

OFFSHORE RENEWABLES JOINT INDUSTRY
PROGRAMME (ORJIP) FOR OFFSHORE WIND



AppSaS – Apportioning seabirds seen-at-sea

WP3 – Method evaluation

February 2024



Contents

ORJIP Offshore Wind	3
Acknowledgements	3
Who we are.....	3
1. Introduction	5
2. Extending summer apportioning to lesser black-backed gull.....	6
Introduction.....	6
Methods	8
Data collection	8
Tracking data	8
Colony data	11
Environmental data	12
Data preparation.....	13
Tracking data	13
Colony data	15
Environmental data	15
Modelling approach	19
Colony-specific habitat use.....	19
Explanatory covariates.....	20
Model selection	21
Estimating usage	21
Comparison with existing apportioning approach	22
Results.....	23
Tracking of lesser black-backed gulls	23
Models of distribution.....	25
Estimated population-level distribution	26
Comparison with existing apportioning approach	29
Discussion.....	32
3. Extending apportioning to winter using geolocation data	35
Introduction.....	35
Methods	36
Geolocation data description.....	36
Statistical methods for analysis of geolocator data	40
Deriving spatial locations from geolocator data	40
Kernel density estimation.....	40
Accounting for uncertainty.....	41
Removing inland locations	41

Statistical methods for evaluating maps derived from geolocator data	43
Visual summaries	43
Comparison against BDMPS tables	43
Visualising density using space time cubes	43
Evaluating potential to produce distributions for untracked colonies	44
Results.....	44
Visual summaries	44
Comparison against BDMPS tables	58
Visualizing density using space time cubes	64
Potential to produce distributions for untracked colonies	69
Discussion.....	71
The use of geolocator-based models in apportioning	71
Wider interpretations	73
4. Conclusions	74
5. Acknowledgements	75
Analysis of GPS data for lesser black-backed gulls	75
Data sources.....	75
General acknowledgements	76
Analysis of geolocator data for guillemot and razorbill	77
6. References	78
7. Appendices.....	84
Examples of extracted environmental covariates.....	86

ORJIP Offshore Wind

The Offshore Renewables Joint Industry Programme (ORJIP) for Offshore Wind is a collaborative initiative that aims to:

- Fund research to improve our understanding of the effects of offshore wind on the marine environment
- Reduce the risk of not getting, or delaying consent for, offshore wind developments
- Reduce the risk of getting consent with conditions that reduce viability of the project.

The programme pools resources from the private sector and public sector bodies to fund projects that provide empirical data to support consenting authorities in evaluating the environmental risk of offshore wind. Projects are prioritised and informed by the ORJIP Advisory Network which includes key stakeholders, including statutory nature conservation bodies, academics, non-governmental organisations and others.

The current stage is a collaboration between The Carbon Trust, EDF Energy Renewables Limited, Ocean Winds UK Limited, Equinor ASA, Ørsted Power (UK) Limited, RWE Offshore Wind GmbH, Shell Global Solutions International B.V., SSE Renewables Services (UK) Limited, TotalEnergies OneTech, Crown Estate Scotland, Scottish Government (acting through the Offshore Wind Directorate and the Marine Directorate), and The Crown Estate Commissioners.

For further information regarding the ORJIP Offshore Wind programme, please refer to the [Carbon Trust website](#), or contact Ivan Savitsky (ivan.savitsky@carbontrust.com) and Žilvinas Valantiejus (zilvinas.valantiejus@carbontrust.com).

Acknowledgements

This document was produced on behalf of ORJIP Offshore Wind by Biomathematics and Statistics Scotland (BioSS), UK Centre for Ecology & Hydrology (UKCEH) and British Trust for Ornithology (BTO). The report was authored by Adam Butler, Lila Buckingham, Chris Thaxter, Esther Jones, Katherine Whyte, Aonghais Cook, Maria Bogdanova, David Ewing, Kate Searle and Francis Daunt.

The project has been advised by the ORJIP Offshore Wind Steering Group, and AppSaS Project Expert Panel. We would like to thank the following organisations for their advice and support of the project via participation on the Project Expert Panel:

- Natural England
- Natural Resources Wales
- NatureScot

This report was sponsored by the ORJIP Offshore Wind programme. For the avoidance of doubt, this report expresses independent views of the authors.

Who we are

Our mission is to accelerate the move to a decarbonised future.

We have been climate pioneers for more than 20 years, partnering with leading businesses, governments and financial institutions globally. From strategic planning and target setting to activation and communication - we are your expert guide to turn your climate ambition into impact.

We are one global network of 400 experts with offices in the UK, the Netherlands, South Africa, China, Singapore and Mexico. To date, we have helped set 200+ science-based targets and guided 3,000+ organisations in 70 countries on their route to Net Zero.

1. Introduction

The aim of this workpackage is to evaluate apportioning methods identified during [WP1](#) by utilising the datasets with the most potential from [WP2](#) to determine their consistency and assess their strengths/weaknesses. [WP1](#) and [WP2](#) demonstrated, however, that the set of methods that have already been used for apportioning in a UK context is very limited, whilst identifying situations in which the data are available to extend these methods or develop new methods. In consultation with the Project Steering Group, WP3 therefore focussed on developing or extending more advanced methods for apportioning in two key situations in which this was identified, based on [WP1](#) and [WP2](#), to be both feasible and a high priority, and then on evaluating these new or extended methods against existing methods.

In the context of breeding season apportioning, the SNH/NatureScot Apportioning Tool is the only apportioning tool that is currently in use for most species – it is a simple method that is easily implemented but makes strong assumptions and does not utilize all of the available data. GPS summer tracking data were previously used to develop spatial distribution maps (Wakefield et al., 2017) and thereby apportioning estimates (Butler et al., 2020) for four species, providing an alternative to the SNH/NatureScot Tool. Extensive GPS tracking data, although not available for all species, are available for species beyond those considered previously ([WP2](#)). Accordingly, we here extend the spatial distribution mapping approach of Wakefield et al. (2017) to include a new species, lesser black-backed gull, in order for these maps to provide an alternative apportioning method to the SNH/NatureScot Tool for this species.

Breeding season apportioning is aided by the strong central place foraging constraint, which necessitates a strong relationship between the spatial distribution of birds and the distance from the colony. Non-breeding season apportioning is more challenging, and the only approach that is currently used in practice for this in the UK is BDMPS, an approach that considers broad spatial regions. We develop an alternative approach, using GLS winter tracking data of guillemot and razorbill (Buckingham et al., 2022) to develop maps that can be used for apportioning. Unlike previous apportioning methods, this approach also provides a quantification of uncertainty – the locational uncertainty in GLS data is potentially large, and we account for this uncertainty explicitly.

We evaluate the new summer GPS-based lesser black-backed gull maps against the NatureScot tool, and the new winter GLS-based maps for guillemot and razorbill against BDMPS. In both cases, the comparison is against the spatial distribution that underpins the methods, not the apportioning percentages i.e., it is an evaluation of the underlying assumptions of the methods rather than a comparison of the actual apportioning results. This is the fundamental analytical progress that is required in order to develop new apportioning estimates, and the most useful, and easily interpretable, way to compare the key differences between the two approaches. As such, we were able to fulfil two main objectives: to broaden estimation of underpinning distributions used in apportioning calculations to new species and seasons, and to compare methods of generating colony-specific distributions underpinning apportioning calculations.

The next workpackage ([WP4](#)) is tasked with developing a tool to calculate apportioning percentages in a suitable format, and will involve developing the code for the apportioning tool based on the methods presented here, including the development of a user interface. As well as the existing methods

(SNH/NatureScot and BDMPS), and the two new methods described here, it will also include a simple extension of the NatureScot method, in which the rate of decay with distance is estimated from published foraging ranges rather than being fixed to be equal to minus two. [WP4](#) will also include a comparison of the apportioning values obtained using different methods, including comparisons of methods for all species for which at least two methods are now available – in contrast, the evaluations here are comparing the colony-specific spatial distribution maps underpinning the methods, and have focused specifically on the two situations in which new or extended methods have been developed. The reason the apportioning calculations are undertaken in [WP4](#), and not in this report, is because it has emerged that for each of the two new/extended methods there is an extra, non-modelling, step required in order to carry the modelling through to apportioning, and it has also become clear that a comparison of the final apportioning results can most usefully and defensibly be undertaken once the methods have been implemented within the tool. For the lesser black-backed gull analysis the additional, non-modelling, step that will be required in order to translate the spatial distribution maps described here into apportioning estimates will be an adjustment for proportion of time spent foraging on land: this was not required for the four species considered in Butler et al. (2020), but is required here because of the substantial proportion of time that lesser black-backed gulls can spend foraging on land, and because this proportion can vary substantially between colonies. For the auk GLS analysis, there are areas of the UK coast that do not lie close to any of the tracked colonies used in creating the GLS maps, and in these situations, we propose that the apportioning method will default to using BDMPS, so that the new apportioning method will use a hybrid of GLS-based maps (where feasible and defensible) and BDMPS (in other situations).

The two main sections of this report focussed on the methods and results associated with developing the two main new analyses outlined above – an analysis of summer distributions of lesser black-backed gulls (Section 2) and an analysis of winter distributions of guillemot and razorbill (Section 3) – and on evaluating these against existing approaches (SNH/NatureScot apportioning and BDMPS, respectively). Within each section a more detailed description of how the results of these analyses will be used in apportioning within [WP4](#) is provided.

2. Extending summer apportioning to lesser black-backed gull

Introduction

Lesser black-backed gulls are an important species in the context of offshore renewables as they use areas identified for offshore wind farms extensively throughout the breeding season both pre- and post-wind farm construction (Thaxter et al. 2015, 2018), and regularly fly at heights where they are at risk of collision with turbine blades. Furthermore, in response to widespread declines at breeding colonies, they have been placed on the Amber List of UK Birds of Conservation Concern (Eaton et al. 2015). Consequently, lesser black-backed gulls have been identified as being potentially vulnerable to collisions with offshore wind turbines (Garthe & Huppop 2002; Furness et al. 2013), and studies have suggested that the cumulative impact of collision has the potential for a significant negative population level impact within the wider North Sea Region (Brabant et al. 2015).

The only breeding season apportioning tool available prior to this project for lesser black-backed gulls was the NatureScot (formerly SNH) Apportioning Tool. This tool makes strong and biologically unrealistic assumptions and fails to quantify uncertainty, so there is a clear need to develop alternative apportioning approaches for this species. The MSS Apportioning Tool (Searle et al., 2020) is based on habitat use models derived from tracking data (Wakefield et al. 2017) for four species of seabird (European shag, black-legged kittiwake, common guillemot, razorbill). A key advantage of this approach is that habitat use is species and colony specific, informed by the environmental availability of each colony, and that these relationships are estimated using empirical data. High-resolution GPS tracking data are available for lesser-black backed gulls across multiple years and colonies, so we have used a similar approach to that in Wakefield et al. (2017) to develop a new apportioning method for lesser black-backed gulls.

The apportioning approach taken for lesser black-backed gulls is similar to that in Wakefield et al. (2017) and Searle et al. (2020): it uses statistical models to link relative abundance, derived from GPS tracking data, to a range of explanatory variables that relate to accessibility, competition and environmental conditions, as in Wakefield et al. (2017), and then converts these into apportioning percentages as in Searle et al. (2020). In particular, we consider the same potential explanatory variables as in Wakefield et al. (2017), although we use updated data sources, for a more recent time period, for these where possible. The statistical modelling and model selection approach is also as in Wakefield et al. (2017). However, during explanatory data analysis it became clear that some differences from the Wakefield et al. (2017) approach were needed, because lesser black-backed gulls, unlike the species considered by Wakefield et al. (2017), can spend a substantial proportion of time foraging on land. We therefore account for this, by adopting a two-stage apportioning approach: spatial distributions at sea are modelled as in Wakefield et al. (2017), and then a post-hoc adjustment is applied to the outputs of these models to account for the proportion of time spent foraging on land. The adjustment is necessary because the usage of land for foraging varies substantially between different types of colonies. We use a simple post-hoc adjustment, rather than modelling locations on land as well as sea, because of the challenges in constructing a habitat model that would provide a realistic characterisation of habitat use both of land and at sea, and because apportioning only needs to be carried out for locations at sea. We calculate the probability of apportioning to a particular colony, for a particular grid cell or area of sea, to be proportional to:

$$\text{Colony size} \quad * \quad \text{Proportion of foraging time birds from this colony spend at sea} \quad * \quad \text{Estimated probability from the model of at sea spatial distributions of being in this grid cell (or area)}$$

For each grid cell or area of sea (e.g. footprint), these quantities will be calculated for every colony, and then rescaled to sum to one. Although we explicitly account for time spent we foraging on land, we follow the approach of Butler et al. (2020) in not explicitly account for time spent at the colony – this is equivalent to an assumption that the proportion of time spent at the colony is approximately the same for all colonies. Future work could involve attempting to relax this assumption by estimating variations

between populations in the proportion of time spent at the colony, but a key challenge would be to do this in a way that would allow the results to be applied to untracked as well as tracked colonies.

Here, we focus upon presenting the data, methods and results of the new statistical modelling of spatial distributions at sea for this species using GPS tracking data. We also compare these estimated spatial distributions against the distribution implied by an inverse distance squared rule, since that is the key assumption that underpins the NatureScot apportioning tool, the existing breeding season apportioning tool for this species.

Methods

Data collection

Tracking data

Tracking data were collected for Lesser Black-backed Gulls (hereafter LBBGU) between 2010 and 2020 for nine sites (hereafter referred to also as 'colonies'; Table 1; Fig. 1(b)). For some sites, LBBGU was listed as a qualifying breeding feature of SPAs (Table 1), including: the Alde-Ore Estuary SPA (Orford Ness), Morecambe Bay and Duddon Estuary SPA (Walney), Skomer, Skokholm and the Seas off Pembrokeshire SPA, Ribble and Alt Estuaries SPA, Bowland Fells SPA, and the Forth Islands SPA (Isle of May, Fidra and Craigleith). Given the need to understand the movements of urban gull populations in the UK (Ross-Smith et al. 2014), urban sites were also studied at Barrow-in-Furness (LBBGU) and Belfast (Table 1). Further information on the colonies tracked and methods used to gather data are available from the following studies: Orford Ness: Thaxter et al. (2014, 2015), Walney: Thaxter et al. (2018), Johnston et al. (2022); Skokholm: Thaxter et al. (2019), Ribble: Langley et al. (2021); Isle of May, Fidra and Craigleith Clewley et al. (2020), Barrow-in Furness: Langley et al. (2022); Belfast (Booth-Jones et al. 2021). Additional tracking data at Bowland Fells were also available from the wider research programme of large gull tracking; however, this colony was not included as the habitat used by the tracked individuals was exclusively inland.

Table 1. Numbers of tags providing initial GPS data in each year for Lesser Black-backed gulls studied between 2010 and 2020. Individuals often contributed data across multiple years; the total number providing data across years for each colony is also shown, further broken down by the tag type used. Tag types are given as: MT= Movetech, UvA =University of Amsterdam. Also shown are the corresponding number of apparently occupied nests (AON) for each site from the Seabird 2000 census, and colony type.

			Individual GPS tracking data by year													
Colony	AON	Colony type	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	Total N	Tag type	
															MT	UvA
Walney	19,487	Coastal	0	0	0	0	24	16	38	24	14	7	0	49	5	44
Skokholm	2,419	Island	0	0	0	0	25	20	14	3	0	0	0	25	0	25
Orford Ness	5,500	Coastal	11	19	15	11	4	2	0	0	0	0	0	24	0	24
Isle of May	1,203	Island	0	0	0	0	0	0	0	0	0	28	20	28	3	25
Craigleith	391	Island	0	0	0	0	0	0	0	0	0	3	1	3	3	0
Fidra	599	Island	0	0	0	0	0	0	0	0	0	5	1	5	5	0
Barrow	85	Urban	0	0	0	0	0	0	9	19	13	6	0	31	31	0
Belfast	63	Urban	0	0	0	0	0	0	0	0	4	5	1	6	6	0
Ribble	4,150	Coastal	0	0	0	0	0	0	9	17	16	18	6	38	30	8
		TOTAL	11	19	15	11	53	44	91	79	56	76	29	238	112	126

Previous work at Orford Ness (Thaxter et al., 2017) showed the relationship between area use and the number of birds tracked, and concluded that likely enough birds had been tracked to estimate area usage for that colony. Sample sizes for other colonies aimed to be at least as high as this, although for logistical reasons this was not always feasible (for example, at Belfast).

Two different types of GPS tag were used to collect telemetry data on the movements of LBBGU in this study: (1) The University of Amsterdam Bird-tracking System (UvA-BiTS), and (2) Movetech Telemetry (a consortium of scientific partners, BTO, the University of Lisbon and the University of East Anglia and development partners Fletronic). Both tags are long-life solar-powered storage devices that record information on GPS locations, timestamps, and other parameters including speed, altitude and further instantaneous measurements on fix quality and acceleration. The two types of tags relayed information back to the user remotely; however, their means of doing so differed. The UvA-BiTS system used a base station and a network of relays that transmitted the data back via a VHF transceiver to a field-based laptop and uploaded data to a cloud-based database. The Movetech system uses GPS-GSM tags that relay to the user via the mobile phone network. For more information on the systems used see Bouten et al. (2013), Thaxter et al. (2019), and Langley et al. (2021, 2022). The use of Movetech devices was often prioritised for sites where use of a base station was prohibited or difficult to implement, including

remote sites and urban locations. Thus, the distribution of tag types differed for the colonies studied (see Table 1).

The two tag types above also varied in their ability to collect data at different sampling rates. The UvA-BiTS system collected data at finer resolutions, primarily using a base rate of five minutes but also including bursts of faster sampling up to 3 seconds. The UvA-BiTS base rate was also lowered at the colony while birds were at the nest or territory, and increased in some areas away from the colony, such as areas of offshore wind farms, both attained through use of 'geofences' implemented through the software, which helped either preserve battery (e.g. at the colony) or direct sampling protocols for research purposes (e.g. at wind farms). Movetech tags were able to collect data up to 5–15 minutes but more typically were 30–60 mins during the daytime; duty cycling was also used for Movetech tags to collect less frequent data, one fix every three hours, during times with lower light levels (ca. 2100-0600), to preserve battery. The UvA-BiTS tags were also programmed to record GPS data at lower sampling frequency during the non-breeding period of the annual cycle, whereas in Movetech tags this was typically unchanged throughout the year. Therefore, at the tail ends of the period of time when birds were 'associated with the colony' (i.e. including pre-breeding and post-breeding movements), UvA-BiTS tags were often still recording information at lower rates (typically ~30–60 minutes) either ahead of departure from the colony (ca. July/August) or after arrival in the spring (ca. February-April). More information on the size of the tags, tag models and weight increments for birds tagged can be found in Thaxter et al. (2021).

