing of covariates. Imprecisions, especially if these vary across space, should be propagated through the analysis so that they are reflected in the final confidence intervals of the SDM parameters and spatial predictions on the distribution of the species (Barry and Elith 2006). This idea has its roots in type II linear regression (Sokal and Rohlf 1995), but can be considerably updated by using the flexible modelling structures of hierarchical Bayesian models. Spatially correlated errors can arise from reconstruction methods (see interpolation methods mentioned earlier) that are often used to generate complete layers of covariate information. Theoretically, the approach for propagating such errors to the final results is similar to the approach of propagating un-correlated uncertainty, with covariance structure for neighbouring locations. However, this is, as yet rarely done for computational reasons.

All these suggestions about uncertainty quantification eventually need to be visualised in conjunction with median predictions. Informatively mapping uncertainty is not straightforward however and communicating uncertainty to policy makers can cause confusion. Interesting approaches to using uncertainty in SDMs for informing policy decisions are actively developed in the field of ecological reserve design (Tulloch et al. 2013).

5.3 Accessibility and density dependence in seabird distribution modelling

For at least some parts of the year, species of seabirds are central place foragers. Their use of different marine locations is therefore likely to be affected by how accessible these locations are from the breeding colony. Accessibility might vary by species and season. The effects of accessibility on distribution have been anticipated theoretically (Matthiopoulos 2003a, 2003b) and found empirically in marine central place foragers such as seabirds (Lewis et al. 2001, Wakefield et al. 2011, Grecian et al. 2012, Thaxter et al. 2012, Waggitt et al. 2019) and pinnipeds (Matthiopoulos et al. 2004, Aarts et al. 2008, Jones et al. 2015). In addition, the coloniality of seabirds leads to foraging aggregations and potential resource depletion in the regions surrounding the colonies (Lewis et al. 2001). Ultimately, the use of particular locations at sea will be determined by the trade-off between commuting costs (as shaped by accessibility) and foraging benefits (as shaped by environmental resources and depletion).

Both accessibility and depletion/interference may be thought of as functions of distance from the colony, but they are complex, highly non-linear processes for distinct reasons. Accessibility is mainly complicated by the fact that different colonies will be placed in locations that are variably affected by the coastline. Hence, colonies on a small island are likely to be unconstrained in every direction, colonies on a relatively straight coastline will only have a semicircle of marine directions available for departure, while colonies in an inlet may be limited to a single water body route into the sea. To capture declines in accessibility with distance, it is possible, as a first approximation, to introduce a distance-decay function, parameterised identically for different colonies (Matthiopoulos et al. 2004, Grecian et al. 2012). However, the fact that the available area of water around each colony will depend on coastal morphology, means that the resulting marine distribution from such a function would not allocate equal numbers of birds at units of area that are the same distance from different colonies.

These behaviours will not be independent of age structure. Seabird populations include a high proportion of immatures that are less competitive than adults so may tend to distribute at sea in areas away from colonies (at least until they start to seek to recruit into a colony themselves). So, we can expect some ‘infilling’ of marine areas away from large colonies by immatures, especially younger age classes of immatures.

Density dependence is complicated initially by resource competition between colony members (Lewis et al. 2001). As the size of the colony grows, individuals need to travel further to escape the density dependent effects of depletion or interference. At the same time, there is evidence that usage from neighbouring colonies can saturate space leading to the appearance of home ranging behaviour at the colony level (Wakefield et al. 2013). This implies that even without the constraints of commuting costs, a colony might not extend its foraging range (and, consequently, its population size) indefinitely. In addition to inter-colony competition with conspecifics, it is possible that individuals from a colony are experiencing competition from neighbouring colonies of other species. In this case, the resulting asymmetries in range will not only be due to relative colony sizes, but also due to trophic niche overlap and competitive dominance between species. All of the above aspects of biology will interact with each other and with environmental productivity. For example, for a given colony size, colonies that are obscured by coastline formations may tend to have greater ranges than island colonies because density dependence is acting over smaller areas close to the obscured colonies.

Non-linearity in the relationship between abundance and its covariates is widely recognised in the statistical literature as well as in the seabird-related literature (e.g. Section 1.1.2 in Oedekoven et al. 2012b) and has prompted the development of semi-parametric, non-linear models based on splines or additive basis functions. These models, broadly known as Generalised Additive Models (Wood 2006) allow the data to “speak for themselves”, i.e. to inform the shape of the relationship with explanatory variables. The greatest advantage of GAMs and their extensions (such as mixed-effects GAMS, or GAMMs), is that their estimation remains embedded in the frameworks of the linear model. Their greatest disadvantage is that they can result in spurious relationships via overfitting. Instead, in some cases we can argue for the derivation of a functional form from biological first principles (see discussion on mechanistic models in Section 3.1). The sophistication that is used for modelling accessibility and density dependence will determine the computational feasibility of the approach (see Section 5.6 below), but we can outline here an idealised approach that would correctly account for both effects. This hypothetical approach would have the following features:

1) **Accessibility** is treated as a distance-based function only. Simpler approaches would use Euclidean distance, but improved approaches would consider distance measures based on at-sea travel (see biological distance in (Matthiopoulos 2003a)). These distances would need to be calculated as spatial layers, specific to each colony. More elaborate measures of accessibility, incorporating landscape resistance (e.g. due to prevailing wind fields) could be calculated (Zeller et al. 2012, 2017). For any given marine location, these metrics could result in asymmetric outward and homeward distances.
2) **Intra-colony competition** may be treated by including seabird density as an auto-covariate (Augustin et al. 1996) in the model, effectively using the response variable *in the neighbourhood* of a point as an explanatory variable for the response variable *at the location* of the point. The apparent circularity of this step requires careful consideration of the model fitting stage, but it is mechanistically equivalent to density dependence (i.e. the feedback effect of a population onto its own distribution). An alternative approach is to introduce an explicit autocorrelation term in the response, but this would need to be made colony-specific (and ideally, colony-size specific).

3) **Inter-colony competition** may be accounted for via simultaneously using the expected usage of one colony as an explanatory variable (possibly with a negative coefficient) for the expected usage of another colony, and vice-versa. The need to couple the usage of different colonies, gives rise to another point of apparent circularity, which can be addressed via simultaneous regression techniques (i.e. modelling multiple response variables at the same time, to allow the use of each response variable as an explanatory variable for the others). The coefficients of these simultaneous regressions would need to be asymmetric, based on the size of each colony (so that large colonies have stronger negative effects on the marine use of smaller colonies).
4) **Inter-colony, interspecific competition** could be dealt with in exactly the same way as intraspecific competition between colonies (see 3, above), allowing for asymmetric effects due to differences in species, as well as differences in colony size.

As we will also discuss in Section 5.6, a model with the above specifications is just about possible to write and simulate from, but not computationally feasible for fitting to data. Instead, a more pragmatic approach would involve the following simplifications:

1) **Accessibility:** Write accessibility as a distance-decay function. Any measure of distance can be used (e.g. Euclidean, travel-time, landscape resistance), because these calculations are not part of model fitting.
2) **Intra-colony competition**: Expand the decay parameter of the previous step, so that it is larger for smaller colonies and vice-versa. This means that as colony size increases, the decay of expected usage with distance from the colony slows down, hence extending the range of the colony to take account of intra-colony density dependence. The function may be allowed to be non-monotonic, so that total usage initially increases with distance, and then eventually starts to decay.
3) **Inter-colony competition:** The formulation from the previous step, can be extended to account for the effects of other colonies but in this case, the effect on a the usage of a given marine point by a focal colony will depend on the distance of that point from the competing colony (as well as the competing colony’s size).
4) **Inter-colony, interspecific competition:** Although biologically, the differences between species are important, from a mathematical point of view, all that is required to capture these effects is a re-parameterization of the model from Step 3, above.

In Section 2.2 of the Vignette that accompanies this report, we develop, graphically explore and spatially simulate from such a model. The resulting model is non-linear but it has modest parameter requirements. For example, for a four-species system, where each species is represented by several colonies, a total of 40 parameters would need to be estimated for all species. If the focus was on a single species with three competing species, then ten parameters would be required. If no inter-specific competition was considered, only four parameters are needed to capture effects of accessibility and density dependence (i.e. Steps 1-3, above). The important benefit from such an approach is that we can learn from the data about the strength of accessibility and density dependence for a particular species. Fitting this model simultaneously with environmental covariates allows us to account for the confounding between intrinsic and environmental regulation of spatial usage.

5.4 Autocorrelation and multi-survey modelling

Informally, residual autocorrelation (RAC) is the spatial or temporal similarity in the direction and magnitude of discrepancies between model and data. The assumption of all statistical regression models (including SDMs) is that all sources of autocorrelation (AC) in the data are the result of autocorrelation in the available explanatory variables[1], and, therefore, that *conditional* on the covariates included in the model, the residuals of the model are independent. This assumption, however, is often violated and the existence of spatial and temporal AC in the residuals casts doubts over the results of inference (Segurado et al. 2006, Dormann et al. 2007). Several methods exist for dealing with the problem (Dormann et al. 2007), but they differ in one crucial aspect: whether they correct the consequences of AC on model dispersion and parameter standard errors, or whether they model AC explicitly. In this section, we will argue that modelling AC explicitly can lead to capital information gains in the analysis of data from multiple surveys.

Spatial autocorrelation

RAC can result from three routes. First, if an influential covariate is missing from the analysis, the fitted values will under-/over-estimate true usage in a spatially aggregated way. Second, even if all relevant covariates are included, they may be misrepresented in the analysis. For example, covariates that have been created by interpolation may inherit over- or under- smoothing to the model's estimates. Over-

[1] More complex approaches, such as state-space models (Buckland et al. 2004, Patterson et al. 2008, Newman et al. 2014), capture additional dependencies by conditioning multiple response variable to each-other.

smoothing is equivalent to using coarse level environmental data to explain responses that are happening at finer spatial scales (de Knegt et al. 2010). Alternatively, it may be that animals are responding to the broader context rather than the finely resolved environmental data offered as explanatory variables (Paton and Matthiopoulos 2018). Third, even if all relevant covariates have been included at the most appropriate scale, there is still the possibility that clustering in usage is caused by intrinsic processes, such as individual memory (Van Moorter et al. 2009), limitations in movement speed (Aarts et al. 2008), or social cues (Riotte-Lambert and Matthiopoulos 2019). In the broadest sense (Beale et al. 2010), spatial AC can be captured either as a covariate , or by introducing a spatially auto-correlated structure in the model's error. Including AC as a covariate could be done by introducing flexible functions of latitude and longitude (Mendel et al. 2019), or by using density as a local auto-covariate (Augustin et al. 1996). Including AC as an auto-correlated error can be one of several components of the error structure of a hierarchical model.