Individual adult LBBGU were captured at the nest site using a walk-in wire-mesh cage trap or a remote release noose. Tags were then attached under license from the BTO Special Methods Technical Panel (SMTP), primarily using a wing-loop harness that has previously been used successfully for LBBGU (Thaxter et al. 2016; Shamoun-Baranes et al. 2017). Initial work at the first site studied, Orford Ness, also conducted trials of harness and tag attachments for different harness configurations (Thaxter et al. 2014). All harnesses were constructed from 6.35 mm tubular Teflon ribbon (Bally Ribbon Mills, Pennsylvania, USA) to minimise abrasion and included a braided nylon core for strength. Initial deployments of tags used a 'permanent' harness design, however, from 2016 onwards, harnesses were modified to include a cotton weak-link element, allowing safe detachment of the tag, after an expected period of approximately two years (Clewley et al. 2022). A requirement of the SMTP license was all additional attachments (colour ring, harness and device combined) must be < 3% of individual body mass at time of capture to minimise the risk of negative effects of tagging (Geen et al. 2019). There was no effect of tags and their attachment on adult survival or breeding success identified through comparisons of tagged birds to a separate cohort of control birds in all studies (e.g. Thaxter et al. 2016).

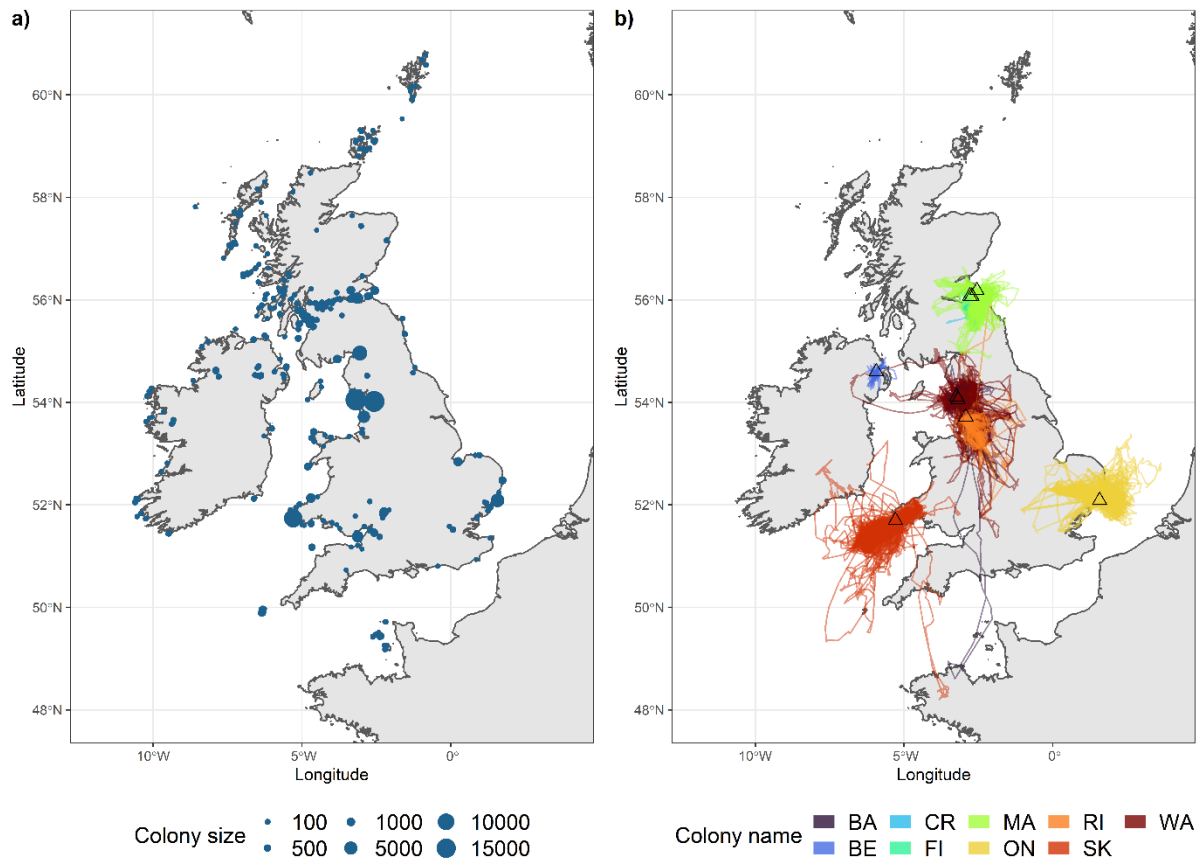


Figure 1. Lesser Black-backed gull data used to estimate the at-sea distribution of birds from UK colonies. (a) Colony size and location data from the Seabird 2000 census of colonies. Point size is proportional to the number of apparently occupied nests. (b) Tracking data of tagged individuals (2010-2020) from nine colonies in the UK (England, Wales, Scotland and Northern Ireland). Colony locations are depicted by triangles for: Isle of May (MA; 2019-2020), Craigleith (CR; 2019-2020), Fidra (FI; 2019-2020), Belfast (BE; 2018-2020), Walney (WA; 2014-2020), Barrow (BA; 2016-2019), Ribble (RI; 2016-2020), Skokholm (SK; 2014-2017) and Orford Ness (ON; 2010-2015).

Colony data

Numbers of apparently occupied nests (AON) of Lesser Black-backed Gulls (Fig. 1(a)) were extracted from the Seabird Monitoring Programme (SMP) database, prioritising the Seabird 2000 (Mitchell et al. 2004) survey of seabird colonies in Britain and Ireland (conducted between 1998 and 2002). It is acknowledged that many sites across the UK have seen large changes in numbers of gulls, such as at traditional coastal colonies and increase in urban locations (Ross-Smith et al. 2014), and further information on the AONs is available from more recent SMP monitoring data, such as for key sites covered in SMP. However, the SMP ongoing annual surveys for both natural and urban nesters is still considered uncertain for production of trends (JNCC 2021) and consequently census data should ideally be relied on for complete UK modelling perspective. The next seabird census is being conducted at present, and data were currently incomplete for inclusion within this work. The previous census (Seabird 2000) also included coverage of urban areas. Given the importance of including urban

locations for LBBGU, these urban colony counts were included within the current study for use within the modelling process.

Environmental data

A total of 10 environmental datasets were sourced and collated (Table 2), as detailed in Wakefield et al. (2017). These were (1) depth, (2) seabed slope, (3) minimum distance to coast, (4) proportion of gravel in sediment, (5) sand:mud ratio in sediment, (6) potential energy anomaly (PEA), (7) proportion of time during which the water column is stratified, (8) sea surface temperature (SST), (9) thermal front gradient density (TFGD), and (10) net primary production (alpha-chlorophyll). These covariates have been previously identified as having potential mechanistic links to different aspects of seabird ecology (Wakefield et al. 2017). Given that the time-span of the LBBGU tracking data used in this study (2010-2020) exceeded Wakefield et al. (2017) (2010-2014), we updated and replaced (deprecated) datasets where possible. The above 10 variables were obtained from four separate data sources:

- (a) ETOPO2 Global Relief 2v2, originally provided by the U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Geophysical Data Center (2006) is now deprecated and has been replaced by a higher resolution 1v1 1 arc-minute dataset <https://www.ngdc.noaa.gov/mgg/global/>; this dataset was used for depth and seabed slope (variables 1, and 2) and is described in more detail in NOAA (2021).
- (b) The British Geological Survey 1:250,000 scale sediment map (Edina digimap <http://digimap.edina.ac.uk>), was used for estimating variables 4 and 5, further translated to a numerical categorical scale (see below).
- (c) The UK Met Office FOAM AMM reanalysis dataset (<http://marine.copernicus.eu/>) (EU Copernicus Marine Service Information, 2021) was used by Wakefield et al. (2017) to estimate the next variables: (6) potential energy anomaly (PEA), and (7) proportion of time water column stratified (following Carroll et al. 2016). The same portal is available but the dataset originally used has also been superseded by a single dataset "NWSHELF_MULTITYEAR_PHY_004_009" and carry the same grid resolution (ca. 7 km) as in Wakefield et al. (2017). We used the UK Met Office Forecasting Ocean Assimilation Model Atlantic Margin model via the MyOcean website (<https://resources.marine.copernicus.eu/>), which contains the necessary key variables of potential salinity and temperature through the water column (3D netcdf dataset) that can be used to derive PEA and the time water column stratified.
- (d) Finally, the remaining four variables (7-10) were extracted from the Natural Environment Research Council Earth Observation Data Acquisition and Analysis Service (NEODAAS), <https://data.neodaas.ac.uk>. The SST data was extracted through an online visualisation tool (<https://data.neodaas.ac.uk/visualisation/>); a variety of sensors can also be specified to create a composite image, but including the Advanced-very high resolution radiometer (AVHRR). Here a multi-sensor 'ocean temp' indicator was used with NASA JPL data provider (global), using a daily interval (Version 4, NRT Refined, L4 analysis), with a resolution of 1 km (temporal data range: 2002-06-01 to 2021-08-22). Similarly, alpha-chlorophyll net primary productivity was extracted through the above visualisation tool using a multi-sensor 'ocean colour' indicator, provided by ESA CCI (global), using a daily interval (Version: Refined-OC-CCI-v5) at a resolution of 1 km (temporal data range: 1997-09-04 to 2020-12-31). This dataset included the long-running European Space Agency Ocean Colour Climate Change Initiative (OC-CCI) product (Sentinal 3A and 3B). The scale of the

accessed data was mg/m^3 for daily maps. For fronts, the AVHRR 11 μm processed by the Plymouth Marine Laboratory Remote Sensing Group (Peter Miller Pers. Com.). Strong front maps were obtained from the NEODAAS Multiview tool (<https://data.neodaas.ac.uk/multiview/>) as seven-day composite images. Together, this data source satisfies the variables of SST, thermal front gradient density (TFGD), following Scales et al. (2014) and Miller and Christodoulou (2014), and net primary production (NPP). It should be noted that there are often many sensors available for SST and net primary production; here it was assumed a multi-sensor approach was best.

Data preparation

Tracking data

Tracking data were extracted and combined from two online repositories: (1) the UvA-BiTS post-greSQL online database for the UvA tags, and (2) the Movebank data repository for the Movetech tags. As part of the data cleaning process, we removed fixes that used just three satellites to derive the geographical location, as these were deemed too imprecise for use in the analysis (Thaxter et al. 2018). We also applied a trajectory maximum speed filter of 30 m/s to remove any unrealistic fixes, as this is thought to be greater than typical maximum gull flight speeds (e.g. Shamoun-Baranes et al. 2016). On occasion, the voltage of some GPS devices dropped too low to continue monitoring of bird movements at programmed GPS sampling rates, resulting in a gap in the GPS datetime record until the tag gained sufficient re-charge to commence recording again. We used a threshold of five hours for a gap duration beyond which we deemed a gap to have occurred, and strings of points were numbered as bursts for valid monitoring periods within a tag's deployment.

Given the variation in GPS tag sampling protocols, we sought to align the datasets across tag types and colonies to a consistent rate to avoid potential sampling biases that could arise. We filtered the datasets to 30-minute, 60-minute and 180-minute resolutions using custom R functions and investigated the number of birds and fixes obtained (Table 3; Table S1). Such decisions were not straight-forward given that the common lowest denominator was a three-hour (180-minute) rate that would match the night-time protocol for Movetech tags (see above). Previous assessment of night-time ranges of birds has suggested they were smaller than daytime ranges as would be expected for a diurnal species (e.g. Scragg et al. 2016). However, no new areas were visited by birds in the night compared to the day so the only difference between night and day was range contraction (Scragg et al. 2016). This suggested that, although constituting a sampling bias for Movetech tags, the influence on relationships with habitat use are likely negligible at the scales of this study. Furthermore, the colonies where such day and night effects were likely greatest (due to the large number of Movetech tags used), were Ribble, Belfast and Barrow (Table 1). Birds tracked from Craigleith and Fidra also exclusively used Movetech tags, although notably in smaller numbers (Table 1). Ribble, Belfast and Barrow exhibited some of the lowest usage of offshore areas across all tracked colonies (see WP4), hence the influence of these effects in the modelling are considered minimal. To maximise data availability, and coverage across and within colonies, GPS data available at a 30-minute resolution were chosen to be used in this study (see Results for further discussion).

The focus of the spatial modelling in this study was the offshore environment; however, LBBGU can use a range of terrestrial, coastal, and marine habitats. Therefore, for this study we identified fixes that were truly 'offshore' by using a low-tide shapefile of the UK using simple GIS overlap analyses (using R packages 'sf' and 'sp', Pebesma & Bivand 2005, Bivand et al. 2013, Pebesma 2018), with fixes marked

as a 1 or 0 for offshore or onshore. We also sought to demarcate estuarine locations and small tributaries at coastal sites as 'inland' that would otherwise be mis-classified through the low-tide shape. Consequently, we also used a 2 m isopleth shapefile (GEBCO 2009) derived from sea depth contours to mask such fixes as inland and exclude them from being labelled as 'offshore'; a 2 m contour was identified as preferable to other contours for this purpose from trials of inclusion/exclusion under different isopleth values.

For each colony, the GPS data were annotated to indicate whether individual birds were at the 'colony' or away on central place foraging 'trips' to and from feeding areas. For gulls and many other coastal species, smaller excursions may be made to nearby bathing pools, estuaries or coastlines (e.g. for preening, self-maintenance), and within this perimeter, activities are less likely to be representative of 'true' foraging trips. A perimeter was therefore identified for each colony using bespoke polygons for each site based on a combination of (1) expert knowledge from field-workers to specify distribution of nests and areas used for non-foraging activities within each gull colony, (2) information on surrounding habitats, and (3) direct information of bird movements from the GPS records around the site. For urban sites this perimeter was typically at the level of individual rooftops where birds nested, facilitated by identified clusters of central place foraging locations. Furthermore, different rooftops were identified as the central place location (i.e. 'colony') within cities (ca. three locations in Belfast, two in Barrow), but were still treated as representative of movements from that city; see Thaxter et al. (2021) for more details. For island sites such as Skokholm and the Isle of May, the perimeter of the island plus a small 25 m buffer (R package *rgeos*, function 'gBuffer', Bivand & Rundel 2021) was taken as the colony perimeter (i.e. to include potential rocks near the colony for perching activity). For subsequent identification of time spent by birds inland and offshore (see [WP4](#)), 'trips' were assigned to birds per colony and year using the identified colony perimeters described above. Trips were sequentially numbered, with fixes at the colony identified as '0'. Only locations identified as being away from the 'colony' were used for modelling in this study; locations at the colony were excluded from analysis.

The data for each colony was initially extracted from the time that birds first arrived at the colony after the non-breeding part of the year, until the last time birds were recorded at the colony. In the case of the first tagging year for a colony, this was from the time of tagging to the end of the season. For any birds recovered later, or in the case of tag malfunction, the GPS record naturally terminated at a different date for that individual. To ensure that the tracking data was most representative of likely breeding activity, i.e. incubation or chick-rearing activity, we sought to refine this range of dates further. We used site-specific nest monitoring data that were collected as part of the wider data collection protocol; these data are normally used to monitor the fate of nests and checks for tag effects through comparison of breeding success to control nests, as mentioned above. These checks varied greatly across the colonies monitored and the years studied (e.g. from daily to 1-2 times a season – see Thaxter et al. 2015, 2019), dependent on risk of disturbance and accessibility to sites (e.g. remote locations, urban rooftops). There can also be difficulties in monitoring in particular habitats (e.g. coastal locations with tall vegetation), and at particular times of the season (e.g. the late stages of chick-rearing when chicks are highly mobile and challenging to assign to a particular nest). The fate of the nests to which individual tagged birds belonged to could not always be reliably attained. Therefore, where possible, we sought to use wider colony monitoring information at the colony level to gain an understanding of the likely wider phases of incubation and chick-rearing for a given colony each year. Monitoring typically did not start until incubation had already commenced, so we used hind-cast and fore-cast projections to the date of first egg, hatch date, and likely dates of 'fledging' based on known incubation and chick-rearing periods (BirdFacts 2005), being 24-27 days incubation and 30-40 days for

the chick-rearing period; here we took the maximum of these ranges for each colony. This resulted in a range of varying start and end dates per colony, accounting for differences in first egg dates for each year and colony. Colonies may also experience complete breeding failure at a given stage in the season, related to several factors such as predation, tidal incursion, heat stress, or food resources facilitating intra-specific aggression (Ross-Smith et al. 2014). These situations occurred at several colonies in some years, such as Walney, Orford Ness, Ribble, and Barrow, and in those situations, no GPS data beyond the point of colony failure were included as part of the likely breeding season assessment. This approach to identifying likely true breeding phases of the season can be considered precautionary by including information on wider nest monitoring beyond those initial tagged birds. However, inevitably some nests of tagged birds may have failed (or birds may have re-laid again and failed) within this wider period and thus movements that were no longer restricted to central place foraging could still be included in this analysis. Further experimental methods to deal with these complexities have been applied in another recent study using trip duration under a certain threshold to indicate central place foraging – see Thaxter et al. (2021) – but were not used here to maximise the data availability from each site given (1) the uncertainty also associated with such further refinement, and (2) the further sub-sampling of data required in this study to use within the modelling stage (as noted above).

Colony data

Colony counts of LBBGU nests (from the Seabird 2000 census) ranged from sites with ~20,000 apparently occupied nests, to sites with only a single occupied nest. To reduce computational challenges with fitting spatial models to a large number of very small colonies, we set a minimum colony size of at least 10 apparently occupied nests. This reduced the number of colonies required to be modelled from ~1100 to ~400; however, it only excluded <1% of the total overall Seabird 2000 census count. The effect of this removal on predicting the spatial distributions of LBBGU around the UK, and its resulting effect on apportioning, is therefore expected to be very limited.

Colony locations were provided as either (1) a single point location, or (2) a start and end location of the survey. For those colonies consisting of start and end locations, the midpoint was calculated for use as the colony's location in the spatial modelling. Of these midpoint locations (~18% of the total colonies), generally the start and end locations were relatively close in space, compared to the scale of the spatial modelling and other datasets used in the analysis (90% of these start-end locations were < ~3 km apart). Reducing these locations to a singular location therefore seemed the most practical approach to include these colonies in the analysis.

Environmental data

Environmental data processing and visualisation was carried out in R (R Core Team, 2021; version 4.1.2) using the packages `fasterize`, `googledrive`, `raster`, `rgdal`, `sf`, `sp`, `terra`, `tidync`, and `weathermetrics` (Pebesma & Bivand 2005, Anderson et al. 2013, Bivand et al. 2013, Pebesma 2018, Ross 2020, Sumner 2020, D'Agostino McGowan & Bryan 2021, Bivand et al. 2022, Hijmans 2022a, Hijmans 2022b). The extents of most datasets were cropped to -13 to 6 degrees longitude, and 46 to 63 degrees latitude. This geographical range was chosen based on the expected UK extent of the LBBGU colonies to include for further model predictions and the potential foraging range of LBBGUs. Some datasets, such as seabed sediments, did not cover the full range of this spatial extent and so were extracted as available.

For sea relief, the NOAA ETOP01 dataset was used as an estimate of sea depth. A further derived layer of change in sea-bed relief was also calculated to provide an up-to-date estimate of 'seabed slope' from this dataset. Land locations were masked as zero for this analysis and negative relief values translated to positive values.

Following Wakefield et al. (2017), the seabed substrate database was translated into categorical scales for two variables of sand:mud ratio and gravel proportion. We converted the shapefiles to 1 km LAEA rasters and then applied a 5 x 5 mean smoothing filter using R package raster (function 'focal'). We used the "LEX_ROCK" attribute of the shapefile data layer, as stated by Digimap as being the two-part code used to label each polygon of the Geology Digimap data. The LEX_ROCK codes matched those listed in Wakefield et al. (2017).

For potential energy anomaly (PEA) and mixed layer depth covariates, multiple products were downloaded from the Copernicus Marine Service dataset. We followed previously available Python scripts available from Plymouth-lab to estimate the PEA, using Python libraries of `numpy` and `matplotlib.pyplot`. For potential temperature and salinity 3D products, 24 bins at depths of 0 to 5000 m were extracted for the potential energy anomaly calculations (below), as well as a separate dataset of mixed layer depth as a single 2D dataset: 'ocean_mixed_layer_thickness_defined_by_sigma_theta', as defined in Kara (2003, reference depth at 3 m instead of 10 m). For estimating the PEA, we used the method outlined in Carroll et al. (2016), defined as the energy per unit depth to mix the water column, i.e. the intensity of thermohaline stratification (Holt et al. 2010). The formula for estimating PEA (φ) is defined in Carroll et al. (2016) as:

$$\varphi = -\frac{g}{h} \int_{z=-h}^0 z \left(\rho(T(z), S(z)) - \rho(\bar{T}, \bar{S}) \right) dz$$

where, g = gravitational acceleration, h = water depth (or 400 m if h exceeds this, Carrol et al. 2016), z = the vertical coordinate (0 indicating the surface, negative values indicating deeper water), ρ = density (calculated using a polynomial function - Jackett et al. 2006, Feistel 2003), T = temperature, S = salinity; here the overbar indicates that the quantity is averaged from h to the surface. Further, as data were available for discrete depths, the integral was evaluated numerically using Simpson's rule as per Carroll et al. (2016). This PEA formula gives units of mechanical energy (J) m^{-3} , and is zero for a fully mixed water column, positive for stable stratification and negative for unstable stratification. In other studies, Carrol et al. (2016), Holt et al. (2010) and Wakefield et al. (2017), for convenience φ is defined to be positive only, for stable stratification. Higher values indicate stronger stratification. 400 m was chosen to represent a consistent comparison between conditions on and off-shelf while also revealing potential deep water mixing and subsequent changes (Holt et al. 2010). Wakefield et al. (2017) also represent the PEA variable in the appendices raised to the power of 0.5, i.e. sqrt-transformed. PEA calculations thus used the accessed information on modelled salinity and temperature through the water column (up to 5000 m) to estimate in situ density.

For mixed layer depth (MLD), Wakefield et al. (2017) followed methods of Monterey and Levitus (1997), which expresses a threshold choice in terms of density change in relation to thermal expansion and thus assigns a threshold of 0.5 °C; this is stated in Kara et al. (2003). However, Kara et al. (2003) use a slightly different definition, where the MLD is when density has changed by a fixed amount, i.e. not directly through temperature. The consequence is that the MLD field is deeper with the Kara approach,

with a larger temperature difference criterion of 0.8°C, and also allows for convective mixing to remove instability from density profiles; the mixed layer depth variable using the Kara et al. (2003) approach was available through the Copernicus Marine Service dataset as a netcdf file ('cmems_mod_nws_phy-mld_my_7km-2D_P1D-m'). Sea depth here was first aligned to the extent of MLD to allow matching grid cells for further calculation, using the raster package functions 'rasterToPoints' and 'rasterize'. Grid cells were flagged as stratified if the mixed layer depth was less than the maximum water depth, and then the mean proportion of days during which stratification occurred was calculated, which results in a proportional variable bounded 0 to 1 (Wakefield et al. 2017).

The sea surface temperature and alpha chlorophyll data were processed and cropped to the same geographical extent as above, to provide daily raster layers of each variable for further use in the modelling. The units of sea surface temperature data were also converted from Kelvin to Celsius, to match the data used in Wakefield et al. (2017).

For thermal fronts, the seven-day composite .png Strong Front images from the NEODAAS Multiview portal (NW Europe area) were first downloaded and stored; maps were obtained in Mercator-projection so were first converted to geographical WGS84 using the appropriate mathematical transformation. This was achieved based on the map width and height pixel count, known longitude (-15, -13 degrees) and latitude (47, 62.999108 degrees) extents at the edges of the .png, and trigonometry; note this extent was later cropped to the main extent listed at the start of this section across covariates. As all .png files were on the same geographical extent, a base raster was determined to map all pixels onto the same grid. Daily seven-day composites were compiled to a raster stack (mosaic for the March to August period per year, at a resolution of ca. 1.2 km using R package raster (function 'stack') and stored as GTiff files using the terra package (Hijmans 2022b). Pixel map units were translated to real front metrics using the known translation from .png image units (0-256) using $\text{raster::values}(r) * 0.001$ to obtain values on the scale of °C km⁻². There are three main variables that can be derived from the front maps as accessed, for further use in habitat use analyses – see Scales et al. (2014) – namely: front density, distance to nearest fronts, and seasonal persistence of fronts over a given time frame. Here, we followed Wakefield et al. (2017) and modelled the density of fronts, over the time frame used in the analysis, thus approximating a combination of density and persistence. We used a Gaussian smooth of 5 pixels to reveal front patterns more clearly (Scales et al. 2014, Wakefield et al. 2017), achieved using R package spatialEco (Evans 2021) function 'raster.gaussian.smooth' (n = 5 pixels, sigma = 2). For simplicity of presentation among other covariates, maps were also tested through translation to a Lambert equal-area projection (see also Wakefield et al. 2017).

Across the study area, the minimum distance to the nearest coastline was also calculated in R. This was carried out using a coastline map of Europe, at a resolution of 0.5 km (as in Wakefield et al. 2017).

As the Wakefield et al. (2017) models did not extend to incorporating dynamic covariates varying through time, dynamic environmental covariates (covariates 6-10 in Table 9) were processed further to produce the final covariate maps used in the spatial modelling. LBBGU tracking data were available from 2010-2020, and covered the breeding period months of May, June, July, and August (see previous section 2.2.1). Dynamic environmental datasets (available at daily or monthly resolutions) were therefore averaged to obtain the mean of these data layers during the months of May-Aug, in the years 2010-2020.

Table 2. Processed environmental data available for use in the spatial models of Lesser Black-backed Gull distribution. Shown are the names of the 10 environmental datasets considered, their units, spatial resolution, and (for dynamic covariates) the time period covered.

	Name of environmental data type	Dates	Units	Resolution
1	Depth	-	m	1/60°
2	Seabed slope	-	°	1/60°
3	Minimum distance to the coast	-	km	0.5 km
4	Proportion of gravel in sediment	-	-	1 km
5	Ratio of sand:mud in sediment	-	-	1 km
6	Potential energy anomaly (PEA)	May-Aug, 2010-2020	J/m ³	~ 7km (0.111 x 0.0667 °)
7	Proportion of time during which water column stratified	May-Aug, 2010-2020	-	~ 7km (0.111 x 0.0667 °)
8	Sea surface temperature (SST)	May-Aug, 2010-2020	°C	1 km
9	Thermal front gradient density (TFGD)	May-Aug, 2010-2020	°C km ⁻¹	~ 1.2 km
10	Net primary production	May-Aug, 2010-2020	mg/m ³ / day	1 km

Modelling approach

Colony-specific habitat use

As in Wakefield et al. (2017), the processed GPS tracking data were first subsampled to select 24-hour bursts of locations to be used in the spatial modelling. This approach ensures that the amount of modelled data is equal across each tagged bird, and is not biased by tagging duration or individual trip duration. This was implemented by, for each tagging year at each colony, randomly selecting a 24-hour burst of at-sea GPS locations per tagged bird. Birds which had less than 24 hours of tracking data were removed from analysis.

The colony-specific habitat use of LBBGUs was modelled using an approach similar to that presented in Wakefield et al. (2017), where habitat use was modelled as a function of habitat accessibility, competition, and environmental conditions. Specifically, the intensity of tracking locations in a given location is modelled as a function of several explanatory covariates. We used the same approach implemented in Wakefield et al. (2017), whereby a Poisson point process is used to model the intensity of locations. This is implemented by constructing a regular LAEA (Lambert Azimuthal equal area) grid across the study area and fitting a weighted GLMM,

$$\lambda_{k,i} \sim \text{Poisson}(\mu_{k,i})$$
$$\log(\mu_{k,i}) = \text{offset}(\log(n_k)) + \beta_0 + \sum_{j=1}^m x_{i,j}\beta_j + u_k$$

where $\lambda_{k,i}$ is the intensity of locations of birds from the k^{th} colony at point i in space, and $\mu_{k,i}$ is the expected number of tracking locations at sea per bird per day per unit area from the k^{th} colony. The expected number of tracking locations $\mu_{k,i}$ is then considered a function of: an intercept term β_0 , an offset term n_k for the number of birds tracked from each colony k , a random colony-level effect u_k to account for differences between colonies, and a vector of explanatory covariates.

The weighting adjusts for variations in sample sizes between colonies, in order to ensure that data from colonies with small sample sizes, at which area use may not be fully characterised, are given less weight in determining relationships than the data from colonies with larger sample sizes.

To consider habitat accessibility in the models, and to estimate the colony-level random effects, it was necessary to consider the area of sea which is accessible to each colony. Following Wakefield et al. (2017), we considered accessible area as lying within d_{max} of the colony location, where d_{max} is defined as 1.1 x the maximum foraging range observed in the tracking data used in this study. For LBBGU, this provided a d_{max} of 675 km. Multiplying the maximum observed distance from the colony by 1.1 provides a buffer around the maximum expected foraging range (Wakefield et al., 2017).

To implement the models, analyses were conducted on a quadrature grid, using the LAEA projection for all datasets. Relative to the species modelled in Wakefield et al. (2017) (shag, kittiwake, razorbill, guillemot), the LBBGU data were of coarser time-resolution (30 minute vs. 100 second), and less individuals had available bursts of at-sea tracking data (due to high land use). To account for this, and to fit models within computationally practicable timescales, a grid size of 10 km was chosen for the analysis of LBBGU habitat use conducted here.

All analyses and modelling were carried out in R (version 4.2.0; R Core Team, 2021), and GLMM models were fit using the R package `lme4` (Bates et al., 2015). For further details of the statistical methodology and background, see Wakefield et al (2017).

Explanatory covariates

Several stages of modelling were carried out, following the approach adopted in Wakefield et al. (2017), to incorporate explanatory covariates. In order, the incorporation of these covariates allowed us to construct models that estimated the relationship between LBBGU habitat use and (1) colony distance, (2) area of available sea, (3) competition, and (4) environmental covariates.

First, as it is assumed that usage of areas will decline as distance from the colony increases (due to central-place foraging theory), distance to the colony d was added as the first potential explanatory covariate (Wakefield et al., 2017). Second, it is assumed that, due to variations in coastal geometry and competition between colonies to access at-sea habitats, the availability of sea to birds from different colonies is likely an important factor in defining distribution (Wakefield et al., 2017). Thus, we considered the cumulative area A of available sea to different colonies as the second potential covariate to be included in the models. Following Wakefield et al. (2017), this covariate was log-transformed to reduce collinearity between A and d .

The third group of potential covariates to be included were those relating to competition. It is assumed that the habitat use of seabirds may be related to the number of breeding birds present at, or nearby to, the colony (Wakefield et al., 2017). Similar to Wakefield et al., (2017) we calculated a number of metrics summarising potential “sympatric competition”: (1) the number N of apparently occupied nests within range d of the colony, (2) $\sqrt{N}$, (3) N weighted by inverse-distance, (4) $\sqrt{N}$ weighted by $\sqrt{\text{inverse-distance}}$, (5) $\sqrt{N}$ weighted by inverse-distance, and (6) N weighted by $\sqrt{\text{inverse-distance}}$. As in Wakefield et al. (2017), we considered the addition of each of these terms as singular terms and as interaction terms with the cumulative area $\log(A)$. See Wakefield et al. (2017) for further details and justifications of each of these variables. Further, it is possible that the density of birds at-sea is an important factor in defining distribution, as birds may avoid particular locations where the density of birds from other colonies is high (termed “parapatric competition” in Wakefield et al. (2017)). As in Wakefield et al., (2017), we used the best-fitting model at this stage to predict the ratio $p_{h,i}$ of the intensity of locations $\lambda_{h,i}$ from the focal site h , relative to the sum of those from all other sites within range ($\sum_{k \neq h} \lambda_{k,i}$), and considered this as a covariate within the models.

In the calculations above, to follow Wakefield et al. (2017), there was a need to calculate distance by sea, in particular for the calculations of accessible habitat from each colony. In Wakefield et al. (2017), as the seabird species studied almost exclusively avoided land, distance by sea was calculated using a UK coastline outline with a high penalty for crossing land. For LBBGUs this approach is not as appropriate, due to the high use of terrestrial habitats and the observation that many tagged individuals crossed large areas of land to reach more distant marine habitats (Fig. 1(b)). To implement these models, in as similar an approach as possible to Wakefield et al. (2017), we instead used an internal (inland) buffer of 20 km of the UK coastline. This approach allowed unpenalised movement across land outside the buffer, but prevents the model allowing movement across the buffer (i.e. stops spillover in predicted distributions between the east and west coasts of the UK). The size of this buffer was chosen after considering different buffer size of 10, 15 and 20 km; 20 km seemed the most appropriate choice given the large areas of land covered by many tracked LBBGUs. The use of this buffer resulted in the

exclusion of three (untagged) colonies from the analysis, as they were considered too far from the true coastline to access marine habitats.

In the last stage of modelling, environmental covariates were considered. Following Wakefield et al. (2017), ten environmental covariates (Table 2) were considered as potential indices describing LBBGU habitat use. To first determine the order in which to add environmental covariates to the model, the best-fitting model at this stage was used to fit a model with each environmental covariate singly; environmental covariates were then added in order of the expected improvement in BA score in the model, from largest to smallest (Wakefield et al., 2017). Different transformations (log, square-root, squared) of each of the environmental covariates were also fit when each environmental covariate was included. To account for potential differences in habitat (environmental) variability between colonies, we also considered interaction terms between each environmental covariate and its expected value at each colony. This was implemented through calculating the covariate's mean in waters accessible from each colony (i.e. habitat within $d_{\max}$), as described in Wakefield et al. (2017). All covariates were standardised prior to model fitting.

Model selection

Model selection followed the same approach as Wakefield et al. (2017), where covariates were added sequentially in the order described above. Due to the large size and computational challenges of fitting such complex spatial models, stepwise forward model selection was considered to be the most practical approach. At each stage, models were evaluated using cross-validation. Each colony was, in turn, left out from the model, and the observed utilisation distribution of birds from that colony was compared to the predicted utilisation distribution from the model fit using the remaining colonies (Wakefield et al., 2017). The weighted Bhattacharyya affinity ($\overline{BA}$) was calculated between the observed and predicted utilisation distributions,

$$\overline{BA} = \frac{\sum_{All\ k} n_k BA_k}{\sum_{All\ k} n_k}$$

where BA_k is the Bhattacharyya affinity for colony k , and n_k is the number of tracked birds from colony k . Colonies with larger numbers of tracked birds are therefore given an increased weighting in this calculation, as it is assumed that the utilisation distribution of colonies with smaller numbers of tracked birds is likely to be an underestimate of the true full area used by the entire colony (Bogdanova et al., 2014; Wakefield et al., 2017). At each step, covariates were retained within the model if they resulted in an increase in $\overline{BA}$; if the $\overline{BA}$ was equal between models, the simpler model was selected. BA (and weighted BA) ranges from 0 (indicating no similarity) to 1 (identical utilisation distributions).

Estimating usage

The best-fitting model (as determined by improvement in $\overline{BA}$ score) was used to estimate LBBGU at-sea density from each of the Seabird 2000 colony sites. Colony densities were also combined to estimate a population-level utilisation distribution across the study area, where utilisation distributions from individual colonies were multiplied by colony size (number of apparently occupied nests) to obtain the overall population-level distribution. To quantify uncertainty, we followed Wakefield et al. (2017) in carrying out a parametric bootstrap. For the best-fitting model, we assumed a multivariate normal distribution and generated 10 random sets of fixed-effects parameters, predicted the utilisation distributions using these parameters, and then calculated the coefficient of variation (CV). Note,

however, that this approach is likely to underestimate uncertainty, potentially substantially, because the models do not account for residual spatial autocorrelation, and so the results (particularly the absolute levels of uncertainty) should be treated with considerable caution.

Comparison with existing apportioning approach

To compare the new models used here against the assumed spatial distribution that underpins the existing approach (the NatureScot apportioning tool), utilisation distributions were generated and mapped for some example colonies. Utilisation distributions with the core 50%, 75% and 90% of estimated usage were calculated using the predictions from the LBBGU distribution models developed in this study. These were compared to some illustrative utilisation distributions to represent the assumptions that underpin the NatureScot apportioning tool approach, by using the same spatial grid to plot the potential density (and utilisation distribution) generated when assuming that the distribution of birds at sea is proportional to the inverse of distance by sea to colony squared (rescaled to sum to 1, for comparison). Note that this is not a comparison against the results actually obtained using SNH/NatureScot apportioning (as that does not explicitly have a utilisation distribution associated with it), but against the assumption of an inverse distance squared relationship to distance from colony that provides the rationale for the SNH/NatureScot apportioning tool.

Results

Tracking of lesser black-backed gulls

Tracking data from a total of 207 birds were available for potential inclusion within the modelling. The selected rate of 5 minutes included more fixes (nearly 1.5 million), but fewest colony spread and number of individuals (given largely being limited to UvA-BiTS data). Alternatively, rates of 180 and 60 minutes included the same number of individuals and most spread across colonies, but greatly reduced the number of fixes available for the modelling. Therefore, the 30-minute rate was selected (167 birds, eight colonies, 49,562 offshore fixes). Note, this 30-minute rate resulted in a reduced number of fixes and birds for Barrow and Belfast (Table 3); however, these sites also exhibited some of the lowest proportions of offshore usage (Table S1).

Table 3. Total number of individual (unique) birds available per colony after filtering GPS data to common rates of five, 30, 60 and 180 minutes, to assess potential sample sizes for inclusion in the modelling.