Often, one of the symptoms of un-modelled heterogeneity generated by RAC is over-dispersion in the residuals. For example, omission of influential covariates is a well-known source of over-dispersion. In seabird analyses, this is frequently addressed by use of over-dispersed distributions (e.g. Lieske et al. 2014) or zero-inflated mixture models (e.g. Oppel et al. 2012, Waggitt et al. 2019). Such approaches, however, assume that the residuals in nearby locations are independent, an assumption that is unlikely to hold. There are therefore arguments for the use of flexible spatial functions in capturing residual spatial complexity (e.g. Sections 1.1.3-4 in Oedekoven et al. 2012b) and, given that formal tests for autocorrelation in residuals can be weak, there is a view that modelling autocorrelation by one of the methods outlined in (Dormann et al. 2007, Beale et al. 2010) should be carried out pre-emptively to avoid biasing the estimated coefficients of available covariates (Fletcher et al. 2016).

However, although in most cases we can be certain that one or more sources of spatial autocorrelation are operating, it is not clear how to interpret the results of explicitly spatial models, and indeed it is not certain that the inclusion of such terms exactly counteracts the causes of the problem. Interpretation of autocorrelation terms is challenging because the processes generating autocorrelation operate at different characteristic scales (Levin 1992) and the observed residual autocorrelation in a species distribution model is the result of all of these influences acting together. It is, therefore, not always easy to interpret autocorrelation patterns biologically or to attribute them to estimation artefacts. The bias-correcting effect of AC terms is also in doubt, because they can be shown to be confounded (and hence, potentially in competition) with fixed-effects terms (Hodges and Reich 2010). In some cases, incorrect specification of spatial errors can lead to severe biases in parameter estimates for fixed effects, even if all the relevant covariates have been included in the model (Beale et al. 2010, Sinclair et al. 2010).The outcome of this competition between spatially structured random effects and covariates, depends on the penalty imposed on the flexibility of random effects. Low penalties encourage over-parameterised spatial terms that diminish (i.e. negatively bias) the coefficient estimates for covariates. This underlines the need for out-of-data validation of fitted models.

Temporal autocorrelation

Temporal autocorrelation causes residual similarity due to temporal proximity between observations, essentially two successive snapshots of a system are not very different from each other because the system has not had enough time to diverge from its original state. This is considered an important violation of the independence assumption, particularly for telemetry studies (Fieberg et al. 2010), but it is less of a concern for surveys, assuming that design recommendations are adhered to, to avoid double counting (see Section 3.2). However, temporal dependence offers us valuable opportunities for exploiting multi-survey data that have been collected at different times. Currently, investigation of multiannual trends and relative changes in usage is usually considered to lie beyond the scope of seabird distribution analyses (Perrow et al. 2015), but such features would be essential if multiannual survey data were to be correctly integrated. Assuming stationary values for the covariates, the counts from two surveys conducted over the same region at different points of time should be expected to be more similar, the closer the surveys were in time. We can add temporal autocorrelation via the fixed effects of a model (as flexible functions of time), or we can assume an autoregressive model (e.g. a random walk in the measurements of residuals). More complex error structures spanning several time lags or continuous time are also possible (Wood 2006).

Spatiotemporal autocorrelation for multi-survey data

In the setting of multiple surveys spatial and temporal autocorrelation can be a real asset. Although habitat and species distribution models are predominantly visualised as maps, most of these models are fitted in environmental space, and are, therefore, neither explicitly spatial nor temporal (Chakraborty et al. 2011). Hence, a classic SDM is unable to recognise the fact that two contemporary surveys that overlapped spatially share more than just the same values for the covariates. Furthermore, within the range of autocorrelation, a model should be able to acquire additional support from the fact that even when two surveys do not exactly coincide in time and space, they can share similar information depending on their spatiotemporal proximity. Therefore, particularly for multi-survey studies it is important to allow environmental, spatial, temporal and spatiotemporal dependencies (Hothorn et al. 2011). However, it should also be recognised that these will not be able to bridge over large distances and intervals. For this, we require transferrable models that can address the non-stationarity of animal responses to different environmental compositions.

5.5 Model transferability

SDMs based on survey models could be asked to perform one of three predictive tasks, presented below in increasing order of difficulty.

Spatial and temporal Interpolation

The most reliable type of SDM prediction deals with the spatial and temporal frames within which the survey data were collected (a task we will call spatio-temporal interpolation). These applications make two, relatively easily satisfied assumptions. First, that the resolution of the surveys is finer than the scales of all autocorrelations used to describe spatial and temporal dependencies in the model. Second, that the coefficients estimated by the models are stationary in space and time. When stationarity is in the temporal dimension, this later assumption is often called the pseudo-equilibrium assumption (Guisan and Zimmermann 2000). If the windows of observation are small and sufficiently densely surveyed, these assumptions are satisfied and the SDM can provide an accurate representation of density at any point in space and any instant in time within the observation ranges (Isojunno et al. 2012).

Environmental interpolation

The SDM models recommended in this report are empirical in nature (albeit with elements of mechanistic modelling – see Section 5.3). Robustness (i.e. precision and accuracy) of such models under prediction is subject to the dangers of extrapolation, particularly when regression is equipped with high functional flexibility (Bell and Schlaepfer 2016). Non-stationarity in environmental responses is a recognised source of variability (Hothorn et al. 2011) and a safe route to unsuccessful model predictions. The difference between explanatory and predictive models is a key area of interest in an discipline such as ecology, which focuses on the consequences of spatial non-stationarity and change over time (Yates et al.

2018). However, for habitat models, spatio-temporal extrapolation need not necessarily be equivalent to environmental extrapolation. A high priority for empirical modelling in general is that sampling effort covers as wide a range of covariate values, and combinations thereof (Oedekoven et al. 2012b). Conceivably, this can be achieved in a single survey, but may be difficult given the logistical constraints on spatial extent (i.e. it may be difficult to survey highly contrasting environments without covering prohibitively large distances). Nevertheless, it may be possible to achieve by pooling multiple surveys together. Matthiopoulos et al. (2011) discuss how a survey can be considered as a single *sampling instance* in a pooled data set. That paper first formalised a generalised model for functional responses in habitat selection (known as a Generalised Functional Response - GFR). Functional responses broadly describe changes in use of an environmental variable as a function of the value of that and other environmental variables (Arthur et al. 1996, Mysterud and Ims 1998, Boyce et al. 1999, Holbrook et al. 2019). For example, in trophic ecology, a functional response describes the consumption of prey by a single predator, as a function of the prey's abundance (Holling 1959). Generalisations of the Holling concept of functional responses to multiple prey give rise to multispecies functional responses (Asseburg et al. 2009, Smout et al. 2010). In a similar way, a GFR describes how preference for a particular habitat changes in response to the availability of *all* habitats in the environment of an individual, or a subpopulation (e.g. a seabird colony). Such models have been shown to bring considerable gains in predictive power for environmental scenarios that are within the range of environmental values recorded in the pooled data (Matthiopoulos et al. 2011, Paton and Matthiopoulos 2018, Holbrook et al. 2019).

Environmental extrapolation

As soon as we require predictions for scenarios found outside the previously observed spatiotemporal and environmental frames of reference (i.e. the environmental profiles used for model-training), several processes begin to interfere with our predictive ability. Sinclair et al. (2010) identify no fewer than ten major biological and methodological obstacles to the success of such predictions. Examples include the appearance of unprecedented environmental domains and concurrent alterations in the community context in which the focal species is embedded.

Although functional response frameworks have been found to increase the robustness of predictions outside the extremes of observed scenarios (Matthiopoulos et al. 2011), it is nevertheless not clear how far they can be pushed. Indeed, there is currently no formal method for measuring the degree to which a particular prediction scenario is one of environmental interpolation or extrapolation.

The challenge of environmental extrapolation is particularly serious for forecasting (Sinclair et al. 2010, Tuanmu et al. 2011), which happens to be the main objective of anticipatory modelling in modern ecology. Considering even more challenging problems such as evolution and adaptation makes it clear just how under-equipped statistical SDMs are in dealing with extrapolation. Arguably, increasing the mechanistic content of SDMs (see Section 3.1) increases predictive ability (at the risk of model misspecification). Hence, there is now a clear tendency in the literature to consider species distributions in the backdrop of the population dynamics of a species (McLoughlin et al. 2010, Ehrlén and Morris 2015a, Matthiopoulos et al. 2015, Turlure et al. 2019), and also in the context of whole ecological communities (Ovaskainen et al. 2010, Calabrese et al. 2014, Distler et al. 2015).

5.6 Computational efficiency

Non-linear and spatial effects

The use of spatially autoregressive models, particularly in combination with non-linear predictors in an MCMC context, is computationally very expensive and shortcuts are necessary (e.g. the "covariate" model in Pacifici et al. (2017), or the list of four methods cited in Section 5.1 of Chakraborty et al. (2011)). For the particular approaches discussed in this report the key challenge lies in estimating biologically interesting and important parameters (pertaining to density dependence, inter-colony effects and inter-specific interactions in distribution – see Section 5.3) at the same time as dealing with spatially and temporally structured residuals. The methodological literature on models with spatiotemporal structure has been revolutionised by approximate Bayesian methods that either deal with fully non-linear models (as is the case with Approximate Bayesian Computation - ABC - Beaumont 2010, Csilléry et al. 2010) or deal with linearised versions of these models (as is the case with Integrated Nested Laplace Approximation – INLA - Martino and Chopin 2007, Rue et al. 2009, Lindgren 2015, Bachl et al. 2019). ABC methods have yet to meet with broad application in SDMs, so there is limited understanding of how to implement them in that setting. On the other hand, INLA methodology was developed initially for spatial point process modelling, so it is ideal for the purposes of SDMs. The question therefore is how best to linearise the models for combined seabird surveys.