Colony / filter rate	5 minute	30 minute	60 minute	180 minutes
BA (Barrow)	8	10	30	30
BE (Belfast)	-	-	6	6
CR (Craigleith)	1	3	3	3
FI (Fidra)	-	3	5	5
MA (Isle of May)	25	28	28	28
ON (Orford Ness)	23	24	24	24
RI (Ribble)	17	26	38	38
SK (Skokholm)	25	25	25	25
WA (Walney)	47	48	48	48
Total birds	146	167	207	207

To standardise the amount of tracking data before modelling, data bursts of 24 hours were randomly selected from each tagged individual in each year of tag deployment (2010-2020). By this approach, a total of 220 movement bursts were selected from the 30-minute resolution tracking data (Table 4). As at least 24 hours of offshore tracking data was required (and many LBBGUs had high use of terrestrial habitats), this process resulted in the removal of many individuals from the tracking dataset. In some cases, this reduced the number of tracking data bursts available per colony to 0 or 1. Initial model fitting with some colonies containing only a single data burst led to challenges in model convergence and problems with cross-validation (used here in model selection); colonies which only had a single burst of tracking data available were therefore removed from analysis. The final data used in the modelling (Table 4) therefore consisted of tracking data from six colonies, with three of the tagged colonies

removed (Belfast, Craigleith, Fidra). It should be noted that, for some colonies, the number of tracking data bursts available is higher than the number of unique individual LBBGUs tagged (Table 3; Table 4). This is because, to improve data availability and coverage across the temporal range of the study, we selected bursts from each tagged individual per year (i.e. if a bird was tagged in two years, a burst from each year may be selected; Table 1). This approach assumes that any bias caused by selecting the same individual multiple times is negligible compared to variation between years.

Table 4. Resulting number of 24-hour GPS data bursts included in the spatial modelling from each tracked colony. Also shown are the number of unique individual birds these data bursts came from (as some individuals were tracked across multiple years).

Colony	Number of tracking data bursts	Number of individual birds
BA (Barrow)	7	5
BE (Belfast)	-	-
CR (Craigleith)	-	-
FI (Fidra)	-	-
MA (Isle of May)	30	23
ON (Orford Ness)	30	19
RI (Ribble)	7	7
SK (Skokholm)	38	21
WA (Walney)	107	44
Total number:	220	119

Models of distribution

The best-fitting model of LBBGU distribution (Table 5) contained: distance to the colony d , cumulative area of available habitat $\log(A)$, $\sqrt{N}$ (a measure of competition), an interaction effect between $\log(A)$ and $\sqrt{N}$, and an environmental effect of $\log(\text{depth})$. At each stage of forward model selection, the inclusion of these variables increased the $\overline{BA}$ score. Although including $\log(\text{depth})$ improved the $\overline{BA}$, the improvement was relatively small, and so a simpler model with no environmental covariates was also included here for comparison. A final decision on which model will be incorporated within the apportioning tool, or whether users will be able to select between models within the tool, will be decided in WP4. Model assessment by cross-validation estimated a $\overline{BA}$ of 0.384 (SD = 0.128) for the final model with $\log(\text{depth})$ and a $\overline{BA}$ of 0.370 (SD=0.125) for the final model with no environmental covariates. Whilst the initial model fitting of singular environmental covariates (to determine the order in which to add covariates) indicated that several of the environmental covariates would increase the $\overline{BA}$ (Table S2), after $\log(\text{depth})$ was added no further additions improved the $\overline{BA}$.

Table 5. Model parameters and model selection process of the resulting final models. Shown are the covariates retained within the final model (name, estimate, standard error and z-statistic). Also shown are the weighted Bhattacharyya affinity ($\overline{BA}$) scores from the cross-validation approach used in model selection. Covariates were added in the order listed below, and model selection stopped when the change in $\overline{BA}$ ($\Delta\overline{BA}$) was no longer positive; $\sqrt{N}$ and the interaction term between $\log(A)*\sqrt{N}$ was added in a single stage, providing a single value for change in $\overline{BA}$ (denoted by the asterisk *).

a) Model with $\log(\text{depth})$

Covariate	Estimate	Standard error	z	$\overline{BA}$	$\Delta\overline{BA}$
Intercept	-17.364	0.804	-21.61	-	-
d	-6.751	0.316	-21.34	0.300	-
$\log(A)$	-1.068	0.053	-20.18	0.334	0.034
$\sqrt{N}$	-0.646	0.560	-1.15	*	*
$\log(A) \times \sqrt{N}$	-0.253	0.039	-6.47	0.370	0.036
$\log(\text{depth})$	-0.931	0.034	-27.15	0.384	0.384

b) Model with no environmental covariates

Covariate	Estimate	Standard error	<i>z</i>	$\overline{BA}$	$\Delta\overline{BA}$
Intercept	-19.707	0.772	-25.51	-	-
<i>d</i>	-8.102	0.344	-23.57	0.300	-
$\log(A)$	-1.270	0.053	-23.74	0.334	0.034
$\sqrt{N}$	-1.290	0.488	-2.64	*	*
$\log(A) \times \sqrt{N}$	-0.650	0.038	-17.11	0.370	0.036

Estimated population-level distribution

Predictions of LBBGU usage were made for each of the 382 colonies and saved for incorporation into [WP4](#). The combined population-level usage map (Fig. 2) shows that gull at-sea usage largely declines with distance from the coast. As expected, areas of relatively high usage tended to coincide with areas of large, or many, LBBGU colonies (Fig. 1(a)); namely around Walney, off Pembrokeshire, and around the central belt of Scotland. Uncertainty in the predicted densities (CV; Fig. 3) generally increased as distance from the coast (and from colonies) increased; note, however, the earlier caveat that uncertainty may be being underestimated within this approach, so that the results of this should be treated with caution.

t

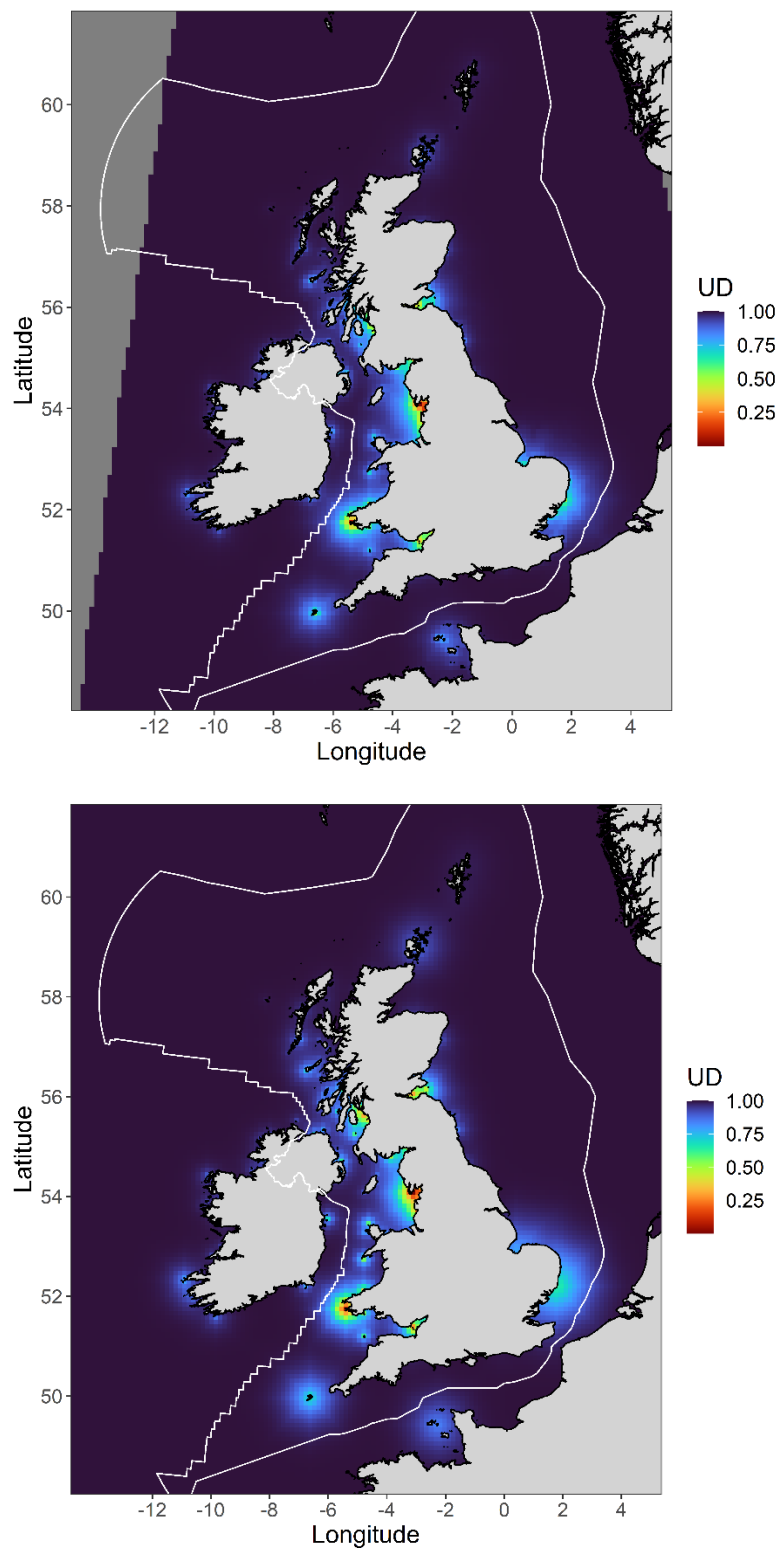


Figure 2. Cumulative at-sea utilisation distribution (UD) of lesser black-backed gulls in the breeding season in Britain and Ireland. Lesser black-backed gull distribution was estimated from a spatial model incorporating colony distance, habitat availability, competition, and with (top) or without (bottom) environmental variables. Warmer colours represent areas of higher usage. Predictions were not made for the areas in grey. The tool will only apply to areas of sea within the UK EEZ (white outline).

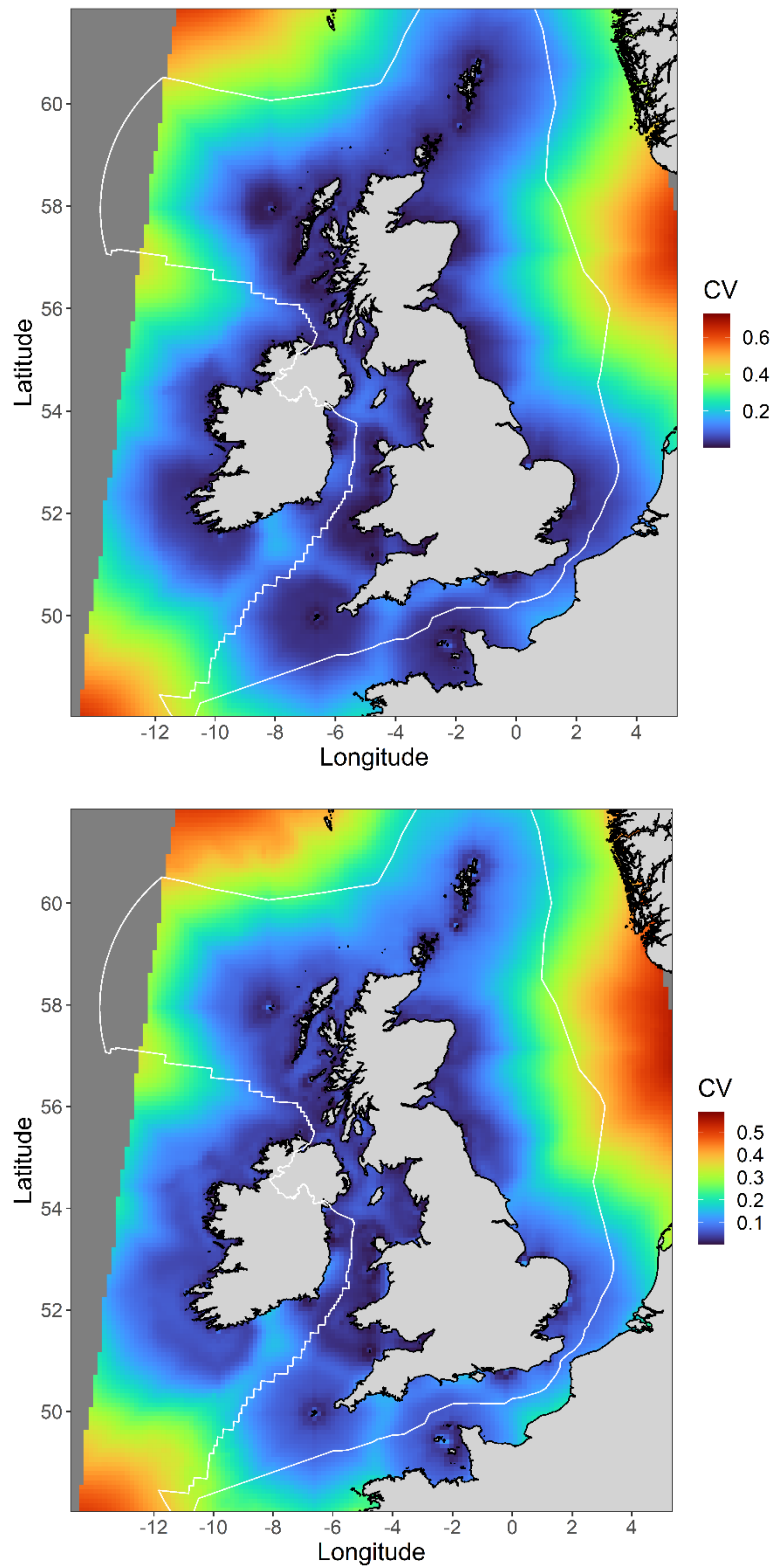


Figure 3. Coefficient of variation (CV) of the at-sea utilisation distribution (UD) of lesser black-backed gulls in the breeding season in Britain and Ireland for models with (top) and without (bottom) environmental variables. CV was estimated by parametric bootstrap using 10 randomly-generated sets of parameters. Warmer colours represent a higher CV on the estimated density. Predictions were not made for the areas in grey. The tool will only apply to areas of sea within the UK EEZ (white outline).

Comparison with existing apportioning approach

Example utilisation distributions were estimated for two of the colonies with GPS tracking data: Skokholm (Fig. 4) and Orford Ness (Fig. 5). It should be noted that the utilisation distributions produced to represent the NatureScot approach correspond to the inverse distance squared assumption that underpins the approach, but the NatureScot approach does not explicitly calculate or use a utilisation distribution, so this comparison is designed to evaluate the realism of this assumption rather than the actual apportioning values calculated using the NatureScot tool (that comparison will only be possible once an adjustment for proportion of time foraging on land has also been applied, in [WP4](#)). Note also that there is some discussion for gull species as to whether it is most appropriate to apply the NatureScot apportioning approach to distance by sea, distance by air, or a hybrid of the two – this will also be evaluated more thoroughly in [WP4](#), once the adjustment for proportion of time foraging on land has been applied.

Given the large potential accessible area by LBBGUs (max foraging range of 675 km), calculating the utilisation distribution using the NatureScot approach (density is proportional to $1 / (\text{distance by sea}^2)$) leads to the inclusion of many low-density grid cells. The estimated utilisation distribution is therefore very sensitive to the cut-off used to exclude grid cells (i.e. when to consider near-zero densities far from the colony as zero); the addition of lots of very small densities changes the predicted extent of the core areas of the utilisation distribution. We therefore included two different utilisation distributions with a range of cutoff values used to exclude grid cells with near-zero densities (Fig. 4(c-d); Fig. 5(c-d)). The core areas of the utilisation distribution for Skokholm (Fig. 4(a-b)) estimated using the LBBGU models developed here differed from that predicted by the NatureScot approach; they were more localised and the overall distribution of the colony is slightly more elongated (Fig. 2), and more closely matching the tracking data from this colony (Fig. 1(b)). In contrast, the distributions predicted for the Orford Ness colony are quite similar between the different approaches (Fig. 5).

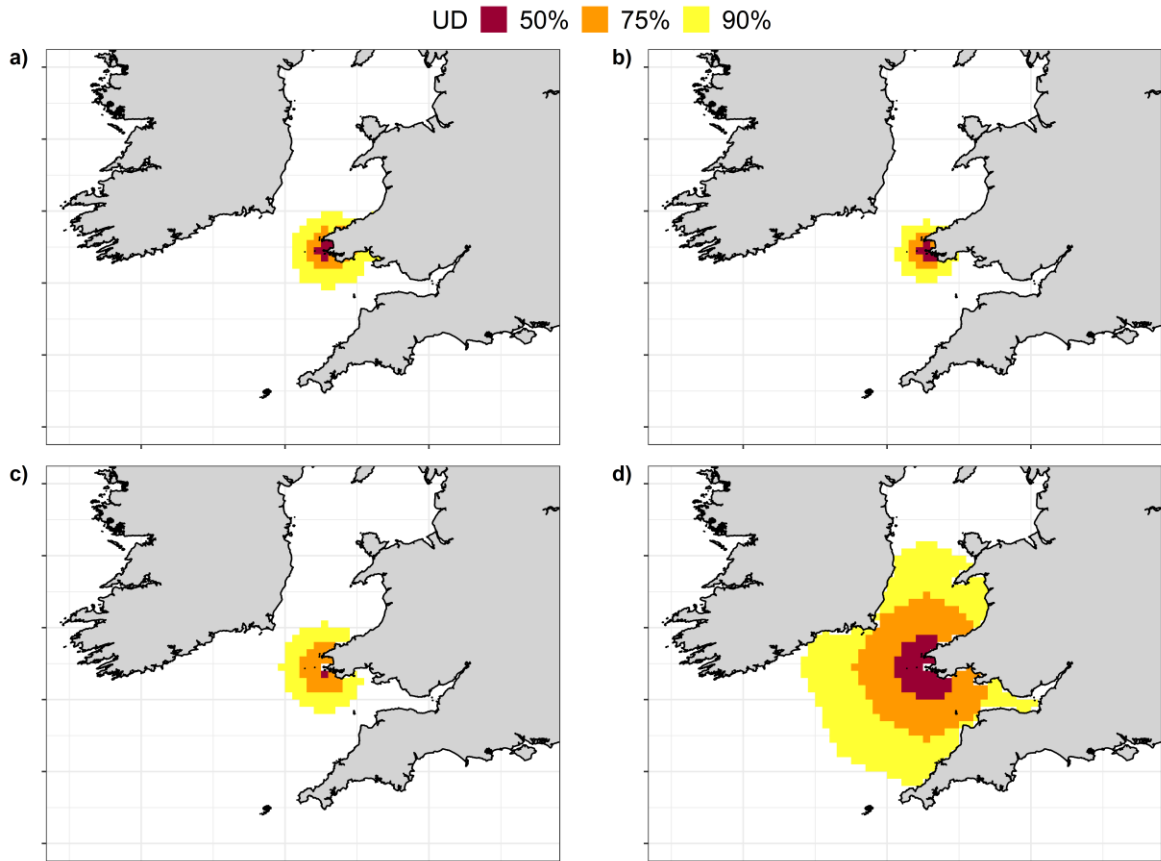


Figure 4. Utilisation distributions of lesser black-backed gulls (LBBGU) from Skokholm colony. (a) Estimated utilisation distribution predicted from a model with log(depth). (b) Estimated utilisation distribution predicted from a model with no environmental covariates. (c-d) Illustrative utilisation distributions using the NatureScot ($1/\text{distance}^2$) approach, using different cutoff values for excluding grid cells with near-zero density: (c) 0.001, (d) 0.0001.

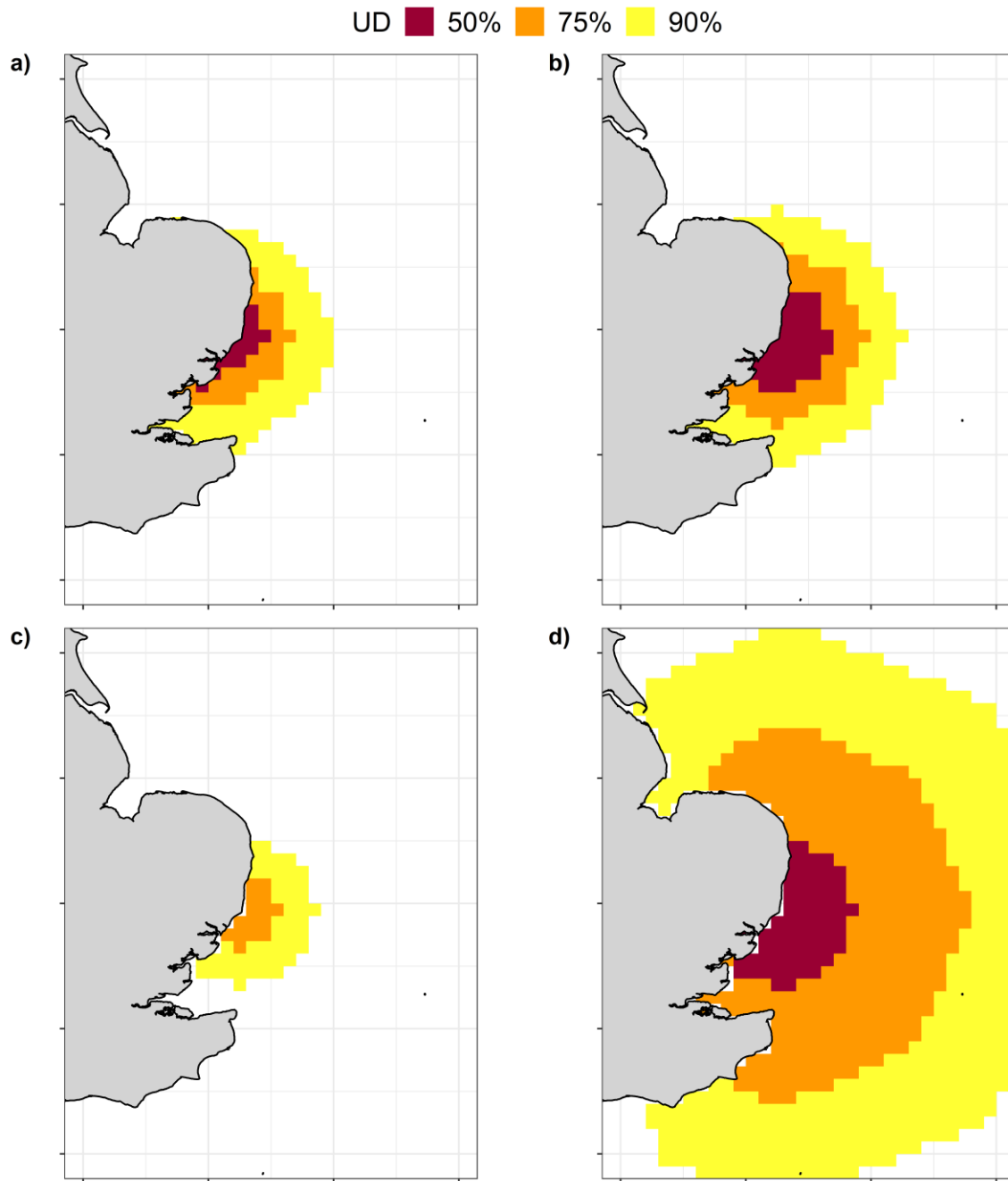


Figure 5. Utilisation distributions of lesser black-backed gulls (LBBGU) from Orford Ness colony. (a) Estimated utilisation distribution predicted from a model with log(depth). (b) Estimated utilisation distribution predicted from a model with no environmental covariates. (c-d) Illustrative utilisation distributions using the NatureScot ($1/\text{distance}^2$) approach, using different cutoff values for excluding grid cells with near-zero density: (c) 0.001, (d) 0.0001.

Discussion

In this study, apportioning approaches were extended to lesser black-backed gulls, by developing and carrying out a method to estimate colony-specific habitat use. The approach used was similar to that proposed by Wakefield et al. (2017), using GPS data from tracked colonies to understand the relationship between bird habitat use and several covariates, and using this model to predict the distribution of birds from untracked colonies around the coasts of Britain and Ireland. This approach represents an improvement on the currently used approach, as it incorporates more biological realism (measures of competition and habitat availability are incorporated), it is based on tracking data from LBBGUs within these colonies, and it is consistent with the approach currently used for other species of seabird (kittiwakes, shags, razorbills, guillemots).

Lesser black-backed gull distribution was found to be related to distance from the colony, habitat availability, competition, and the at-sea environment (depth). The best-fitting LBBGU model was relatively more similar to those of shags and razorbills in Wakefield et al. (2017), as distribution was related to metrics of “sympatric” competition, but not related to the estimated measure of “parapatric” competition. Cross-validation, using the weighted Bhattacharyya affinity ($\overline{BA}$), indicated that the LBBGU models ($\overline{BA} = 0.38$) performed more poorly than the Wakefield et al. (2017) models for shags, kittiwakes and guillemots ($\overline{BA} = 0.52 - 0.53$), but marginally better than the Wakefield et al. (2017) models for razorbills ($\overline{BA} = 0.34$).

The approach developed here used LBBGU colony counts and locations from the Seabird 2000 census. It is however acknowledged that the methods used within Seabird 2000 and subsequent monitoring, predominantly being vantage point methodology (78% counts, others being aerial and ground-based surveys, Mitchell et al. 2004), could underestimate number of urban gull populations. Further work has investigated the influence of survey methodology in deriving reliable urban gull population estimates for urban sites in England and Wales for Herring and Lesser Black-backed Gull (Woodward et al. 2020, Burnell et al. 2021b). At the time of writing, a UK-wide set of population estimates for urban areas was not yet available, but further studies are ongoing using different survey methodologies within a model-based approach (Burnell et al. 2021a, b). Further, there have been large changes in the numbers of gulls at different colonies over the last two decades, in particular at urban colonies (Ross-Smith et al. 2014). Updating the colonies used within this modelling framework, in particular determining a robust and up-to-date estimate of LBBGU colonies across Britain and Ireland would be a valuable improvement to the estimates given here, although we acknowledge that this is not necessarily straightforward.

Tracking data were available for LBBGUs from nine colonies; however, due to variability in tagging time-resolutions, the criteria used in selecting suitable data for modelling, the high use of terrestrial areas by tagged individuals, and the statistical modelling framework used, only six of these colonies were able to be used in the final models of gull distribution. It should therefore be noted that, whilst predictions are made for all colonies, these are based on GPS-tracking data from these six colonies. These colonies represent a mix of urban, coastal, and island colonies, but are largely based on tracking data from coastal and island locations. Whilst this represents the best information that we have on LBBGU offshore movements, further tracking data from urban colonies, and from spatial regions with limited/no tracking data would help improve our understanding of LBBGU at-sea distribution. The modelling of area use of LBBGU also included movements of birds during all behaviours. For example, commuting to and from a foraging location over the ocean would also be included in association with environmental features. Such relationships are expected to be strongest, however, when birds are

actively engaged in searching or foraging activity, hence further refinement to models could also incorporate behaviour to more appropriately characterise habitat relationships.

The analysis adjusted for variations in sample sizes between colonies via weighting, and the data collection made use of previous analyses (Thaxter et al., 2017) that aimed to characterise the minimum sample size needed to capture area use, but it would be useful to investigate representativeness further, not least because we acknowledge that the smaller ranges seen at some of the urban sites may be the result of relatively small sample sizes. Simple approaches (e.g. looking at the relation between UD contour areas and number of birds for each colony) are not particularly meaningful or feasible in this context, in part because the models are too computationally intensive to be refitted many times, and in part because the models are not fitted separately for each colony (so that the predicted area use for a colony does not depend only on the number of birds tracked at that colony, but also on the total number of tracked colonies and the numbers of birds tracked at other colonies). A more thorough evaluation of the impacts of different sample designs would therefore be required.

The models used here quantify seabird distribution in relation to a range of covariates of interest, and here we also adapted this model for LBBGUs, allowing movement across areas of land when assessing habitat accessibility. Although it may involve substantial statistical development work, and would require testing to ensure that the models remain quantitatively robust, one potential improvement to these models would be to incorporate the ability to use tracking data of different time-resolutions (e.g. to include 5-minute, 30-minute, 60-minute tracking data within the same framework, using fine or coarse-scale tracking data as available). This would maximise the use of available tracking data for different species. One potential way this could be approached is through including an additional offset term within the models, similar to that used to offset densities by the number of tracked individuals at each colony. Further work with the existing models could also be carried out to investigate the potential effects of other covariates, e.g. density of fishing vessels, which may affect species distribution. As part of this work, distance to fronts was also extracted using the methods of P. Miller pers comm (see Scales et al. 2014) but for simplicity was not included here for alignment with Wakefield et al. (2017); this could also be investigated in further work. Finally, considering the differences observed in colony locations, in colony trends, in terrestrial habitat use, and in maximum foraging ranges between urban and non-urban gulls, considering a modelling framework where these two groups are considered and modelled separately (and then integrated), may be a future research avenue. For example, it is possible that the relationship between gull density and marine environmental variables may be different for coastal vs. urban gulls.

Overall, the work carried out in this WP has developed and produced estimates of lesser black-backed gull distributions for colonies around Britain and Ireland. Due to the differences in LBBGU use of land vs. sea between different types of colonies (e.g. island, urban, coastal), the predicted densities generated here will be adjusted by the estimated proportion of time spent foraging on land per colony, implemented in the apportioning tool within [WP4](#).

The comparison against the inverse distance squared assumption that underpins the NatureScot apportioning method suggest broad similar patterns between the GPS-based maps and those obtained under this assumption, reflecting the strong central place foraging constraint for this species in summer months. The distributions do differ, however, and the inverse distance decay squared method is sensitive to the choice of cut-off (imposed here as a minimum small probability below which the probability is assumed to be zero). The comparison against the NatureScot method, and, in particular, the implications of this sensitivity, will be explored more in [WP4](#); after adjustment for proportion of time

spent foraging on land the results of the GPS-based approach can be compared directly to apportioning estimates from the NatureScot method. [WP4](#) will also include a simple extension of the NatureScot method to allow the rate of decay with distance to be estimated based on published foraging ranges, rather than fixed to be equal to minus two, and this extended method will also be compared against the GPS-based maps.

3. Extending apportioning to winter using geolocation data

Introduction

A significant challenge with apportioning seabirds in the non-breeding season to protected SPA colonies is that standard methods of estimating distribution (i.e. at-sea surveys) provide no information on the provenance of birds. As a result, the current method for apportioning in the non-breeding season (Biologically Defined Minimum Population Scales or BDMPS; Furness 2015) makes strong and unrealistic assumptions about the spatial distribution of birds in relation to colony provenance and does not always make use of the most recent available data.

Bird-borne instrumentation provides information on provenance, and the availability of loggers that can be carried by individuals throughout the annual cycle provide new opportunities to undertake apportioning of seabirds at sea in winter with greater accuracy. Geolocation devices have been widely used to identify the wintering locations of a range of seabird species from different colonies (e.g. Fort et al., 2012; Frederiksen et al., 2012; Glew et al., 2018). Geolocation loggers record ambient light levels with day length being used to estimate latitude and the timing of midday relative to GMT being used to estimate longitude (Phillips et al., 2004). The loggers have the advantage of being considerably smaller and lighter than GPS tags meaning that they can be deployed on a wider range of species for longer time periods. However, the reliance on light levels to estimate location in combination with a limited number of fixes per day means that there can be considerable uncertainties surrounding the inferred locations, in the range of 100-200 km (Merkel et al., 2016). Despite these uncertainties, geolocation data can offer valuable insights into the movement and distribution of seabirds of known colony provenance during the non-breeding season that could not be obtained using other methods.

In order to develop data-driven methods for apportioning birds outside the breeding season, a new opportunity has arisen from a rich geolocation data set from multiple colonies in the UK for guillemots and razorbills (see next section for details). As geolocation data lack the spatial and temporal resolution of GPS data, habitat modelling approaches, such as that of Wakefield et al. (2017) are markedly more challenging to use. Alternatively, these data can be used to estimate colony-specific Utilisation Distributions (UDs) for tracked colonies, and these distributions can then be extrapolated in some way to also provide distributions for untracked colonies (e.g. through a weighted average of distributions from tracked colonies). These distributions, in combination with estimates of colony size, can enable birds to be apportioned back to their breeding colonies in the non-breeding season. Modelling of spatial distributions (e.g. colony specific UD) from geolocation data has similarities with the modelling of GPS data, but there are some important differences:

- (a) Geolocation data are much lower frequency than GPS data – typically 1-2 records per day, which means detailed modelling of behaviour and local spatial movement is not possible.
- (b) Levels of observation error in geolocation data are much higher than for GPS data, and are sufficiently large that models which ignore observation error are unlikely to be defensible.
- (c) The levels of observation error in geolocation data are likely to be heterogeneous, and it seems that as they vary according to known factors (e.g. time of year) this variability can be modelled.

These differences mean that the methods used to build models that can be used to apportion in the non-breeding season will necessarily differ from those used in the breeding season.

Methods

Geolocation data description

Geolocation (hereafter 'geolocators'; Biotrack model MK3006 and Migrate Technology Intigeo models C65, F100 and C65-Super) were deployed on 695 common guillemots (hereafter 'guillemots') *Uria aalge* and 339 razorbills *Alca torda* during three breeding seasons (June-July 2017, 2018 and 2019) at twelve breeding colonies around the north of the UK, which were relatively evenly distributed through the breeding region of guillemots and razorbills in the UK (Buckingham et al. 2022). The geolocators measured light levels, salt water immersion and sea-surface temperature, from which daily locations were estimated throughout the non-breeding season. Deployments were made under licence from the British Trust for Ornithology.

Individuals were caught at the nest using a noose pole during late incubation or chick rearing (mid-June – early July). Birds were equipped with a unique metal ring (if not already present) and a geolocator mounted on a plastic colour ring on the opposite leg. The maximum mass of the geolocator plus colour ring (guillemots = 4.8g; razorbills = 4g) comprised 0.63% (guillemots) and 0.79% (razorbills) of the minimum body mass recorded in breeding adults of the two species in Britain (guillemots: 765g; razorbills: 505g; Wagner 1999, Harris et al. 2000). Adults were recaptured during the 2018, 2019, 2020 and 2021 breeding seasons, resulting in a deployment duration of one to four years. For analyses, data were pooled from three non-breeding seasons: 2017-18, 2018-19 and 2019-20 (Table 6). In all cases, handling times did not exceed ten minutes during either device deployment or retrieval. There was no concrete evidence that the tags had any adverse impacts on the birds' behaviour. Observations of birds following logger deployment was that they returned to normal breeding behaviour soon after release. However, it was not possible to monitor all birds post-release in a systematic way because they would have increased disturbance. At retrieval, all tagged birds were examined for any leg injury caused by tags or colour rings they had been carrying, and no injuries were observed. It was not possible to quantify the potential effect of carrying the logger on foraging efficiency and demographic rates such as productivity and survival. However, three tagged birds were found dead on the shore over the 2017-18 winter, and it was important to establish whether this was a higher than expected number. Typically, under 10% of adult guillemots and razorbills will die each year. Since 436 birds were tagged in 2017, we would expect ~40 of those to die the following winter. The reporting rate of ringed guillemots and razorbills that die is approximately 10%, so we would have expected approximately four of the 40 birds that die to be reported, which is close in number to the three found dead on the shore. This suggests that the presence of the colour ring and tag did not have a measurable impact on survival.

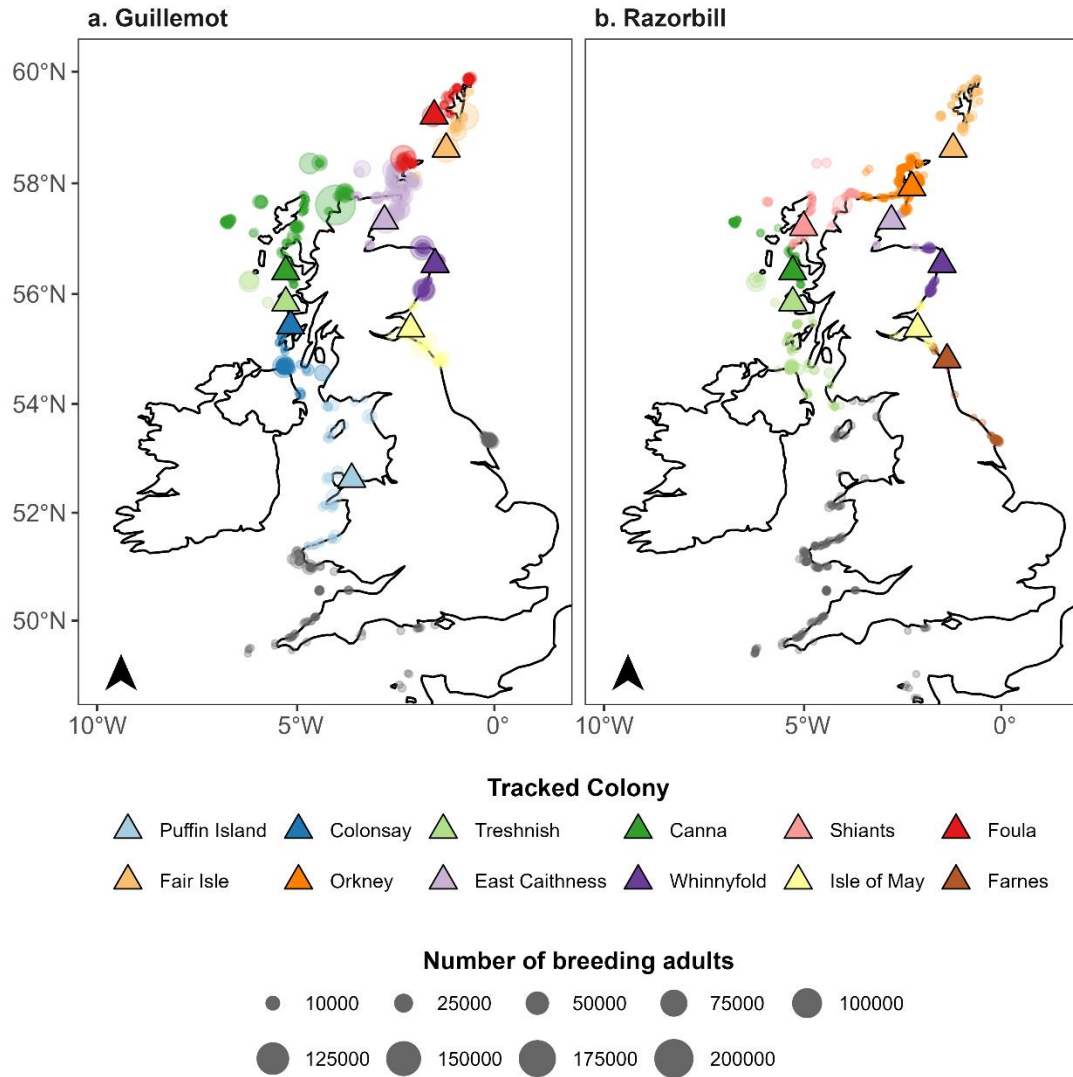


Figure 6. Tracked colonies (triangles) and Seabird 2000 sites (circles) for each species (common guillemots and razorbills). Seabird 2000 sites are coloured by the nearest tracked colony within 270km whilst grey circles indicate colonies that were further than 270km from the nearest tracked colony. Size of the Seabird 2000 site indicates the number of breeding adults at that colony, which we converted to number of breeding adults by multiplying by 0.67 (to estimate the number of pairs) and doubling (Mitchell *et al.* 2004).