The models implemented in the joint survey R library accompanying this report illustrate the MCMC implementation of both of these features but are very difficult to use for extensive spatio-temporal predictions. Nevertheless, the ability to fit these models implies that we can get good posterior estimates for the parameters participating in the non-linear parts of the models, while correctly accounting for spatiotemporal autocorrelation. This suggests three steps forward, in terms of computation: First, a pilot model-fitting exercise can be used within a limited spatial range to obtain posteriors for parameters in the non-linear components of the model (Section 5.3). This would follow the protocol of model SPATIAL in the attached JointSurvey R package and described in the accompanying vignette. Second, having obtained the posteriors for their parameters, these non-linear functions can then be considered as constructed covariates in a linear model (see further details in package vignette). Third, this linearised model can be fitted in INLA to enable wider inference for non-stationary processes (e.g. under a GFR framework) but, most importantly, to allow the treatment of spatial and temporal structures for inference and prediction.

An interesting development in this area is the ongoing extension of INLA under the INLABRU library (Bachl et al. 2019) to enable users to fit non-linear responses very similar to the ones outlined in this report. Pending computational evaluation of this facility, it would allow the workflow, in its entirety, to be ported into INLA.

Model selection

Models with high computational demands become particularly cumbersome if they need to be fitted repeatedly for the purposes of model selection (e.g. in order to derive likelihood ratios or information criteria). A particularly computationally expedient approach recently suggested by Renner et al. (2019) combines i) likelihood maximisation with ii) localised models of spatial effects (such as area-interaction models) and iii) shrinkage methods in-lieu of model selection. Likelihood maximisation is generally faster than Bayesian methods because its task is to find optimal parameter values, rather than to describe full posterior distributions for the parameters. Of course, this goes counter to the requirement for uncertainty propagation to the final results, but approximate measures of uncertainty can be obtained. Localised models of spatial effects capture the spatial dependencies that matter most, and therefore avoid calculations pertaining to negligible long-distance effects. Finally, shrinkage estimators reduce the value of parameters for non-influential explanatory variables to zero, hence performing the task of model selection without stepwise forward addition or backward elimination of variables.

6 Best practice for multi-survey analyses of seabird distributions

Ideally, the data used for robust spatial prediction of species distributions should be both high-resolution and spatially expansive. However, logistical trade-offs between spatiotemporal extent and resolution mean that such in-depth and geographically broad data are rarely available in practice. Instead, researchers need to piece together data from different places, times, or survey methods. Such integration presents several challenges (see Section 5, above), but it also offers remarkable opportunities. For example, data from different places and times, can allow us to increase the spatial extent of our maps, and our historical reconstructions, but importantly, they allow us to model the focal species under distinct and different circumstances, hence increasing the transferability of our model predictions. Also, if the survey designs are different (e.g. different resolutions and different field methodologies), simultaneous analysis has the potential to allow different surveys to effectively ground-truth (i.e. calibrate) each-other. We will approach the objective of this project in four stages.

The following recommendations build upon the closing section of Fletcher et al. (2019). We have arrived at these based on our literature review (see above), but also practical experimentation with realistic simulated data (see accompanying R library, manual and vignette).

6.1 Appropriate response and explanatory variables

Keep the highest-grade form of data

Even if occasionally true, the notion that occupancy models are more robust than models based on abundance can be misleading, since occupancy represents lower-grade information. Intentionally thresholding abundance data into presence and absence represents considerable information loss, precludes predictions of spatial distribution (instead, yielding surfaces for the probability of presence) and is therefore best avoided. If individual detections are available, these may be used in preference to aggregated counts (Section 5.1). Similar arguments apply to downgrading explanatory data (Section 5.2).

Analyse even low-grade data as if originating from abundance

Many data types may be curtailed at the stage of data collection. Citizen scientists may record species presence only once and transect surveys may aggregate counts of birds in each transect segment. Therefore, we may not have the option to analyse high-grade data, but this should not preclude us from modelling the underlying data-generating process as an intensity surface, referring to expected abundance. Treating this surface as latent, but common to all surveys in the pooled data set, enables the integration of multiple surveys and data types into a common statistical platform, a pre-requisite for pooled analyses.

Avoid inflated error structures until the end of modelling

Zero-inflated and over-dispersed data are the norm in spatial ecology. Often, this leads to hurdle analyses (e.g. modelling spatial occupancy first and conditional abundance second) or use of over-dispersed likelihood models (such as the negative binomial). However, the decision of whether this is an issue with a particular data set should not be taken a-priori. Modelling with covariates will generally explain some of that variability and use of spatially and temporally auto-correlated errors will account for unexplained hot- and cold-spots in distribution.

Partly missing covariates should not necessarily lead to data censoring

When parts of a spatial covariate layer are missing, the tendency is to curtail the data set, either by removing the covariate or by reducing the number of points to a subset for which complete covariate data exist. This may prove necessary in the end, however, it may be worth attempting to reconstruct the covariate either as a separate interpolation step, or as part of an integrated analysis with partially missing data (Section 5.2).

6.2 Treatment of survey design attributes and observation errors

Use distance sampling

Distance sampling techniques have a long pedigree in ecological surveys and facilitate the pooling of surveys with different protocols by reducing them into a common (if, numerically different) set of detectability characteristics (Section 3.2). The extensions of distance sampling that deal with transect design and the incorporation of covariates facilitate the correction of errors intrinsic in the observation process (Section 5.1).

Prioritise cross-calibration between surveys

Different surveys may rank differently in terms of their detectability (accuracy/precision) and spatiotemporal span. These qualities must often be tradedoff at the design stage. However, the joint analysis of multiple surveys, allows the combination of high detectability and high span (Section 5.1). Surveys for which the detectability errors have been quantified (e.g. by multiple observer platforms), should be prized highly in this process because they can be used within a joint analysis to cross-calibrate other, less detailed surveys that may have happened close in space and time. Such calibration may be shared hierarchically by all the surveys in the data, stepping-stone-fashion, depending on proximity to each other.

Consider state-space approaches

State-space approaches acknowledge both the dynamic nature of marine distribution data (Section 5.4), but also the importance of modelling complex observation processes explicitly (Sections 5.1 & 5.2). In this way, rather than "correcting" the observations for biases, prior to the formal analysis, a statistical observation model is included in the model likelihood to effect the necessary correction in an integrated way (i.e. together with parameter estimation). This has the advantages that both the biological and the observation models are tuned with regard to each other, and that uncertainty propagation from the observation model to the final predictions happens automatically. Although we have not reviewed this option extensively in this report, it will be worth considering as available software becomes optimised and the computation times of multi-survey models decline.

6.3 Treatment of space time

Use point process models

Point process models allow space-time to be modelled jointly and continuously. They can also subsume all other valid approaches to species distribution modelling (Section 3.1). Finally, they are compatible with other features of modelling developed to enhance predictive power (Section 5.5). Heterogeneous point process approaches are fast becoming the gold standard for spatiotemporal analyses, and their implementation in speed-optimised frameworks such as INLA has attracted a lot of interest from management practitioners.

Use autocorrelated structures

Spatially and temporally autocorrelated structures can perform a multiplicity of tasks. They can account for missing covariates (hence explaining residual over-dispersion – Section 5.4). They can also be used to impute gaps in covariate values (Section 5.2). However, most importantly, they can be used to communicate to models of pooled survey data information about the spatiotemporal proximity between abundance observations. In this way, even if exact replication is not part of the survey design, an indirect form of replication can be achieved. There are caveats associated with the implementation and interpretation of auto-correlated structures, and their use is far from automatic (Section 5.4). However, the rewards, particularly for multi-survey data sets are very high.

Take dynamics into account

The pseudo-equilibrium assumption for SDMs is difficult to justify in applications that require more than spatial interpolation in the time-frame of data collection (Section 5.5). For example, if we need to account for multi-survey data that include before-and-after control impact, it is important to account for temporal non-stationarity. In some cases, non-linearity in the habitat responses of a species can be captured by simple extensions such as statistical interaction terms in the linear predictors of models. In other cases, a more explicitly biological model may be required. Temporal autocorrelation structures (see above) are also helpful in this respect.

6.4 Accessibility and density dependence

Use realistic distance measures

For colonial species, accessibility and density dependence in spatial usage are most often represented as non-linear transformations of distance of points at sea from colony locations. Therefore, using appropriate distance measures is essential, if birds don’t transit between locations in straight lines. Depending on the species, if we are concerned that they avoid flying over land, or if due to glide-flight they rely on prevalent wind direction, it is important that we account for these effects in the measure of distance. This is particularly relevant for behaviours such as avoidance of anthropogenic structures, where birds need to circumnavigate. The distribution of usage may alter in the vicinity of structure but an SDM will be unable to capture the changes without an appropriate measure of distance.

In the present, Use abstracted models for density dependence

We consider that, currently, the computational demands of a fully spatially explicit model of intra-colony, inter-colony and interspecific competition are prohibitive for the purposes of applied SDMs. We have therefore provided an illustration (in the project vignette), of how a pragmatic model for these processes can be developed and incorporated into the linear predictor of an SDM. We recognise that such models are crude approximations of the truth, but even such relatively simple formulations are currently missing from most seabird SDM approaches.

In the future, consider spatially explicit models for density dependence

As computational approaches (particularly in the area of Approximate Bayesian Computation and Integrated Nested Laplace Approximation) become more widespread in the field of SDMs, it may become possible to model competition in a fully spatially explicit way. For example, INLA is already capable of modelling multiple, coupled response variables. This would allow the spatial interactions of different colonies to be captured as part of simultaneous regression where the distribution of animals from any given colony is allowed to affect and be affected by the distributions of members of other colonies and species.

6.5 Inferential Platforms

Use hierarchical models

Three important features of multi-survey models described above rely on hierarchical models. Specifically, cross-calibration of observation models, covariate imputation and latency and use of spatio-temporal proximity to allow the predictions to borrow strength from multiple surveys.

Use Bayesian approaches

Computer-intensive Bayesian model-fitting deserves attention because it is implemented in flexible software frameworks (such as JAGS or Stan), that allow state-space and hierarchical structures. More importantly, Bayesian inference permits the elicitation of expert opinion in the form of parameter priors. The expert knowledge on the attributes of field survey practices will prove invaluable at this stage for specifying parameter priors for the observation models.

Use Data integration

Although approaches for using multiple data sources could take the form of a comparison (e.g. so that predictions derived from an expansive data set are validated by use of a localised, high resolution set of data), this is a relatively weak approach that does not make best use of the combined data. The alternative approach of joint inference, whereby both data sets are analysed simultaneously to extract maximum power. These approaches are also particularly useful for extending the analyses to non-survey data (Section 7).