Table 6. Number of geolocators deployed and retrieved for each species (G = common guillemot, R = razorbill), colony and year.

		Deployed					Retrieved					Retrieval rate (%)
		2017	2018	2019	2020	Total	2018	2019	2020	2021	Total	
Puffin Island	G	-	-	25	-	25	-	-	8	4	12	48.0
Colonsay	G	-	30	30	-	60	-	14	29	3	46	76.7
	R	-	9	-	-	9	-	1	-	-	1	11.1
Treshnish	G	-	20	30	-	50	-	12	19	4	35	70.0
	R	-	20	-	-	20	-	12	1	2	15	75.0
Canna	G	90	40	11	-	141	37	26	-	6	69	48.9
	R	20	22	-	-	42	4	16	-	3	23	54.8
Shiants	R	-	20	-	-	20	-	13	-	-	13	65.0
Foula	G	40	-	-	-	40	13	-	-	-	13	32.5
	R	10	-	-	-	10	1	-	-	-	1	10.0
Fair Isle	G	25	-	-	-	25	10	5	-	-	15	60.0
	R	21	-	-	-	21	9	2	-	-	11	52.4
Orkney	R	30	22	-	-	52	4	10	-	-	14	26.9
East Caithness	G	40	40	30	-	110	18	27	30	4	79	71.8
	R	30	30	-	-	60	13	7	2	2	24	40.0
Whinnyfold	G	40	40	30	-	110	24	27	25	9	85	77.3

	R	20	19	-	-	39	2	8	-	-	10	25.6
Isle of May	G	30	34	36	30	130	10	22	27	5	64	49.2
	R	30	30	-	-	60	11	11	-	-	22	36.7
Farnes	G	4	-	-	-	4	1	-	-	-	1	25.0
	R	6	-	-	-	6	3	1	-	-	4	66.7
Total	G	269	204	192	30	695	113	133	138	35	419	60.3
	R	167	172	0	0	339	47	81	3	7	138	40.7

Statistical methods for analysis of geolocator data

Deriving spatial locations from geolocator data

Two locations per day were derived from the geolocators using the R package 'probGLS' (Merkel *et al.* 2016) following the methods from Buckingham *et al.* (2022). ProbGLS uses an iterative algorithm to determine the most likely track based on light, salt water immersion and temperature data sampled by the geolocator. This method of combining several data types has been shown to reduce the amount of error associated with geolocation (Phillips *et al.* 2004; Halpin *et al.* 2021) and allows locations during the equinox to be estimated (approximately: autumn: 23rd September, spring: 20th March) and surrounding periods, unlike methods that solely rely on light data. The 'prob_algorithm' function was run using 100 iterations (Dunn *et al.* 2020; Buckingham *et al.* 2022).

The use of multiple iterations within probGLS is designed to capture the error associated with estimating spatial locations from the GLS data. For each GLS observation, at each iteration, the twilight time and solar angle are simulated from distributions that are designed to capture the uncertainty in these quantities: the resulting simulated twilight time and solar angle are then used to derive a simulated location associated with each observation within each iteration. Looking across all iterations gives a cloud of points, representing the locational uncertainty associated with that GLS observation. This approach is then applied across the whole track, to capture the locational uncertainty associated with each track.

Kernel density estimation

To estimate the density of birds from each colony within the tracked range, the utilisation distribution for birds from each tracked colony during each month of the non-breeding season (July – March) was calculated. We used kernel density estimation, implemented via the 'kernelUD' function in the R package 'adehabitatHR' (Calenge 2006), with bivariate normal kernels and ad-hoc smoothing and a grid cell size of 100km². Kernel density estimation is designed to estimate the spatial distribution from the data in a way that assumes that the distribution varies smoothly over space, but, otherwise, makes relatively few assumptions. This differs, for example, from the habitat association models used in the analysis of GPS data for lesser black-backed gulls, as the habitat association models attempt to link the spatial distribution to explanatory variables – the habitat association modelling approach has the potential to be more generalizable (i.e. it provides a basis for using the model to directly predict distributions at untracked sites, as in Wakefield *et al.*, 2017) but also relies on stronger assumptions. The rationale for using kernel density estimation, rather than habitat association modelling, to model the GLS data is that non-breeding season distributions of seabirds will typically be harder to relate to explanatory variables (because the central place foraging constraint means that distance to colony is typically strongly related to the spatial distribution in the breeding season), and that the high level of locational uncertainty in the GLS data would also make it difficult to identify relationships with explanatory variables. To determine whether sample sizes were sufficient to capture colony- and year-specific distributions for 2017-18 and 2018-19, Buckingham *et al.* (2022) conducted a bootstrapping procedure whereby they calculated 50% kernel density contours for each dataset, following the above methods, using randomly sampled selections of birds and allowing individual replacement, starting with a sample size of 1 bird and increasing until the total number of birds for that dataset had been reached. This step was repeated 1000 times. The point where the increase in the median kernel contour area used (km²) levelled off was then determined for each dataset. Any datasets where this point was not reached, indicating they did not have confidence that the sample size of tracked birds was sufficient to reliably estimate the core distribution of individuals

from that breeding colony for that period in the given year, were eliminated. This was the case for guillemots on Farnes, and razorbills on Farnes, Foula and Colonsay. The findings of Buckingham et al. (2022) were used here to inform the selection of colonies for kernel analysis based on samples sizes of all three study years combined, allowing razorbills on Farnes to be added to the suite, but not Farnes guillemots, Foula razorbills or Colonsay razorbills.

Accounting for uncertainty

The kernel density estimation approach could be applied only to the “most probable track” (which can be extracted from the results of “prob_algorithm”) but this would mean that the substantial locational uncertainty in the GLS observations was not accounted for. Since probGLS already simulates the uncertainty associated with each track, our approach is to propagate this uncertainty through into the kernel density estimation, and hence ensure that it is quantified within the outputs of the spatial mapping. We achieve this by applying ‘kernelUD’ separately to the simulated tracks within each of the 100 iterations of probGLS. This leads to the production of 100 simulated kernel density estimates for each tracked colony for each month of the non-breeding season. We then present uncertainty by summarizing across these iterations: for example, by calculating the 25% and 75% quantiles across these simulations for each point on each map, in order to show the uncertainty in a spatial way. To show the uncertainty in summary statistics (e.g. the proportion of the spatial distribution that lies within a particular polygon, such as a BDMPS region), we calculate this statistic separately for each iteration, and then, again, summarize by looking at quantiles across these simulations (e.g. 2.5% and 97.5% quantiles, which define a 95% confidence interval).

Removing inland locations

We removed inland distributions predicted by the utilisation distribution (Figure 7), as the majority of smoothed locations were located in the sea (95.0% of 318857 points), thus inland distributions were mostly a result of the kernel density smooth (and are extremely unlikely for our study species).

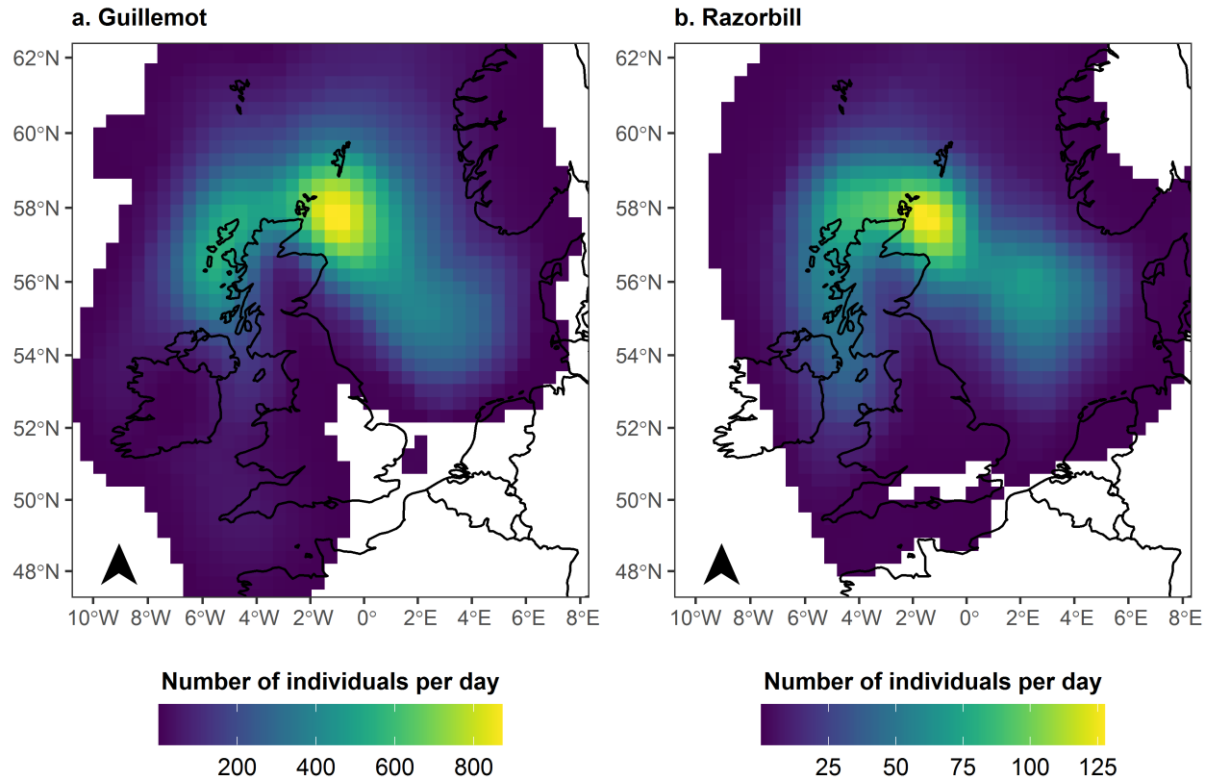


Figure 7. Distribution (mean number of birds present per day) using 75% kernel density contours of common guillemots and razorbills during the non-breeding season (July – March), without exclusion of land-based distributions. These maps are representative of adult birds that breed at the colonies within our range (i.e. tracked colonies and untracked colonies within 270km of a tracked colony; see Figure 6) and should not be used to infer distributions or energetic requirements of colonies outside of this range or other age classes.

Statistical methods for evaluating maps derived from geolocator data

Visual summaries

We produce maps to visualize the GLS utilisation distributions (UDs), and the uncertainty associated with these. The bootstrapping approach (of running kernel density estimation on the simulated tracks produced by each iteration of probGLS) yields 100 realisations of the UD for each species/colony/month combination, so that uncertainty in location could be characterised. Nine tracked colonies for guillemot (Farnes was excluded as only 1 tag was retrieved) and nine tracked colonies for razorbill (Colonsay and Foula were excluded as 1 tag was retrieved at each colony) were included in the comparison analysis (Table 6). All analyses were carried out using R v4.2.1 (R Core Team, 2022) and RStudio release 7872775e, 2022-07-22. Seasonal maps are calculated and summarized, with the seasons defined as in BDMPS (to ensure consistency with the seasonal definitions used in the current apportioning tool). For guillemot, the non-breeding season was defined from August to February. Two non-breeding seasons were defined for razorbill of migration from August to October and January to March, and winter in November and December. UD's (including all bootstrap realisations) were aggregated by species, colony, and season and summary statistics of mean, upper and lower 95%CI's were calculated by cell using packages 'rgdal' 1.5.32 (Bivand et al. 2022), 'raster' 3.6.3 (Hijmans, 2022a), and 'terra' 1.6.17 (Hijmans, 2022b), and 'stringr' 1.4.0 (Wickham, 2019). Maps of mean, lower and upper 95%CI were produced for each population in each season, and give a visual representation of the level of uncertainty relative to the magnitude of the mean.

Comparison against BDMPS tables

BDMPS is the current main non-breeding season apportioning tool, so it is important to compare the spatial distributions that have been derived from GLS data against those used in BDMPS. BDMPS takes a regional approach to characterising the spatial distribution, by specifying the proportion of birds from each population (individual UK SPAs, together with large-scale aggregations of UK non-SPA populations and non-UK populations) that are in each region in each season. BDMPS areas were overlaid onto each realisation of the GLS-based maps to apportion estimates in the regions UK North Sea and Channel waters, UK Western waters, and Outside of the UK. Proportions were calculated for each area. Summary tables were collated by corresponding species, SPAs, and season from the BDMPS tables with estimated mean and lower and upper 95%CI's proportions. The proportion of breeding adults visiting each region was compared between the GLS-based maps and BDMPS.

Visualising density using space time cubes

Visual tools can also be used as a diagnostic tool to explore the assumptions that underpin the GLS-based maps. Maps are usually presented as static visualisations, making it difficult to determine whether there are temporal shifts in space use, particularly at a fine-scale. Understanding and determining temporal variation in distribution is important when considering apportioning criteria because temporal variation can be used to characterise uncertainty when apportioning (areas) and also because areas of high importance (e.g. foraging grounds) which are visited periodically, may not be shown as areas of high usage when mean values are taken over snapshots of time. We developed space time cubes to characterise and visualise fine-scale temporal shifts in space use, and thereby to investigate the plausibility of the assumption, which underpins the GLS-based maps, that temporal change – aside from systematic differences between months/seasons – can safely be ignored.

Space time cubes showing how the estimated distribution of the birds changed over time were produced for guillemots from two select colonies, Isle of May and Canna, using the methodology of Demšar et al (2015) and R code adapted from Demšar (2015). This approach calculates space time densities from a set of movement trajectories, which were estimated from the geolocator data using the probGLS R package (Merkel 2018). The most probable track was extracted for each bird from each colony, with the 0.1% most extreme locations in each of the four cardinal directions filtered from the dataset to remove some extreme outliers stemming from the substantial uncertainty in the geolocator positions. The tracks were split by month to assist with ease of visualisation. The voxel size (the spatial resolution at which the cubes were displayed) was chosen to be 10km to balance the ability to detect relatively fine scale changes in spatial utilisation, particularly around the coasts, against the high computational cost of estimating densities over a large area. A Gaussian kernel with a width of 75km was chosen through trial and error as this was observed to give what appeared to be an appropriate level of smoothing of the densities. More sophisticated methods of choosing this kernel size were considered; however, the currently developed methodology only allows a one-dimensional kernel and so this method appeared to give the most robust results across colonies. The calculated densities were then corrected to remove any density over land and the densities were reweighted accordingly. The final estimated densities were then plotted in ParaView (Ahrens et al 2005) for each combination of month and colony and the images were stitched together into gif format. Space time cubes are 3D objects that can be manipulated within the software. We chose visualisations of space in two dimensions (x and y) and day of the month in the z-axis so that a 3D utilisation distribution cloud represented fine-scale (daily) space usage.

Evaluating potential to produce distributions for untracked colonies

In order to calculate apportioning percentages, it will be necessary to derive the spatial distributions associated with all colonies, not only those for which GLS data are available. To evaluate the potential to extrapolate the GLS-based spatial distributions from tracked colonies to untracked colonies, we extracted the locations and colony counts for each UK breeding colony that were included in the Seabird 2000 census (Mitchell *et al.* 2004). We calculated the distance by sea between each tracked colony and Seabird 2000 site and assigned each site to its nearest tracked colony, since it is likely that colonies that are located closer together have more similar distributions and behaviour (Fauchald *et al.* 2021; Buckingham *et al.* 2022). We visually explored the distribution of distances between tracked colonies and Seabird 2000 sites.

Results

Visual summaries

From the UD's available, nine colonies of guillemot and nine colonies of razorbill were processed. Paucity in underlying data prevented simulating realisations at the remaining colonies. Colony names were defined from the locations where geolocator tags were attached to individuals in colonies (Table 6).

Figs. 8-16 show the estimated population density of adult breeding guillemot with 95% CIs at each tracked colonies in the non-breeding season (August-February). Figs. 17-25 show the estimated population density of adult breeding razorbill with 95% CIs at each tracked colonies in the non-breeding migration seasons (August-October and January-March). Figs. 18-34 show the estimated population density of adult breeding razorbill with 95% CIs at each tracked colonies in the non-breeding winter season

(November-December). The maps were cropped for visualisation but the full utilisation densities were used to calculate the summary statistics for Tables 7, 8, and 9.