Fully propagate uncertainty to the final predictions

We have outlined methods that can usefully reconstruct or account for biases, imprecisions and autocorrelations in explanatory and response data, as well as coarse, misaligned, partly or wholly missing covariates. Such methods can go a considerable way towards 1) correcting predictions, 2) realistically representing inherent uncertainties and 3) increasing the spatial extent of model fitting regions, by allowing more of the data to be retained in (i.e. not censored out of) the analysis. However, there is always a limit to how much missing information can be statistically imputed and therefore some prudence may be needed in determining which variables to include in the analysis. This is best illustrated in the case of SDM forecasts that are based on dynamic environmental variables. It may be biologically known that a particular environmental variable is shaping the distribution of a species, but if that variable is not available for future predictions, then it will either need to be excluded from the original analysis, or its effect integrated out of the final predictions.

6.6 Computational platforms

Support open source

As a matter of process, all code developed by government funding should be made available to the scientific community. We have used R (R Core Team 2019) to develop the demo libraries for this project. It is a good idea to keep all functions within a single environment and push for the standardisation and quality control of these libraries.

Ensure strong interface with Geographic Information systems

Much of the effort in preparing for modelling goes into interfacing the analysis framework with the raw data. This would greatly be assisted by establishing stable protocols for data formatting, and by using the GIS functionality in R to keep all data processing on a single platform.

Parameterise non-linear model components with exact methods

We have used a flexible MCMC approach to implement autocorrelation structures and non-linear features of biology within statistical models for inference. The prototype models presented in the jointSurvey library are computationally greedy, but they have the best chance of retrieving the difficult parameters pertaining to density dependence and competition. The JAGS environment used here interfaces seamlessly with R, and therefore keeps model usage (if not model development) to the same platform.

Implement large scale predictions using fast approximate methods

To generate large scale predictions with the spatio-temporal autocorrelation features stipulated above, it will be imperative to move towards efficient methods, such as INLA. These can be deployed from within R and are therefore the next logical step for real-world applications. In addition, exchange of information between the JAGS models already enclosed in the jointSurvey library and the more efficient INLA models used for large-scale predictions would be both necessary and efficient under the proposed scheme.

7 Future extensions

7.1 The promise of integrated hierarchical models

The use of hierarchical approaches in SDMs (Keil et al. 2013, Hefley and Hooten 2016, Pacifici et al. 2017, Fletcher et al. 2019) follows the principles set out in previous sections of this report. In particular, it assumes a true but unknown underlying distribution (usually, the continuous intensity surface of the heterogeneous Poisson process), which is observed by one or more methods that may be incomplete or imbalanced in their spatiotemporal coverage, biased in consistent ways and imprecise in other ways. This separation between the biology (which forms the objective of statistical inference) and the observation processes that generate data from it, allows us to do two useful things: first, to allocate proportionate modelling effort to the formulation of the methodological imbalances, biases and imprecisions. This leads to error models that separate natural stochasticity (a biological source of uncertainty) from methodological artefacts. Second, it allows the use of multiple observation models for a single underlying truth. We have seen in this report that such complementary use of different surveys and, possibly different methods, can lead to benefits such as the cross-calibration of methodologies, the combination of high spatiotemporal extent and resolution, and the ability to reinforce predictions that are explicitly spatial and temporal.

From the perspective of seabird SDMs, hierarchical models can perform three types of integration. They can bring together data from multiple line transect surveys (as explored in this report and the accompanying vignette), they can combine survey data with distribution data of fundamentally different types , such as occupancy data from citizen science records (e.g. Keil et al. 2013, Hefley and Hooten 2016, Pacifici et al. 2017, Fletcher et al. 2019), but they can also combine distribution data with supporting information that underpins the analysis with more mechanistic principles. Such precedents of multi-data integration are considerably more developed in the area of population dynamics (Buckland et al. 2004, 2007, Newman et al. 2014, Zipkin and Saunders 2018). In this section, we briefly explore possibilities for data integration beyond the multi-survey context.

7.2 Multi-species surveys

Surveys at sea offer the opportunity to track multiple species. This is an alternative interpretation of the multi-survey idea, in the sense that the same platform provides multiple datasets. Interest in hotspots of biodiversity has led to the idea of stacking single-species SDM models (Calabrese et al. 2014, D'Amen et al. 2015, Distler et al. 2015). Although stacking is not an integrated analysis in the sense outlined in this report, it has been useful in demonstrating the magnitude and duration of seabird aggregations or partitioning in the open sea from both survey (Nur et al. 2011) and tracking (Jones et al. 2015, Grecian et al. 2016) data. However, an interesting research direction lies in allowing data sets from multiple species to gain strength from each-other. We outlined earlier how spatiotemporal proximity can be used to borrow strength by jointly analysing a collection of surveys that have been carried out within a defined geographic region and time window. The same idea could be extended to develop hierarchical models using taxonomic or functional proximity (Kindsvater et al. 2018) . Multispecies SDMs could be developed to quantify the (apparent) associations between species (Guisan and Zimmermann 2000, Ovaskainen et al. 2016, Thorson et al. 2016), and then these could be used to reconstruct and predict the distribution for any-and-all of the species participating in the model. This approach can also have potential as a cross-calibration method in correcting for errors due to detectability, or unknown observation effort (Chambert et al. 2018, Peel et al. 2019).

7.3 Combination with vantage point data

Several data types could come under this category. The most important is derived from on-shore observation stations (e.g. by use of total station comprising theodolite and distancer). These could be important sources of information for near-shore distribution. Their combination with line transect survey data is relatively straightforward since both data types belong to the broader class of transect methods (Buckland et al. 2001). Terrestrial habitat preferences for seabirds are a considerably less studied aspect of their biology, but one that is particularly pertinent for determining the placement of potential new colonies and for examining nest placement within colonies. For example, (Clark et al. 2019) used integrated modelling of transect and burrow occupancy data to map out the distribution of a cryptic seabird on the colony. Of particular relevance for studying human-seabird interactions is the terrestrial distribution of scavenging species such as gulls, as it shifts away from marine foraging.

Other data could come from methods of detection/non detection such as acoustic stations or camera trapping (Ngoprasert et al. 2019). Their integration with survey data is equivalent to the combination between occupancy and abundance (Keil et al. 2013, Hefley and Hooten 2016, Pacifici et al. 2017, Fletcher et al. 2019).

7.4 Combination with citizen science data

Methods for coordination of citizen science programmes are flourishing in ecology (Bonney et al. 2009, Amano et al. 2014, Chase and Levine 2016, Giraud et al. 2016, Kosmala et al. 2016, Wald et al. 2016, La Sorte et al. 2018) and so is the development of statistical methodologies for dealing with the fundamental restrictions in the quality of such opportunistic data (Hochachka et al. 2012, Bird et al. 2014). The main issue with citizen scientist data isn't so much the higher level of bias or imprecision in species identification that might arise in some cases, but rather, the heterogeneity in those across individual observers, through space and time. Although it is possible in principle to account for such heterogeneities in analysis frameworks, the task is made difficult by the frequent absence of information on effort, precision and accuracy. Such gaps in knowledge then need to be supplemented by proxies (such as plausible assumptions about the behaviour and distribution of citizen observers or more detailed models of these). It is also possible that integrated analysis of multispecies surveys (see Section 7.2) may help by allowing the collective detections of all species to act as an approximation of the effort distribution. The combination of citizen science data with survey data may happen either by using the opportunistic data to fine-tune the design of surveys

(Reich et al. 2018) or by analysing them together in an integrated platform using different observation models (Nelli et al. 2019).

7.5 Combination with telemetry data

The idea of combining survey with tracking data is as old as the early days of satellite telemetry. This combination however has proved particularly challenging. There is a fundamental difference between those two data types: Surveys focus on particular regions of space and can (in principle) observe any individual animal in the population. Telemetry studies focus on particular individuals and can (in principle) observe any region in space. Therefore, we have a situation of incompatibility, which (like many of the data-pooling problems discussed in this report) could be turned into a situation of complementarity, although as yet that has not been achieved. Studies that have attempted this marriage in the seabird literature have often tended to inflict heavy censoring on the data. For example, (Louzao et al. 2009) found it necessary to convert survey data into occupancy and to select a single foraging trip from each tagged bird, achieving a form of indiscriminate pooling (see Section 4.1). There are papers (e.g. Carroll et al. 2019) that have taken an ad-hoc comparative approach (see Section 4.2) and papers (e.g. Yamamoto et al. 2015, Zipkin and Saunders 2018) that have followed more powerful approaches of post-hoc combination (see Section 4.3). However, none of the current approaches have achieved fully integrated inference. A major obstacle to joint inference is the incongruence between frameworks used for these two data types. Telemetry data are most conveniently analysed via step selection functions (SSFs), while resource selection functions (RSFs) are most appropriate for survey data. A fundamental problem with these approaches is that they do not, by default, lead to the same results. Specifically, scaling up by simulation the microscopic model obtained via SSFs does not yield the same expected distribution generated by an RSF (Signer et al. 2017). A promising development in this area is the convergence between the frameworks of resource selection and step selection analyses (Michelot et al. 2019b, 2019c). This work has established the conditions under which SSF and RSF frameworks agree, and has derived methods for joint inference (Michelot et al. 2019a).

7.6 Combination with mark-recapture data

Mark-recapture data have rarely been used to map seabird distributions and fit habitat models (Camphuysen et al. 2004), however, they are a potentially valuable repository of spatial data that are also individually referenced. In a sense therefore, mark-recapture data carry intermediate information between point transects and telemetry tracking and could, in the longer-term benefit from current developments in the integration between these two (see Section 7.5).

7.7 Combination with non-spatial data

Integration into SDMs of non-spatial data can nevertheless be valuable for spatial prediction. We have argued at several points above that SDMs can benefit by being embedded in the dynamics of the population and community that they refer to. There is now a consistent move to think more mechanistically about the constraints of species distributions by connecting them to those other aspects of ecology (Morales et al. 2010, Ehrlén and Morris 2015b, Matthiopoulos et al. 2015, Mcloughlin et al. 2018, Zipkin and Saunders 2018, Yen et al. 2019).

8 References

Aarts, G., J. Fieberg, and J. Matthiopoulos. 2012. Comparative interpretation of count, presence-absence and point methods for species distribution models. Methods in Ecology and Evolution 3:177–187.

Aarts, G., M. MacKenzie, B. McConnell, M. Fedak, and J. Matthiopoulos. 2008. Estimating space-use and habitat preference from wildlife telemetry data. Ecography 31:140–160.

Amano, T., J. D. Lamming, and W. J. Sutherland. 2014. Spatial Gaps in Global Biodiversity Information and the Role of Citizen Science. BioScience 66:393–400.

Araújo, M. B., and M. New. 2007. Ensemble forecasting of species distributions. Trends in Ecology and Evolution 22:42–47.

Arthur, S. M., B. F. J. Manly, L. L. Mcdonald, W. Gerald, S. Ecology, N. Jan, S. M. Arthur, B. F. J. Manly, and G. W. Garner. 1996. Assessing Habitat Selection when Availability Changes. Ecology 77:215–227.

Asseburg, C., J. Harwood, J. Matthiopoulos, and S. Smout. 2009. The functional response of generalist predators and its implications for the monitoring of marine ecosystems. Top Predators in Marine Ecosystems:262–274.

Augustin, N. H., M. A. Mugglestone, and S. T. Buckland. 1996. An autologistic model for the spatial distribution of wildlife. Journal of Applied Ecology 33:339–347.

Bachl, F. E., F. Lindgren, D. L. Borchers, and J. B. Illian. 2019. inlabru: an R package for Bayesian spatial modelling from ecological survey data. Methods in Ecology and Evolution 10:760–766.

Bahn, V., and B. J. Mcgill. 2007. Can niche-based distribution models outperform spatial interpolation? Global Ecology and Biogeography 16:733–742.

Barry, S., and J. Elith. 2006. Error and uncertainty in habitat models. Journal of Applied Ecology 43:413–423.

Beale, C. M., M. J. Brewer, and J. J. Lennon. 2014. A new statistical framework for the quantification of covariate associations with species distributions. Methods in Ecology and Evolution 5:421–432.

Beale, C. M., J. J. Lennon, J. M. Yearsley, M. J. Brewer, and D. A. Elston. 2010. Regression analysis of spatial data. Ecology Letters 13:246–264.

Beaumont, M. A. 2010. Approximate Bayesian Computation in Evolution and Ecology. Annual Review of Ecology, Evolution, and Systematics 41:379–406.

Bell, D. M., and D. R. Schlaepfer. 2016. On the dangers of model complexity without ecological justification in species distribution modeling. Ecological Modelling 330:50–59.

Bird, T. J., A. E. Bates, J. S. Lefcheck, N. A. Hill, R. J. Thomson, G. J. Edgar, R. D. Stuart-Smith, S. Wotherspoon, M. Krkosek, J. F. Stuart-Smith, G. T. Pecl, N. Barrett, and S. Frusher. 2014. Statistical solutions for error and bias in global citizen science datasets. Biological Conservation 173:144–154.

Bolker, B. M. 2008. Ecological models and data in R. Princeton University Press.

Bonney, R., C. B. Cooper, J. Dickinson, S. Kelling, T. Phillips, K. V. Rosenberg, and J. Shirk. 2009. Citizen Science: A Developing Tool for Expanding Science Knowledge and Scientific Literacy. BioScience 59:977–984.

Boyce, M. S., and L. L. McDonald. 1999. Relating populations to habitats using resource selection functions. Trends in Ecology & Evolution 14:268–272.

Boyce, M. S., L. L. McDonald, and B. F. J. Manly. 1999. Relating populations to habitats-Reply. Trends In Ecology & Evolution 14:490.

Bradbury, G., M. L. Burt, and R. Hexter. 2011. Digital aerial surveillance of inshore waterbirds in Liverpool Bay Special Protection Area.

Bradbury, G., M. Trinder, B. Furness, A. N. Banks, R. W. G. Caldow, and D. Hume. 2014. Mapping Seabird Sensitivity to offshore wind farms. PLoS ONE 9.

Buckland, S. T., D. R. Anderson, K. . Burnham, J. L. Laake, D. Borchers, and L. Thomas. 2001. Introduction to Distance Sampling: Estimating Abundance of Biological Populations. Oxford University Press.

Buckland, S. T., D. R. Anderson, K. P. Burnham, and J. L. Laake. 2008. Advanced distance sampling.

Buckland, S. T., M. L. Burt, E. A. Rexstad, M. Mellor, A. E. Williams, and R. Woodward. 2012. Aerial surveys of seabirds: The advent of digital methods. Journal of Applied Ecology 49:960–967.

Buckland, S. T., J. L. Laake, and D. L. Borchers. 2010. Double-observer line transect methods: Levels of independence. Biometrics 66:169–177.

Buckland, S. T., K. B. Newman, C. Fernández, L. Thomas, and J. Harwood. 2007. Embedding Population Dynamics Models in Inference. Statistical Science 22:44–58.

Buckland, S. T., K. B. Newman, L. Thomas, and N. B. Koesters. 2004. State-space models for the dynamics of wild animal populations. Ecological Modelling 171:157–175.

Burnham, K. P., and D. R. Anderson. 2004. Multimodel inference understanding {AIC and BIC} in model selection. Sociological methods & research 33:261–304.

Burnham, K. P., D. R. Anderson, and K. P. Huyvaert. 2011. {AIC} model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons. Behavioral Ecology and Sociobiology 65:23–35.

Burt, M. L., E. A. Rexstad, and S. T. Buckland. 2009. Comparison of visual and digital aerial survey results of avian abundance for Round 3, Norfolk Region.

Calabrese, J. M., G. Certain, C. Kraan, and C. F. Dormann. 2014. Stacking species distribution models and adjusting bias by linking them to macroecological models. Global Ecology and Biogeography 23:99–112.

Camphuysen, C. J., B. Calvo, J. Durinck, K. Ensor, A. Follestad, R. W. Furness, S. Garthe, G. Leaper, H. Skov, M. L. Tasker, and C. J. N. Winter. 1995. Consumption of discards by seabirds in the North Sea.

Camphuysen, K. C. J., T. A. D. Fox, M. M. F. Leopold, and I. K. Petersen. 2004. Towards standardised seabirds at sea census techniques in connection with environmental impact assessments for offshore wind farms in the U.K. Cowrie - Bam - 02-2002:1–38.

Carroll, M. J., E. D. Wakefield, E. S. Scragg, E. Owen, S. Pinder, M. Bolton, J. J. Waggitt, and P. 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.

Certain, G., and V. Bretagnolle. 2008. Monitoring seabirds population in marine ecosystem: The use of strip-transect aerial surveys. Remote Sensing of Environment 112:3314–3322.

Chakraborty, A., A. E. Gelfand, A. M. Wilson, A. M. Latimer, J. A. Silander, S. Journal, R. Statistical, S. Series, C. A. Statistics, and A. Chakraborty. 2011. Point pattern modelling for degraded presence-only data over large regions. Journal of the Royal Statistical Society. Series C: Applied Statistics 60:757–776.

Chambert, T., E. H. C. Grant, D. A. W. Miller, J. D. Nichols, K. P. Mulder, and A. B. Brand. 2018. Two-species occupancy modelling accounting for species misidentification and non-detection. Methods in Ecology and Evolution 9:1468–1477.

Chase, S. K., and A. Levine. 2016. A framework for evaluating and designing citizen science programs for natural resources monitoring. Conservation Biology 30:456–466.

Clark, T. J., J. Matthiopoulos, A. S. Bonnet-Lebrun, L. Campioni, P. Catry, I. Marengo, S. Poncet, and E. Wakefield. 2019. Integrating habitat and partial survey data to estimate the regional population of a globally declining seabird species, the sooty shearwater. Global Ecology and Conservation 17:1–15.

Clarke, E. D., L. B. Spear, M. L. McCracken, F. F. C. Marques, D. L. Borchers, S. T. Buckland, and D. G. Ainley. 2003. Validating the use of generalized additive models and at-sea surveys to estimate size and temporal trends of seabird populations. Journal of Applied Ecology 40:278–292.

Csilléry, K., M. G. B. Blum, O. E. Gaggiotti, and O. François. 2010. Approximate Bayesian Computation (ABC) in practice. Trends in Ecology and Evolution 25:410–418.

D'Amen, M., A. Dubuis, R. F. Fernandes, J. Pottier, L. Pellissier, and A. Guisan. 2015. Using species richness and functional traits predictions to constrain assemblage predictions from stacked species distribution models. Journal of Biogeography 42:1255–1266.

Distler, T., J. G. Schuetz, J. Velásquez-Tibatá, and G. M. Langham. 2015. Stacked species distribution models and macroecological models provide congruent projections of avian species richness under climate change. Journal of Biogeography 42:976–988.

Dormann, C. F., J. M. Calabrese, G. Guillera-Arroita, E. Matechou, V. Bahn, K. Bartoń, C. M. Beale, S. Ciuti, J. Elith, K. Gerstner, J. Guelat, P. Keil, J. J. Lahoz-Monfort, L. J. Pollock, B. Reineking, D. R. Roberts, B. Schröder, W. Thuiller, D. I. Warton, B. A. Wintle, S. N. Wood, R. O. Wüest, and F. Hartig. 2018. Model averaging in ecology: a review of Bayesian, information-theoretic, and tactical approaches for predictive inference. Ecological Monographs 88:485–504.

Dormann, C. F., J. M. McPherson, B. Araújo, Miguel, R. Bivand, J. Bolliger, G. Carl, R. G. Davies, A. Hirzel, W. Jetz, W. Daniel Kissling, I. Kühn, R. Ohlemüller, P. P. Peres-Neto, B. Reineking, B. Schröder, F. M. Schurr, and R. Wilson. 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. Ecography 30:609–628.

Ehrlén, J., and W. F. Morris. 2015a. Predicting changes in the distribution and abundance of species under environmental change. Ecology Letters 18:303–314.

Ehrlén, J., and W. F. Morris. 2015b. Predicting changes in the distribution and abundance of species under environmental change. Ecology Letters 18:303–314.

Elith, J., S. J. Phillips, T. Hastie, M. Dudík, Y. E. Chee, and C. J. Yates. 2011. A statistical explanation of MaxEnt for ecologists. Diversity and Distributions 17:43–57.

Fieberg, J., J. Matthiopoulos, M. Hebblewhite, M. S. M. S. Boyce, and J. L. J. L. Frair. 2010. Correlation and studies of habitat selection: problem, red herring or opportunity? Philosophical Transactions of the Royal Society of London B: Biological Sciences 365:2233–2244.

Fieberg, J. R. J. R., J. D. J. D. Forester, G. M. G. M. Street, D. H. D. H. Johnson, A. A. A. ArchMiller, and J. Matthiopoulos. 2018. Used-habitat calibration plots: a new procedure for validating species distribution, resource selection, and step-selection models. Ecography 41:737–752.