Guillemot at the Canna colony in the non breeding season

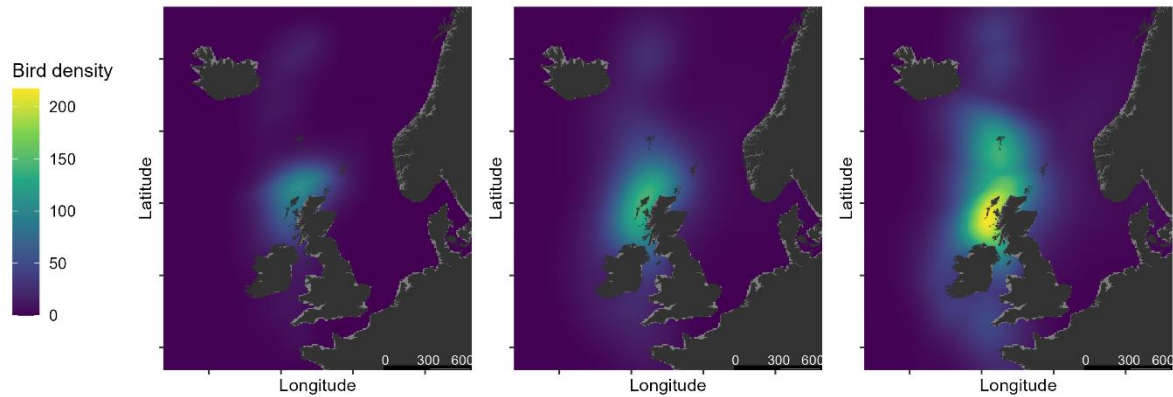


Figure 8. Estimated population density (L) lower 95% CI, (C) mean, (R) upper 95% CI of adult breeding guillemot in the Canna colony during the non-breeding season (Aug-Feb)

Guillemot at the Colonsay colony in the non breeding season

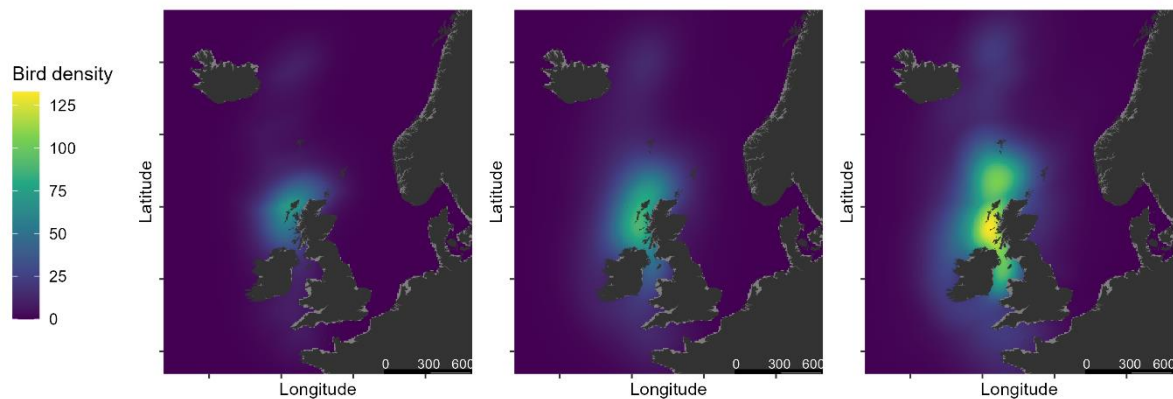


Figure 9. Estimated population density (L) lower 95% CI, (C) mean, (R) upper 95% CI of adult breeding guillemot in the Colonsay colony during the non-breeding season (Aug-Feb)

Guillemot at the East Caithness colony in the non breeding season

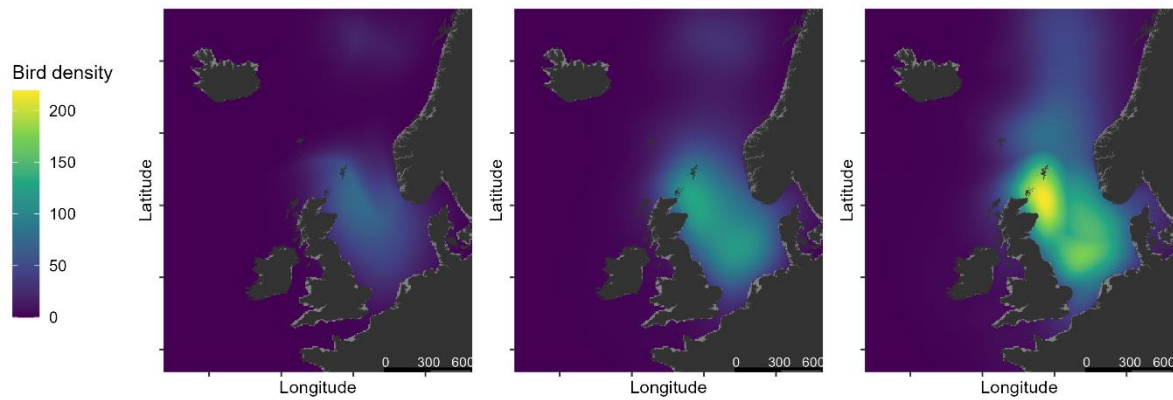


Figure 10. Estimated population density (L) lower 95% CI, (C) mean, (R) upper 95% CI of adult breeding guillemot in the East Caithness colony during the non-breeding season (Aug-Feb)

Guillemot at the Fair Isle colony in the non breeding season

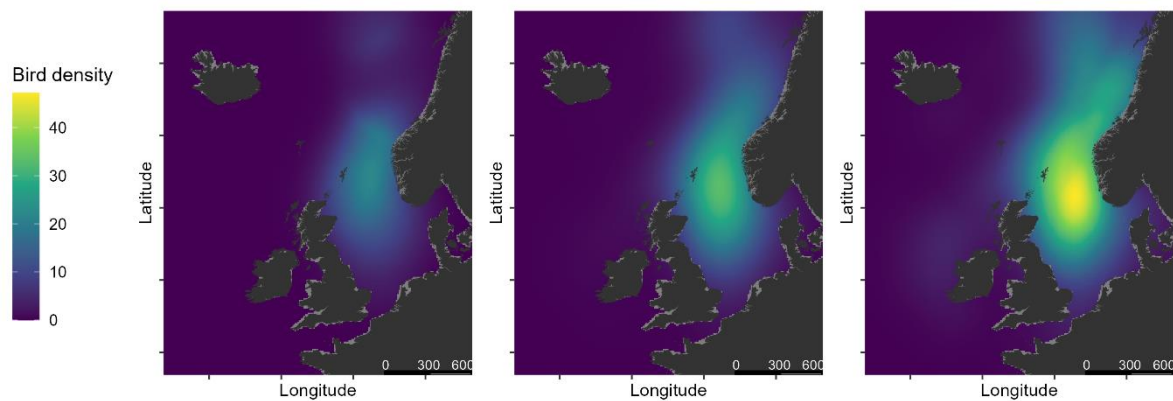


Figure 11. Estimated population density (L) lower 95% CI, (C) mean, (R) upper 95% CI of adult breeding guillemot in the Fair Isle colony during the non-breeding season (Aug-Feb)

Guillemot at the Foula colony in the non breeding season

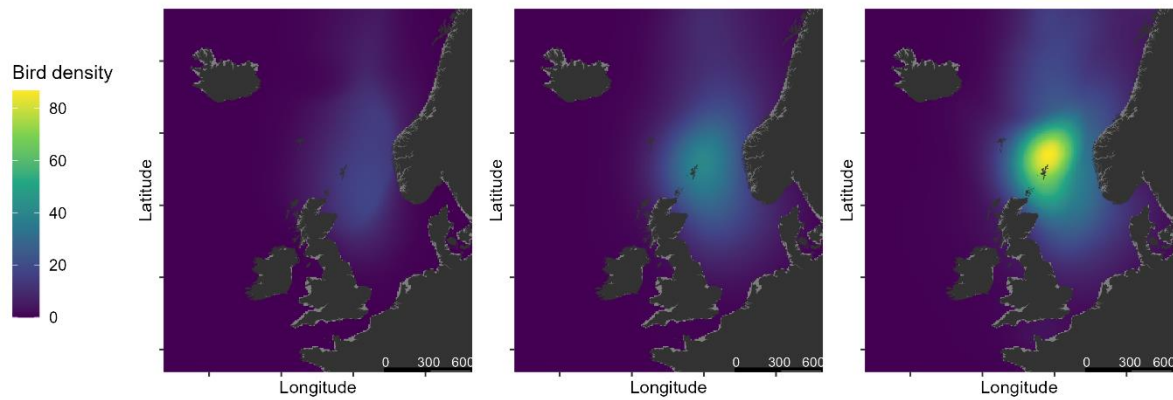


Figure 12. Estimated population density (L) lower 95% CI, (C) mean, (R) upper 95% CI of adult breeding guillemot in the Foula colony during the non-breeding season (Aug-Feb)

Guillemot at the Isle of May colony in the non breeding season

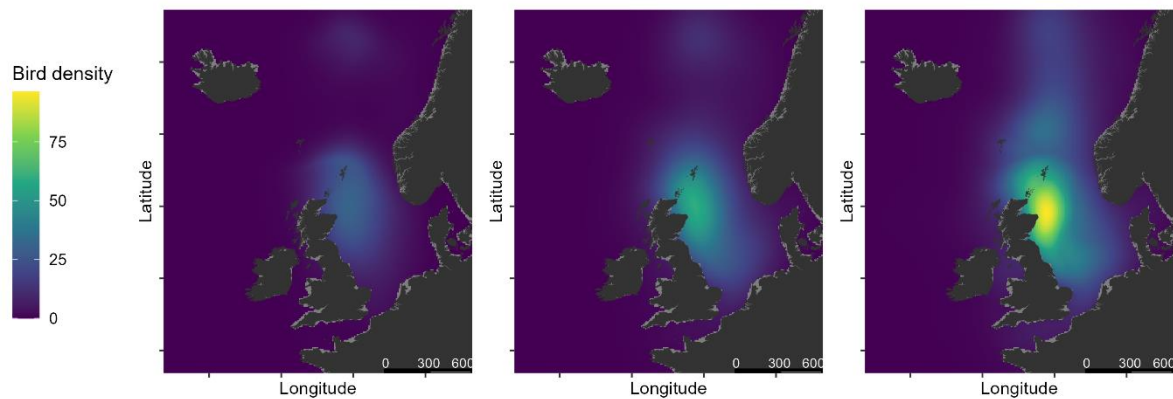


Figure 13. Estimated population density (L) lower 95% CI, (C) mean, (R) upper 95% CI of adult breeding guillemot in the Isle of May colony during the non-breeding season (Aug-Feb)

Guillemot at the Puffin Island colony in the non breeding season

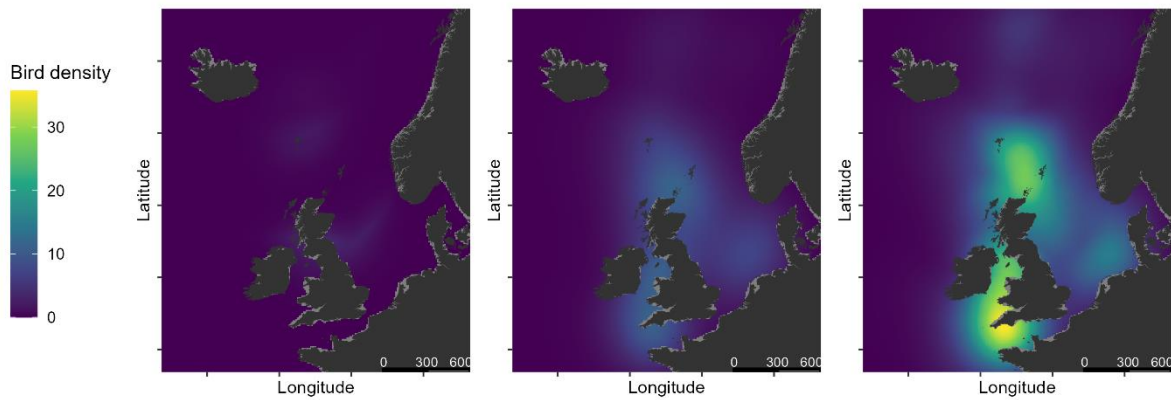


Figure 14. Estimated population density (L) lower 95% CI, (C) mean, (R) upper 95% CI of adult breeding guillemot in the Puffin Island colony during the non-breeding season (Aug-Feb)

Guillemot at the Treshnish colony in the non breeding season

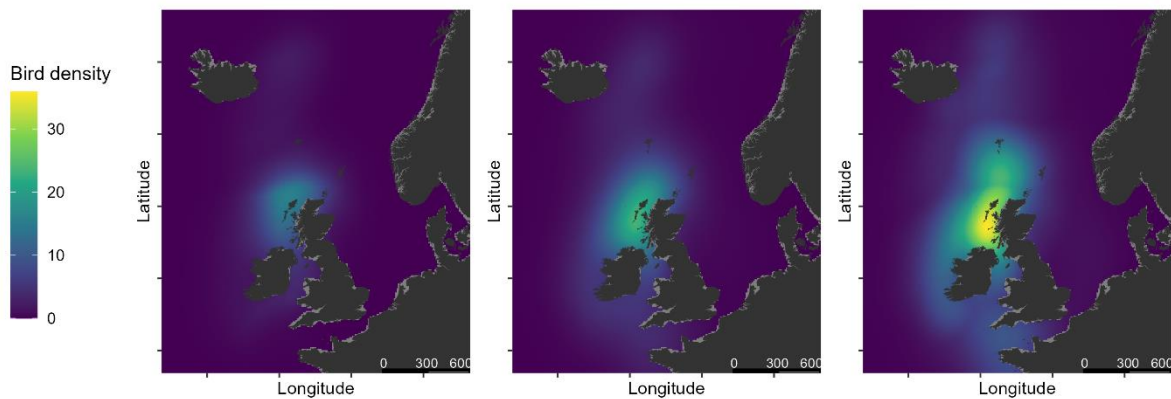


Figure 15. Estimated population density (L) lower 95% CI, (C) mean, (R) upper 95% CI of adult breeding guillemot in the Treshnish colony during the non-breeding season (Aug-Feb)

Guillemot at the Whinnyfold colony in the non breeding season

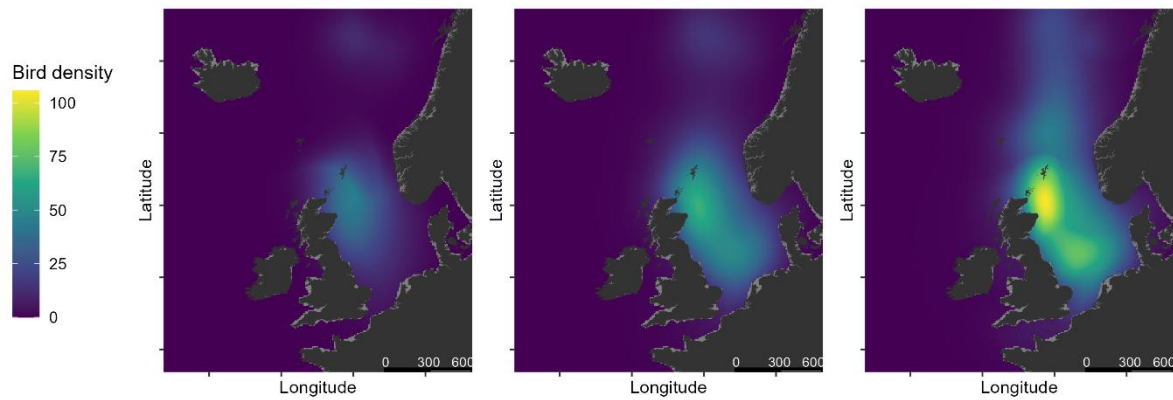


Figure 16. Estimated population density (L) lower 95% CI, (C) mean, (R) upper 95% CI of adult breeding guillemot in the Whinnyfold colony during the non-breeding season (Aug-Feb)

Razorbill at the Canna colony in the migration seasons

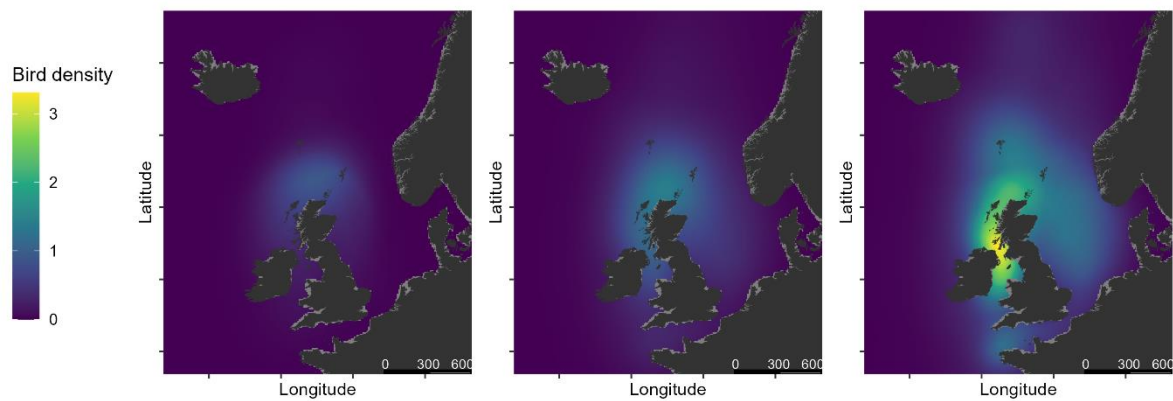


Figure 17. Estimated population density (L) lower 95% CI, (C) mean, (R) upper 95% CI of adult breeding razorbill in the Canna colony during the non-breeding migration seasons (Aug-Oct; Jan-Mar)

Razorbill at the East Caithness colony in the migration seasons

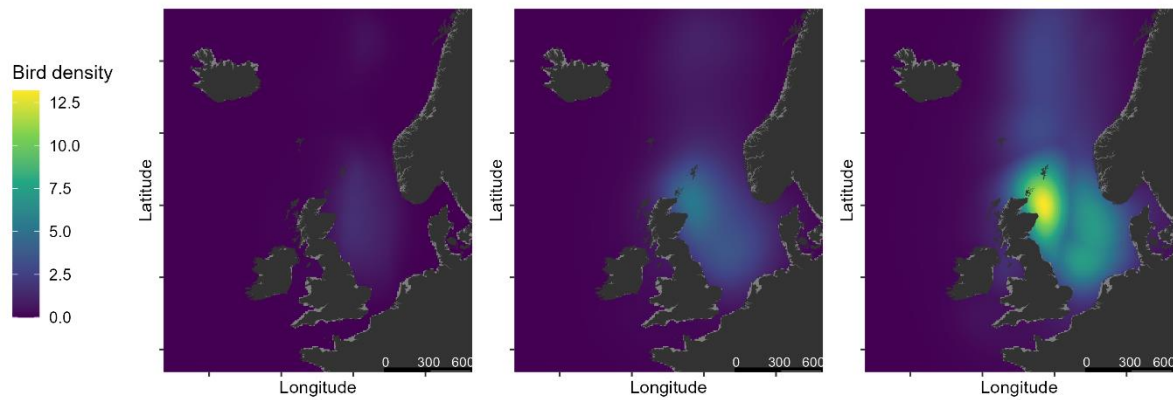


Figure 18. Estimated population density (L) lower 95% CI, (C) mean, (R) upper 95% CI of adult breeding razorbill in the East Caithness colony during the non-breeding migration seasons (Aug-Oct; Jan-Mar)