Fithian, W., and T. Hastie. 2013. Finite-sample equivalence in statistical models for presence-only data. Annals of Applied Statistics 7:1917–1939.

Fletcher, R. J., T. J. Hefley, E. P. Robertson, B. Zuckerberg, R. A. McCleery, and R. M. Dorazio. 2019. A practical guide for combining data to model species distributions. Ecology 0:e02710.

Fletcher, R. J., R. A. McCleery, D. U. Greene, and C. A. Tye. 2016. Integrated models that unite local and regional data reveal larger-scale environmental relationships and improve predictions of species distributions. Landscape Ecology 31:1369–1382.

Frair, J. L., J. Fieberg, M. Hebblewhite, F. Cagnacci, N. J. DeCesare, and L. Pedrotti. 2010. Resolving issues of imprecise and habitat-biased locations in ecological analyses using {GPS} telemetry data. Philosophical Transactions of the Royal Society of London B: Biological Sciences 365:2187–2200.

Furnas, B. J., D. S. Newton, G. D. Capehart, and C. W. Barrows. 2019. Hierarchical distance sampling to estimate population sizes of common lizards across a desert ecoregion. Ecology and Evolution 9:3046–3058.

Furness, R. W. 2016. Third party review of Fehmarnbelt fixed link plausibility check report and suitability of baseline plus check data on marine birds and mammals.

Giraud, C., C. Calenge, C. Coron, and R. Julliard. 2016. Capitalizing on opportunistic data for monitoring relative abundances of species. Biometrics 72:649–658.

Grecian, W. J. J., M. J. J. Witt, M. J. J. Attrill, S. Bearhop, P. H. H. Becker, C. Egevang, R. W. W. Furness, B. J. J. Godley, J. González-Solís, D. Grémillet, M. Kopp, A. Lescroël, J. Matthiopoulos, S. C. C. Patrick, H.-U. Peter, R. A. A. Phillips, I. J. J. Stenhouse, and S. C. C. Votier. 2016. Seabird diversity hotspot linked to ocean productivity in the Canary Current Large Marine Ecosystem. Biology Letters 12.

Grecian, W. J., M. J. Witt, M. J. Attrill, S. Bearhop, B. J. Godley, D. Grémillet, K. C. Hamer, and S. C. Votier. 2012. A novel projection technique to identify important at-sea areas for seabird conservation: An example using Northern gannets breeding in the North East Atlantic. Biological Conservation 156:43–52.

Guillera-Arroita, G., and J. J. Lahoz-Monfort. 2012. Designing studies to detect differences in species occupancy: Power analysis under imperfect detection. Methods in Ecology and Evolution 3:860–869.

Guisan, A., and W. Thuiller. 2005. Predicting species distribution: offering more than simple habitat models. Ecology Letters 8:993–1009.

Guisan, A., W. Thuiller, and N. E. Zimmermann. 2017. Habitat suitability and distribution models: with applications in R. Cambridge University Press.

Guisan, A., and N. E. Zimmermann. 2000. Predictive habitat distribution models in ecology. Ecological Modelling 135:147–186.

Hedley, S. L., and S. T. Buckland. 2004. Spatial models for line transect sampling. Journal of Agricultural, Biological, and Environmental Statistics 9:181–199.

Hefley, T. J., and M. B. Hooten. 2016. Hierarchical Species Distribution Models. Current Landscape Ecology Reports 1:87–97.

Hochachka, W. M., D. Fink, R. A. Hutchinson, D. Sheldon, W. K. Wong, and S. Kelling. 2012. Data-intensive science applied to broad-scale citizen science. Trends in Ecology and Evolution 27:130–137.

Hodges, J. S., and B. J. Reich. 2010. Adding spatially-correlated errors can mess up the fixed effect you love. American Statistician 64:325–334.

Holbrook, J. D., L. E. Olson, N. J. DeCesare, M. Hebblewhite, J. R. Squires, and R. Steenweg. 2019. Functional responses in habitat selection: clarifying hypotheses and interpretations. Ecological Applications 29:e01852.

Holling, C. S. 1959. Some characteristics of simple types of predation and parasitism. The Canadian Entomologist 91:385–398.

Hothorn, T., J. Müller, B. Schröder, T. Kneib, and R. Brandl. 2011. Decomposing environmental, spatial, and spatiotemporal components of species distributions. Ecological Monographs 81:329–347.

Isojunno, S., J. Matthiopoulos, and P. G. H. H. P. G. H. Evans. 2012. Harbour porpoise habitat preferences: Robust spatio-temporal inferences from opportunistic data. Marine Ecology Progress Series 448:155–170.

Jones, E. L. L., B. J. J. McConnell, S. Smout, P. S. S. Hammond, C. D. D. Duck, C. D. D. Morris, D. Thompson, D. J. F. J. F. Russel, C. Vincent, M. Cronin, R. J. J. Sharples, and J. Matthiopoulos. 2015. Patterns of space use in sympatric marine colonial predators reveal scales of spatial partitioning. Marine Ecology Progress Series 534.

Kearney, M., and W. Porter. 2009. Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. Ecology letters 12:334–350.

Keil, P., J. Belmaker, A. M. Wilson, P. Unitt, and W. Jetz. 2013. Downscaling of species distribution models: A hierarchical approach. Methods in Ecology and Evolution 4:82–94.

Keil, P., A. M. Wilson, and W. Jetz. 2014. Uncertainty, priors, autocorrelation and disparate data in downscaling of species distributions. Diversity and Distributions 20:797–812.

Kent, M., R. A. Moyeed, C. L. Reid, R. Pakeman, and R. Weaver. 2006. Geostatistics, spatial rate of change analysis and boundary detection in plant ecology and biogeography. Progress in Physical Geography 30:201–231.

Kindsvater, H. K. H. K., N. K. N. K. Dulvy, C. Horswill, M.-J. M. J. M.-J. Juan-Jordá, M. Mangel, and J. Matthiopoulos. 2018. Overcoming the Data Crisis in Biodiversity Conservation. Trends in Ecology and Evolution 33:676–688.

de Knegt, H. J., F. van Langevelde, M. B. Coughenour, A. K. Skidmore, W. F. de Boer, I. . M. A. Heitkönig, N. M. Knox, R. Slotow, C. van der Waal, and H. H. T. Prins. 2010. Spatial autocorrelation and the scaling of species — environment relationships. Ecology 91:2455–2465.

Koshkina, V., Y. Wang, A. Gordon, R. M. Dorazio, M. White, and L. Stone. 2017. Integrated species distribution models: combining presence-background data and site-occupancy data with imperfect detection. Methods in Ecology and Evolution 8:420–430.

Kosmala, M., A. Wiggins, A. Swanson, and B. Simmons. 2016. Assessing data quality in citizen science. Frontiers in Ecology and the Environment 14:551–560.

Lahoz-Monfort, J. J., G. Guillera-Arroita, and B. A. Wintle. 2014. Imperfect detection impacts the performance of species distribution models. Global Ecology and Biogeography 23:504–515.

Lele, S. R., M. Moreno, and E. Bayne. 2012. Dealing with detection error in site occupancy surveys: What can we do with a single survey? Journal of Plant Ecology 5:22–31.

Levin, S. A. 1992. The Problem of Pattern and Scale in Ecology: The Robert H. MacArthur Award Lecture. Ecology 73:1943–1967.

Lewis, S., T. N. Sherratt, K. C. Hamer, and S. Wanless. 2001. Evidence of intra-specific ccompetition for food in a pelagic seabird. Nature 412:816–819.

Lieske, D. J., D. A. Fifield, and C. Gjerdrum. 2014. Maps, models, and marine vulnerability: Assessing the community distribution of seabirds at-sea. Biological Conservation 172:15–28.

Lindgren, F. 2015. Journal of Statistical Software Bayesian Spatial Modelling with R-INLA. Journal Of Statistical Software 63:1–25.

Louzao, M., J. Bécares, B. Rodríguez, K. D. Hyrenbach, A. Ruiz, and J. M. Arcos. 2009. Combining vessel-based surveys and tracking data to identify key marine areas for seabirds. Marine Ecology Progress Series 391:183–197.

Manly, B. F. J. 2003. Estimating a resource selection function with line transect sampling. Pages 213–228 Journal of Applied Mathematics and Decision Sciences.

Marques, F. F. C., and S. T. Buckland. 2003. Incorporating Covariates into Standard Line Transect Analyses. Integration The Vlsi Journal:924–935.

Martino, S., and N. Chopin. 2007. Approximate Bayesian Inference for Latent Gaussian Models Using Approximate Bayesian Inference for Latent Gaussian Models Using Integrated Nested Laplace Approximations. Journal of the Royal Statistical Society Series B:1–28.

Matthiopoulos, J. 2003a. The use of space by animals as a function of accessibility and preference. Ecological Modelling 159:239–268.

Matthiopoulos, J. 2003b. Model-supervised kernel smoothing for the estimation of spatial usage. Oikos 102:367–377.

Matthiopoulos, J., and G. Aarts. 2007. The Spatial analysis of marine mammal abundance. Pages 27–33 *in* I. L. Boyd, W. D. Bowen, and S. J. Iverson, editors. Marine Mammal Ecology And Conservation: A Handbook of Techniques (Oxford Biology) (Techniques in Ecology & Conservation). Oxford University Press.

Matthiopoulos, J., L. Cordes, B. Mackey, D. Thompson, C. Duck, S. Smout, M.

Caillat, and P. Thompson. 2014. State-space modelling reveals proximate causes of harbour seal population declines. Oecologia 174:151–162.

Matthiopoulos, J., J. Fieberg, G. Aarts, H. L. H. L. Beyer, J. M. J. M. J. M. Morales, and D. T. D. T. D. T. Haydon. 2015. Establishing the link between habitat selection and animal population dynamics. Ecological Monographs 85:413–436.

Matthiopoulos, J., C. Field, and R. MacLeod. 2019. Predicting population change from models based on habitat availability and utilization. Proceedings of the Royal Society B: Biological Sciences 286.

Matthiopoulos, J., M. Hebblewhite, G. Aarts, and J. Fieberg. 2011. Generalized functional responses for species distributions. Ecology 92:583–589.

Matthiopoulos, J., B. Mcconnell, C. Duck, and M. Fedak. 2004. Using satellite telemetry and aerial counts to estimate space use by grey seals around the British Isles. Journal of Applied Ecology 41:476–491.

McLoughlin, P. D., D. W. Morris, D. Fortin, E. Vander Wal, and A. L. Contasti. 2010. Considering ecological dynamics in resource selection functions. Journal of Animal Ecology 79:4–12.

Mcloughlin, P. D., D. W. Morris, D. Fortin, E. Vander Wal, A. L. Contasti, P. D. Mcloughlin, D. W. Morris, D. Fortin, E. Vander Wal, and A. L. Contasti. 2018. Considering ecological dynamics in resource selection functions. Journal of Animal Ecology 79:4–12.

Mendel, B., P. Schwemmer, V. Peschko, S. Müller, H. Schwemmer, M. Mercker, and S. Garthe. 2019. Operational offshore wind farms and associated ship traffic cause profound changes in distribution patterns of Loons (Gavia spp.). Journal of Environmental Management 231:429–438.

Merow, C., M. J. Smith, and J. A. Silander. 2013. A practical guide to MaxEnt for modeling species’ distributions: what it does, and why inputs and settings matter. Ecography 36:1058–1069.

Michelot, T., P. G. Blackwell, S. Chamaillé-Jammes, and J. Matthiopoulos. 2019a. Inference in MCMC step selection models. Biometrics 1:1–10.

Michelot, T., P. G. Blackwell, and J. Matthiopoulos. 2019b. Linking resource selection and step selection models for habitat preferences in animals. Ecology 100:1–12.

Michelot, T., P. Gloaguen, P. G. Blackwell, and M. P. Étienne. 2019c. The Langevin diffusion as a continuous-time model of animal movement and habitat selection. Methods in Ecology and Evolution 10:1894–1907.

Miller, D. A., J. D. Nicholas, M. B. T., E. H. Campbell Grant, and L. L. Bailey. 2011. Improving occupancy estimation when two types of observational error occur: non-detection and species misidentificatio. Ecology 92:1422–1428.

Miller, D. A. W., K. Pacifici, J. S. Sanderlin, and B. J. Reich. 2019. The recent past and promising future for data integration methods to estimate species' distributions. Methods in Ecology and Evolution 10:22–37.

Monestiez, P., L. Dubroca, E. Bonnin, J. Durbec, and C. Guinet. 2006. Geostatistical modelling of spatial distribution of Balaenoptera physalus in the Northwestern Mediterranean Sea from sparse count data and heterogeneous observation efforts. Ecological Modelling 193:615–628.

Moorcroft, P. R. 2012. Mechanistic approaches to understanding and predicting mammalian space use: recent advances, future directions. Journal of Mammalogy 93:903–916.

Moorcroft, P. R., A. Barnett, P. R. Moorcroft, and A. Barnett. 2008. Mechanistic Home Range Models and Resource Selection Analysis : A Reconciliation and Unification 89:1112–1119.

Moorcroft, P. R., M. A. Lewis, and R. L. Crabtree. 1999. Home Range Analysis Using a Mechanistic Home Range Model 80:1656–1665.

Moorcroft, P. R., M. A. Lewis, and R. L. Crabtree. 2006. Mechanistic home range models capture spatial patterns and dynamics of coyote territories in Yellowstone. Proceedings of the Royal Society B: Biological Sciences 273:1651–1659.

Van Moorter, B., D. Visscher, S. Benhamou, L. Börger, M. S. Boyce, and J. M. Gaillard. 2009. Memory keeps you at home: A mechanistic model for home range emergence. Oikos 118:641–652.

Morales, J. M. M., P. R. R. Moorcroft, J. Matthiopoulos, J. L. L. Frair, J. G. G. Kie, R.

A. A. Powell, E. H. H. Merrill, and D. T. T. Haydon. 2010. Building the bridge between animal movement and population dynamics. Philosophical Transactions of the Royal Society B Biological Sciences 365:2289–2301.

Morris, L. R., K. M. Proffitt, and J. K. Blackburn. 2016. Mapping resource selection functions in wildlife studies: Concerns and recommendations. Applied Geography 76:173–183.

Munson, M. A., R. Caruana, D. Fink, W. M. Hochachka, M. Iliff, K. V. Rosenberg, D. Sheldon, B. L. Sullivan, C. Wood, and S. Kelling. 2010. A method for measuring the relative information content of data from different monitoring protocols. Methods in Ecology and Evolution:no-no.

Mysterud, A., and R. A. Ims. 1998. Functional responses in habitat use: availability influences relative use in trade-off situations. Ecology 79:1435–1441.

National Archives. 2022. Open Government Licence https://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/ Accessed 30/03/2022

NaturalEngland. 2019. Digital video aerial surveys of red-throated diver in the Outer Thames Estuary Special Protection Area 2018.

Nelli, L., H. M. Ferguson, and J. Matthiopoulos. 2019. Achieving explanatory depth and spatial breadth in infectious disease modelling: Integrating active and passive case surveillance. Statistical Methods in Medical Research:096228021985638.

Newman, K. B., S. T. Buckland, B. J. T. Morgan, R. King, D. L. Borchers, D. J. Cole, P. Besbeas, O. Gimenez, and L. Thomas. 2014. Modelling Population Dynamics: Model Formulation, Fitting and Assessment Using State-Space Methods. Springer.

Ngoprasert, D., G. A. Gale, and A. J. Tyre. 2019. Abundance estimation from multiple data types for group-living animals: An example using dhole (Cuon alpinus). Global Ecology and Conservation 20:e00792.

Nur, N., J. Jahncke, M. P. Herzog, J. Howar, K. D. Hyrenbach, J. E. Zamon, D. G. Ainley, J. A. Wiens, K. Morgan, L. T. Ballance, N. Nur, J. Jahncke, M. P. Herzog, J. Howar, K. D. Hyrenbach, J. E. Zamon, D. G. Ainley, J. A. Wiens, K. Morgan, L. T. Ballance, and D. Stralberg. 2011. Where the wild things are : predicting hotspots of seabird aggregations in the California Current System 21:2241–2257.

Oedekoven, C. S., M. L. Mackenzie, L. A. S. Scott-Hayward, and E. Rexstad. 2012a. Marine Scotland Science Report 05/14 Statistical Modelling of Seabird and Cetacean Data: Literature Review.
Oedekoven, C. S., M. L. Mackenzie, L. A. S. Scott-Hayward, and E. Rexstad. 2012b. Marine Scotland Science Report 04 / 14 Statistical Modelling of Seabird and Cetacean Data : Guidance Document.

Oppel, S., A. Meirinho, I. Ramírez, B. Gardner, A. F. O'Connell, P. I. Miller, and M. Louzao. 2012. Comparison of five modelling techniques to predict the spatial distribution and abundance of seabirds. Biological Conservation 156:94–104.

Ovaskainen, O., N. Abrego, P. Halme, and D. Dunson. 2016. Using latent variable models to identify large networks of species-to-species associations at different spatial scales. Methods in Ecology and Evolution 7:549–555.

Ovaskainen, O., J. Hottola, and J. Siitonen. 2010. Modeling species co-occurrence by multivariate logistic regression generates new hypotheses on fungal interactions. Ecology 91:2514–21.

Pacifici, K., B. J. Reich, D. A. W. Miller, B. Gardner, G. Stauffer, S. Singh, A. McKerrow, and J. A. Collazo. 2017. Integrating multiple data sources in species distribution modeling: A framework for data fusion. Ecology 98:840–850.

Pacifici, K., B. J. Reich, D. A. W. Miller, and B. S. Pease. 2019. Resolving misaligned spatial data with integrated species distribution models. Ecology 100:1–15.

Paton, R. S., and J. Matthiopoulos. 2018. Defining the scale of habitat availability for models of habitat selection. Ecology 97:1113–1122.

Patterson, T. A. T. A., L. Thomas, C. Wilcox, O. Ovaskainen, and J. Matthiopoulos. 2008. State-space models of individual animal movement. Trends in Ecology & Evolution 23:87–94.

Pearce, J. L., and M. S. Boyce. 2006. Modelling distribution and abundance with presence-only data. Journal of Applied Ecology 43:405–412.

Pearman, P. B., A. Guisan, O. Broennimann, and C. F. Randin. 2008. Niche dynamics in space and time. Trends in Ecology and Evolution 23:149–158.

Peel, S. L., N. A. Hill, S. D. Foster, S. J. Wotherspoon, C. Ghiglione, and S. Schiaparelli. 2019. Reliable species distributions are obtainable with sparse, patchy and biased data by leveraging over species and data types. Methods in Ecology and Evolution 10:1002–1014.
Perrow, M. R., A. J. P. Harwood, E. R. Skeate, E. Praca, and S. M. Eglington. 2015. Use of multiple data sources and analytical approaches to derive a marine protected area for a breeding seabird. Biological Conservation 191:729–738.

Petersen, I., and R. Nielsen. 2011. Abundance and distribution of selected waterbird species in Danish marine areas.

Peterson, A. T., J. Soberón, R. G. Pearson, R. P. Anderson, E. Martinez-Meyer, M. Nakamura, and M. B. Araújo. 2011. Ecological niches and geographic distributions. Princeton University Press.

R Core Team. 2019. {R}: A language and environment for statistical computing. {Vienna, Austria; 2014}. URL http://www. R-project. org.

Reich, B. J., K. Pacifici, and J. W. Stallings. 2018. Integrating auxiliary data in optimal spatial design for species distribution modelling. Methods in Ecology and Evolution 9:1626–1637.

Renner, I. W., J. Elith, A. Baddeley, W. Fithian, T. Hastie, S. J. Phillips, G. Popovic, and D. I. Warton. 2015. Point process models for presence-only analysis. Methods in Ecology and Evolution 6:366–379.

Renner, I. W., J. Louvrier, and O. Gimenez. 2019. Combining multiple data sources in species distribution models while accounting for spatial dependence and overfitting with combined penalized likelihood maximization. Methods in Ecology and Evolution 2019:2118–2128.
Renner, I. W., and D. I. Warton. 2013. Equivalence of MAXENT and Poisson point process models for species distribution modeling in ecology. Biometrics 69:274–281.

Rexstad, E. A., and S. T. Buckland. 2012. SOSS-01. Displacement analysis boat surveys Kentish Flats.

Riotte-Lambert, L., and J. Matthiopoulos. 2019. Communal and efficient movement routines can develop spontaneously through public information use. Behavioral Ecology 30:408–416.

Robertson, M. P., N. Caithness, M. H. Villet, and N. J. Mar. 2001. A PCA-Based Modelling Technique for Predicting Environmental Suitability for Organisms from Presence Records A PCA-based modelling technique for predicting environmental suitability for organisms from presence records. Diversity 7:15–27.
Robinson, N. M., W. A. Nelson, M. J. Costello, J. E. Sutherland, and C. J. Lundquist. 2017. A systematic review of marine-based Species Distribution Models (SDMs) with recommendations for best practice. Frontiers in Marine Science 4:1–11.

Rue, H., S. Martino, and N. Chopin. 2009. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 71:319–392.

Sansom, A., L. J. Wilson, R. W. G. Caldow, and M. Bolton. 2018. Comparing marine distribution maps for seabirds during the breeding season derived from different survey and analysis methods. PLoS ONE 13:1–17.

Sardà-Palomera, F., L. Brotons, D. Villero, H. Sierdsema, S. E. Newson, and F. Jiguet. 2012. Mapping from heterogeneous biodiversity monitoring data sources. Biodiversity and Conservation 21:2927–2948.

Segurado, P., M. B. Araújo, and W. E. Kunin. 2006. Consequences of spatial autocorrelation for niche-based models. Journal of Applied Ecology 43:433–444.

Signer, J., J. Fieberg, and T. Avgar. 2017. Estimating utilization distributions from fitted step-selection functions. Ecosphere 8.

Sinclair, S. J., M. D. White, and G. R. Newell. 2010. How useful are species distribution models for managing biodiversity under future climates? Ecology and Society 15.

Smout, S., C. Asseburg, J. Matthiopoulos, C. Fernández, S. Redpath, S. Thirgood, and J. Harwood. 2010. The Functional Response of a Generalist Predator. PLoS ONE 5:7.

Sokal, U. R. R., and U. F. J. Rohlf. 1995. Biometry. W. H. Freeman.

La Sorte, F. A., C. A. Lepczyk, J. L. Burnett, A. H. Hurlbert, M. W. Tingley, and B. Zuckerberg. 2018. Opportunities and challenges for big data ornithology. The Condor 120:414–426.

Strindberg, S., and S. T. Buckland. 2004. Zigzag survey designs in line transect sampling. Journal of Agricultural, Biological, and Environmental Statistics 9:443–461.

Talluto, M. V., I. Boulangeat, A. Ameztegui, I. Aubin, D. Berteaux, A. Butler, F. Doyon, C. R. Drever, M. J. Fortin, T. Franceschini, J. Liénard, D. Mckenney, K. A. Solarik, N. Strigul, W. Thuiller, and D. Gravel. 2016. Cross-scale integration of knowledge for predicting species ranges: A metamodelling framework. Global Ecology and Biogeography 25:238–249.

Tasker, M. L., P. H. Jones, T. Dixon, and B. F. Blake. 1984. Counting Seabirds at Sea from Ships: A Review of Methods Employed and a Suggestion for a Standardized Approach. The Auk 101:567–577.

Thaxter, C. B., B. Lascelles, K. Sugar, A. S. C. P. Cook, S. Roos, M. Bolton, R. H. W. Langston, and N. H. K. Burton. 2012. Seabird foraging ranges as a preliminary tool for identifying candidate Marine Protected Areas. Biological Conservation 156:53–61.

Thiers, L., M. Louzao, V. Ridoux, M. Le Corre, S. Jaquemet, and H. Weimerskirch. 2014. Combining methods to describe important marine habitats for top predators: Application to identify biological hotspots in tropical waters. PLoS ONE 9:1–23.

Thomas, L., S. T. Buckland, E. A. Rexstad, J. L. Laake, S. Strindberg, S. L. Hedley, J. R. B. Bishop, T. A. Marques, and K. P. Burnham. 2010. Distance software: Design and analysis of distance sampling surveys for estimating population size. Journal of Applied Ecology 47:5–14.

Thorson, J. T., J. N. Ianelli, E. A. Larsen, L. Ries, M. D. Scheuerell, C. Szuwalski, and E. F. Zipkin. 2016. Joint dynamic species distribution models: a tool for community ordination and spatio-temporal monitoring. Global Ecology and Biogeography 25:1144–1158.

Tuanmu, M.-N., A. Viña, G. J. Roloff, W. Liu, Z. Ouyang, H. Zhang, and J. Liu. 2011. Temporal transferability of wildlife habitat models: implications for habitat monitoring. Journal of Biogeography 38:1510–1523.

Tulloch, V. J., H. P. Possingham, S. D. Jupiter, C. Roelfsema, A. I. T. Tulloch, and C. J. Klein. 2013. Incorporating uncertainty associated with habitat data in marine reserve design. Biological Conservation 162:41–51.

Turlure, C., N. Schtickzelle, Q. Dubois, M. Baguette, R. L. H. Dennis, and H. Van Dyck. 2019. Suitability and transferability of the resource-based habitat concept: A test with an assemblage of butterflies. Frontiers in Ecology and Evolution 7.

Tyre, A. J., B. Tenhumberg, S. A. Field, D. Niejalke, K. Parris, and H. P. Possingham. 2003. Improving precision and reducing bias in biological surveys: Estimating false-negative error rates. Ecological Applications 13:1790–1801.

Waggitt, J. J., P. G. H. Evans, J. Andrade, A. . Banks, O. Boisseau, M. Bolton, G. Bradbury, T. Brereton, C. J. Camphuysen, J. Durinck, T. Felce, R. C. Fijn, I. Garcia-Baron, S. Garthe, S. C. . Geelhoed, A. Gilles, M. Goodall, J. Haelters, S. Hamilton, L. Hartny-Mills, N. Hodgins, K. James, M. Jessopp, A. S. Kavanagh, M. Leopold, K. Lohrengel, M. Louzao, N. Markones, J. Martinez-Cediera, O. O'Cadhla, S. L. Perry, G. J. Pierce, V. Ridoux, K. P. Robinson, M. B. Santos, C. Saavedra, H. Skov, E. W. M. Stienen, S. Sveegaard, P. Thompson, N. Vanermen, D. Wall, A. Webb, J. Wilson, S. Wanless, and J. G. Hiddink. 2019. Distribution maps of cetacean and seabird populations in the North-East Atlantic. Journal of Applied Ecology 205:1365-2664.13525.

Wakefield, E. D., T. W. Bodey, S. Bearhop, J. Blackburn, K. Colhoun, R. Davies, R. G. Dwyer, J. A. Green, D. Grémillet, A. L. Jackson, M. J. Jessopp, A. Kane, R. H. W. Langston, A. Lescroël, S. Murray, M. Le Nuz, S. C. Patrick, C. Péron, L. M. Soanes, S. Wanless, S. C. Votier, and K. C. Hamer. 2013. Space partitioning without territoriality in gannets. Science 341:68–70.

Wakefield, E. D. E. D., R. A. R. A. Phillips, P. N. P. N. Trathan, J. Arata, R. Gales, N. Huin, R. Graham, S. M. S. M. Waugh, H. Weimerskirch, and J. Matthiopoulos. 2011. Habitat preference, accessibility, and competition limit the global distribution of breeding Black-browed Albatrosses. Ecological Monographs 81:141–167.

Wald, D. M., J. Longo, and A. R. Dobell. 2016. Design principles for engaging and retaining virtual citizen scientists. Conservation Biology 30:562–570.

Warton, D. I., L. C. Shepherd, and others. 2010. Poisson point process models solve the "pseudo-absence problem''' for presence-only data in ecology." The Annals of Applied Statistics 4:1383–1402.

Webb, A., and G. Nehls. 2019. Surveying Seabirds. Page 330 *in* M. R. Perrow, editor. Wildlife and windfarms, Conflicts and solutions.

Wood, S. N. 2006. Generalized Additive Models: An Introduction with R. Chapman & Hall/CRC.

Yamamoto, T., Y. Watanuki, E. L. Hazen, B. B. Nishizawa, H. Sasaki, and A. Takahashi. 2015. Statistical integration of tracking and vessel survey data to incorporate life history differences in habitat models. Ecological Applications 25:2394–2406.

Yates, K. L., P. J. Bouchet, M. J. Caley, K. Mengersen, C. F. Randin, S. Parnell, A. H. Fielding, A. J. Bamford, S. Ban, A. M. Barbosa, C. F. Dormann, J. Elith, C. B. Embling, G. N. Ervin, R. Fisher, S. Gould, R. F. Graf, E. J. Gregr, P. N. Halpin, R. K. Heikkinen, S. Heinänen, A. R. Jones, P. K. Krishnakumar, V. Lauria, H. Lozano-Montes, L. Mannocci, C. Mellin, M. B. Mesgaran, E. Moreno-Amat, S. Mormede, E. Novaczek, S. Oppel, G. Ortuño Crespo, A. T. Peterson, G. Rapacciuolo, J. J. Roberts, R. E. Ross, K. L. Scales, D. Schoeman, P. Snelgrove, G. Sundblad, W. Thuiller, L. G. Torres, H. Verbruggen, L. Wang, S. Wenger, M. J. Whittingham, Y. Zharikov, D. Zurell, and A. M. M. Sequeira. 2018. Outstanding Challenges in the Transferability of Ecological Models. Trends in Ecology and Evolution 33:790–802.

Yen, J. D. L., Z. Tonkin, J. Lyon, W. Koster, A. Kitchingman, K. Stamation, and P. A. Vesk. 2019. Integrating multiple data types to connect ecological theory and data among levels. Frontiers in Ecology and Evolution 7:1–7.

Zeller, K. A., K. McGarigal, and A. R. Whiteley. 2012. Estimating landscape resistance to movement: A review. Landscape Ecology 27:777–797.

Zeller, K. A., T. W. Vickers, H. B. Ernest, and W. M. Boyce. 2017. Multi-level, multi-scale resource selection functions and resistance surfaces for conservation planning: Pumas as a case study. PLoS ONE 12:1–20.

Zipkin, E. F., and S. P. Saunders. 2018. Synthesizing multiple data types for biological conservation using integrated population models. Biological Conservation 217:240–250.

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

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

This publication is available at **www.gov.scot**

Any enquiries regarding this publication should be sent to us at

The Scottish Government
St Andrew's House
Edinburgh
EH1 3DG

ISBN: 978-1-80435-210-6 (web only)

Published by The Scottish Government, June 2022

Produced for The Scottish Government by APS Group Scotland, 21 Tennant Street, Edinburgh EH6 5NA
PPDAS1035130 (06/22)