Razorbill at the Fair Isle colony in the migration seasons

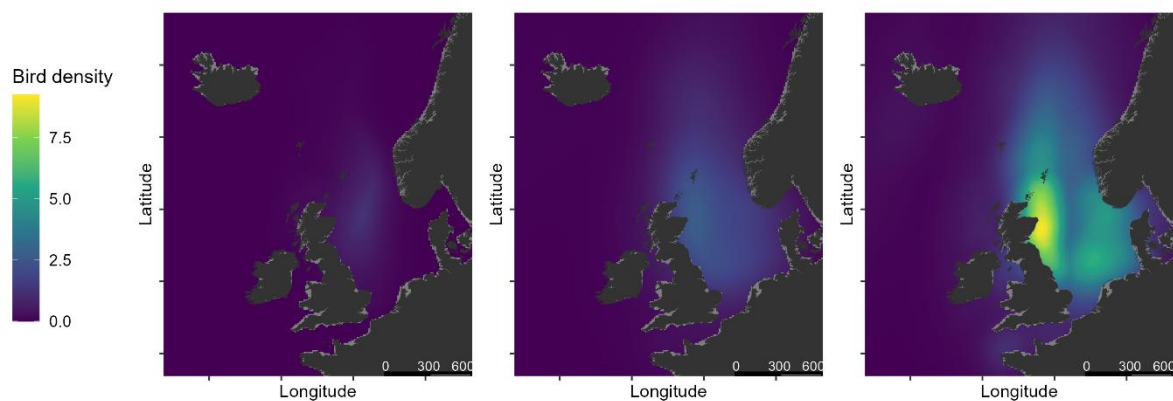


Figure 19. Estimated population density (L) lower 95% CI, (C) mean, (R) upper 95% CI of adult breeding razorbill in the Fair Isle colony during the non-breeding migration seasons (Aug-Oct; Jan-Mar)

Razorbill at the Farnes colony in the migration seasons

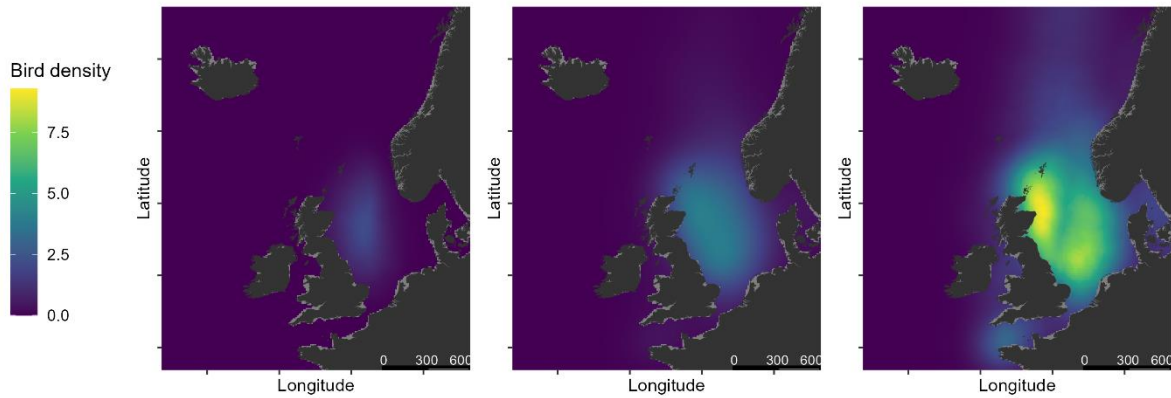


Figure 20. Estimated population density (L) lower 95% CI, (C) mean, (R) upper 95% CI of adult breeding razorbill in the Farnes colony during the non-breeding migration seasons (Aug-Oct; Jan-Mar)

Razorbill at the Isle of May colony in the migration seasons

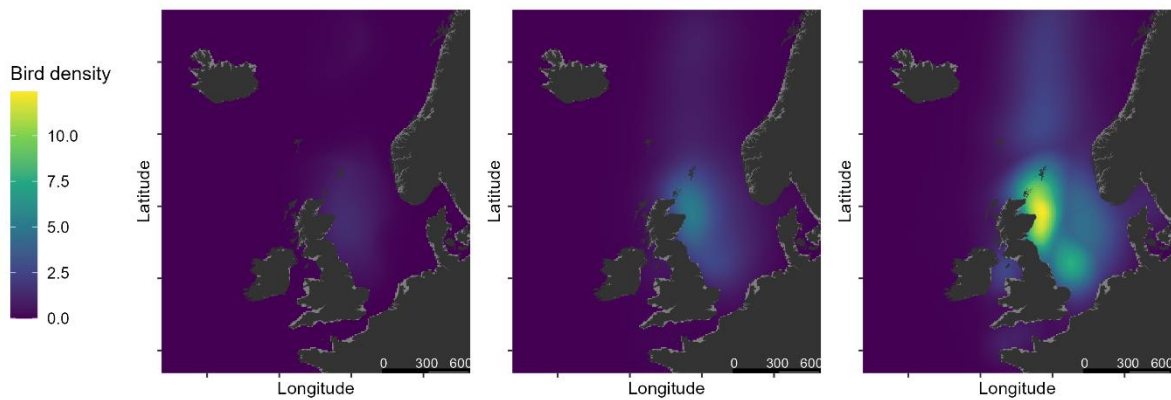


Figure 21. Estimated population density (L) lower 95% CI, (C) mean, (R) upper 95% CI of adult breeding razorbill in the Isle of May colony during the non-breeding migration seasons (Aug-Oct; Jan-Mar)

Razorbill at the Orkney colony in the migration seasons

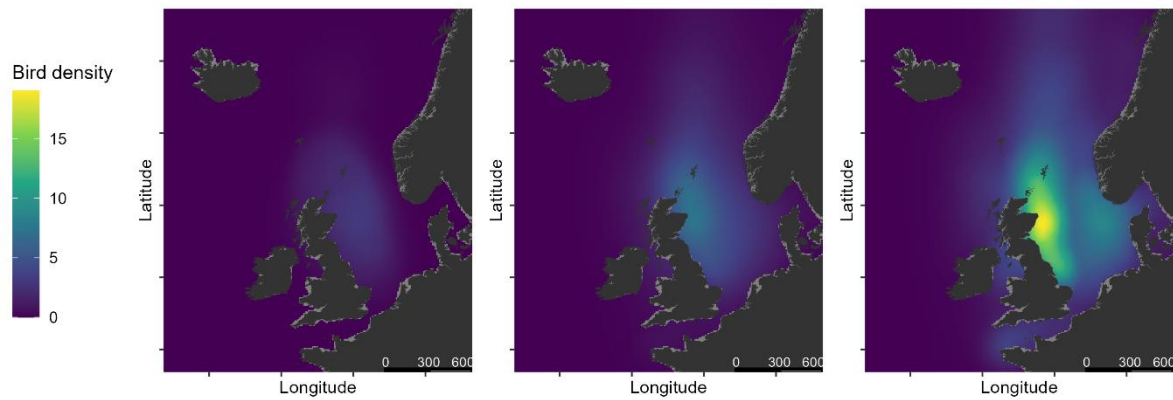


Figure 22. Estimated population density (L) lower 95% CI, (C) mean, (R) upper 95% CI of adult breeding razorbill in the Orkney colony during the non-breeding migration seasons (Aug-Oct; Jan-Mar)

Razorbill at the Shiant's colony in the migration seasons

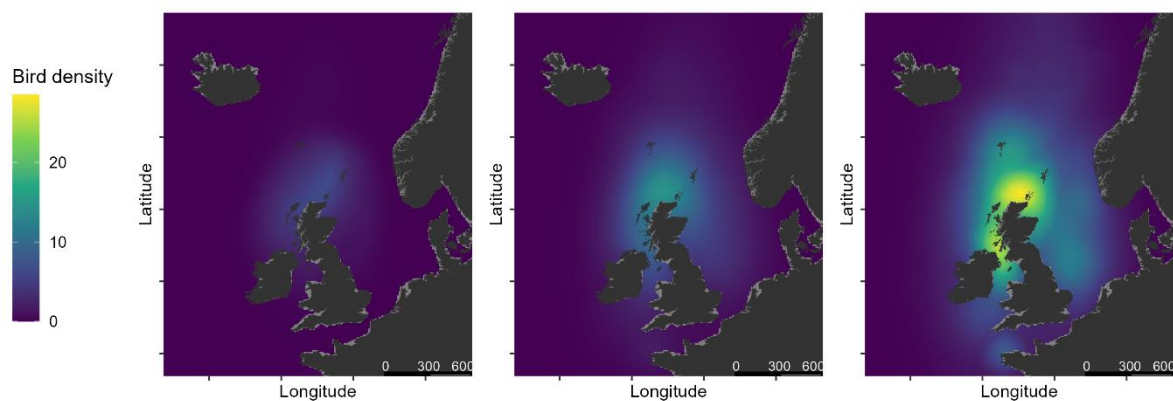


Figure 23. Estimated population density (L) lower 95% CI, (C) mean, (R) upper 95% CI of adult breeding razorbill in the Shiant's colony during the non-breeding migration seasons (Aug-Oct; Jan-Mar)

Razorbill at the Treshnish colony in the migration seasons

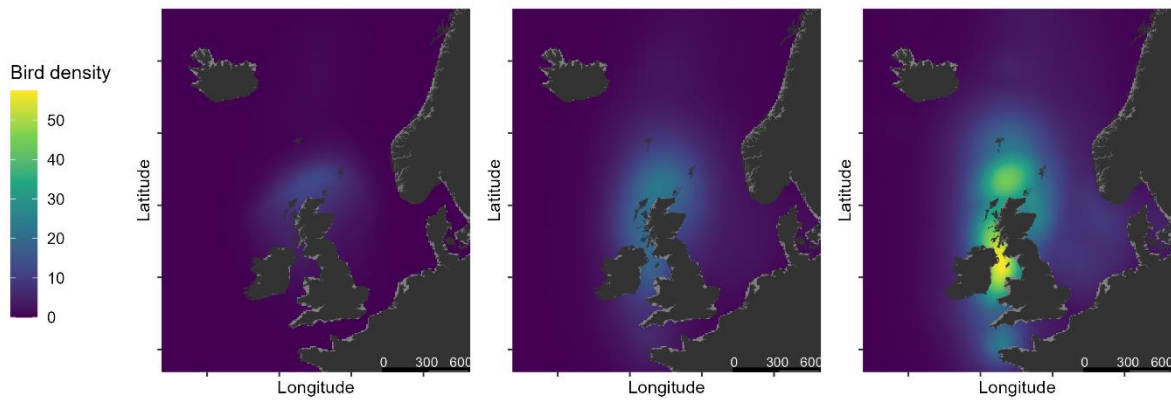


Figure 24. Estimated population density (L) lower 95% CI, (C) mean, (R) upper 95% CI of adult breeding razorbill in the Treshnish colony during the non-breeding migration seasons (Aug-Oct; Jan-Mar)

Razorbill at the Whinnyfold colony in the migration seasons

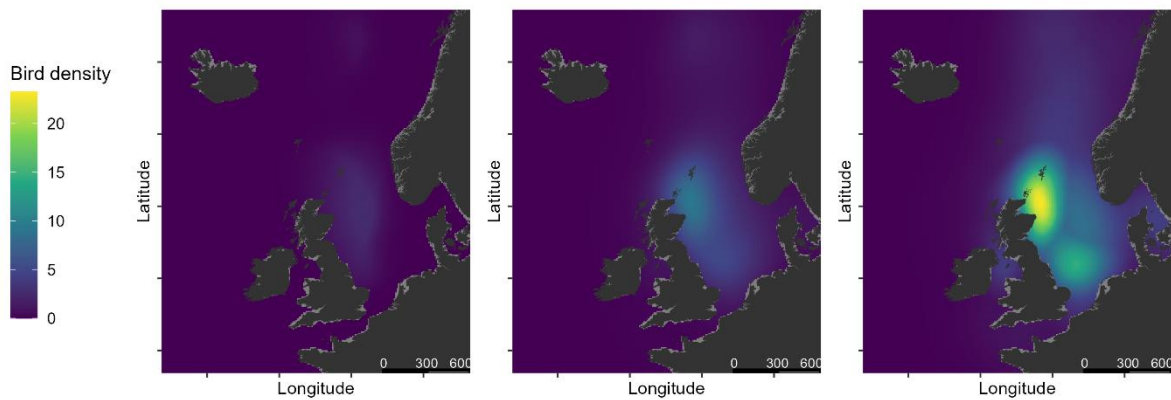


Figure 25. Estimated population density (L) lower 95% CI, (C) mean, (R) upper 95% CI of adult breeding razorbill in the Whinnyfold colony during the non-breeding migration seasons (Aug-Oct; Jan-Mar)

Razorbill at the Canna colony in the winter

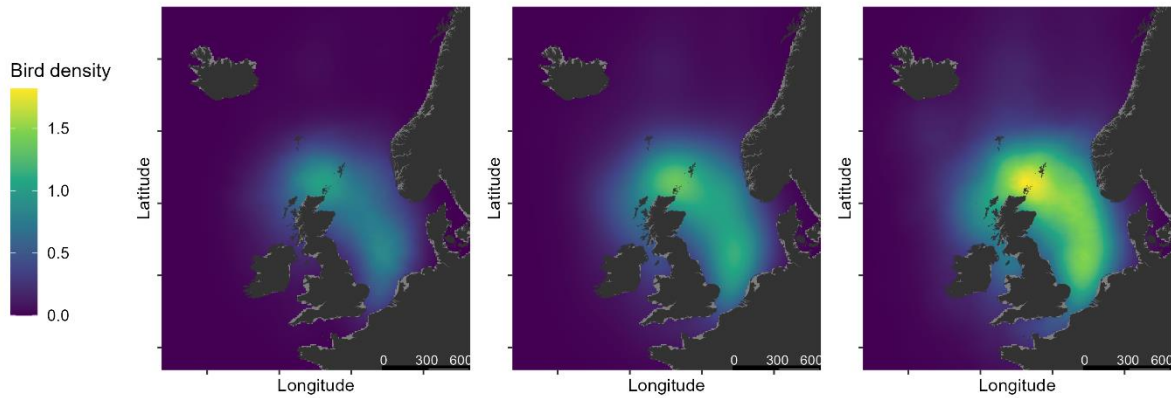


Figure 26. Estimated population density (L) lower 95% CI, (C) mean, (R) upper 95% CI of adult breeding razorbill in the Canna colony during the non-breeding winter seasons (Nov-Dec)

Razorbill at the East Caithness colony in the winter

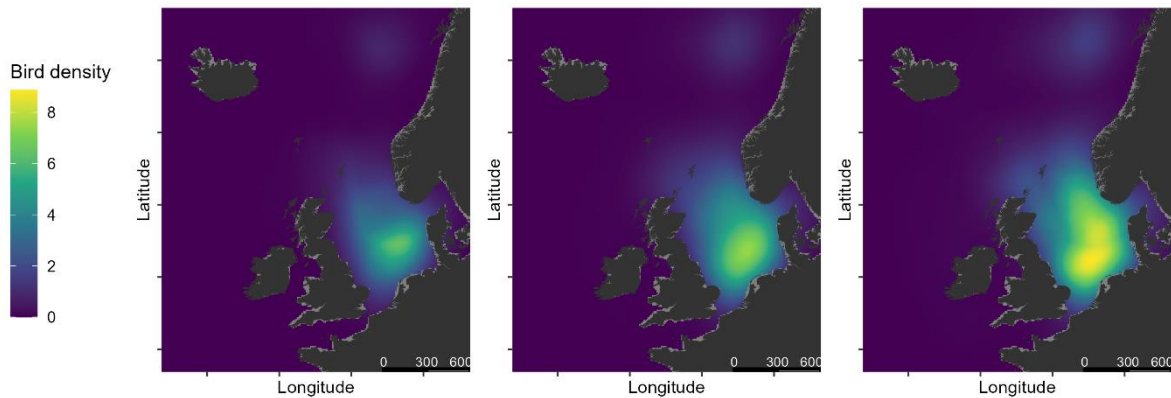


Figure 27. Estimated population density (L) lower 95% CI, (C) mean, (R) upper 95% CI of adult breeding razorbill in the East Caithness colony during the non-breeding winter seasons (Nov-Dec)

Razorbill at the Fair Isle colony in the winter

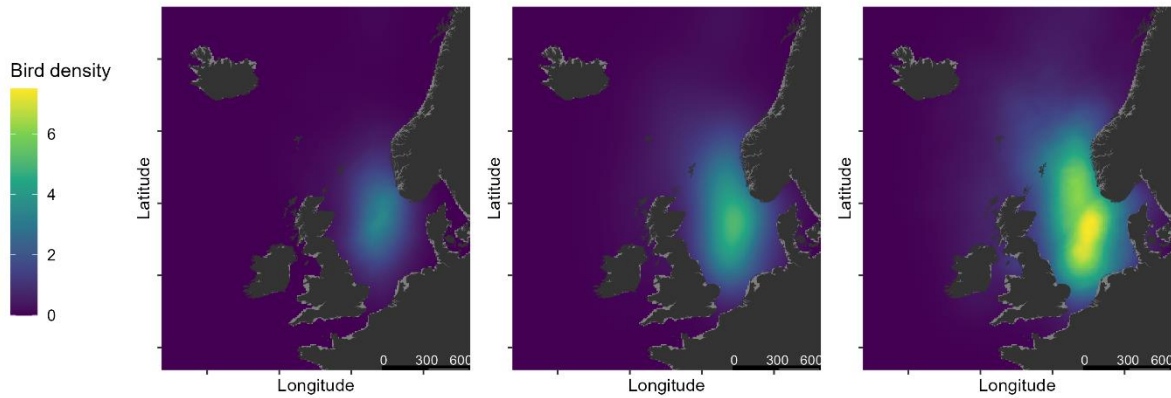


Figure 28. Estimated population density (L) lower 95% CI, (C) mean, (R) upper 95% CI of adult breeding razorbill in the Fair Isle colony during the non-breeding winter seasons (Nov-Dec)

Razorbill at the Farnes colony in the winter

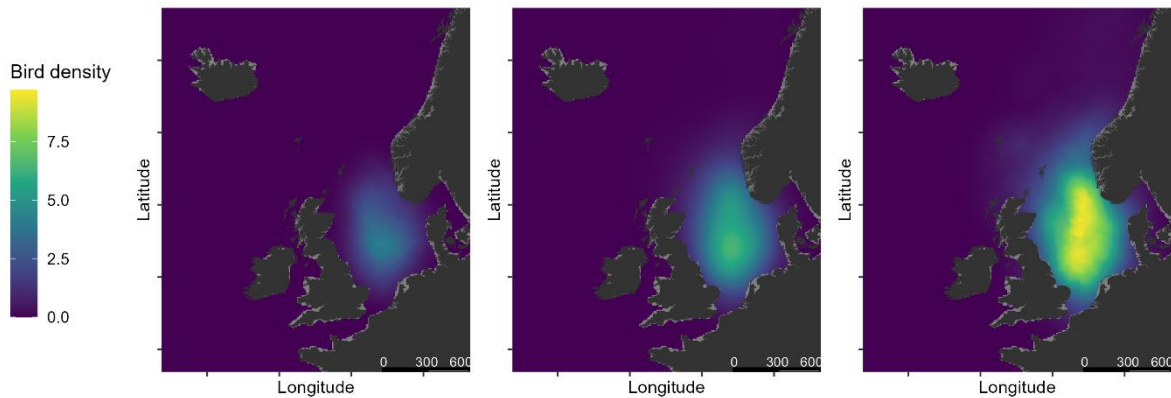


Figure 29. Estimated population density (L) lower 95% CI, (C) mean, (R) upper 95% CI of adult breeding razorbill in the Farnes colony during the non-breeding winter seasons (Nov-Dec)

Razorbill at the Isle of May colony in the winter

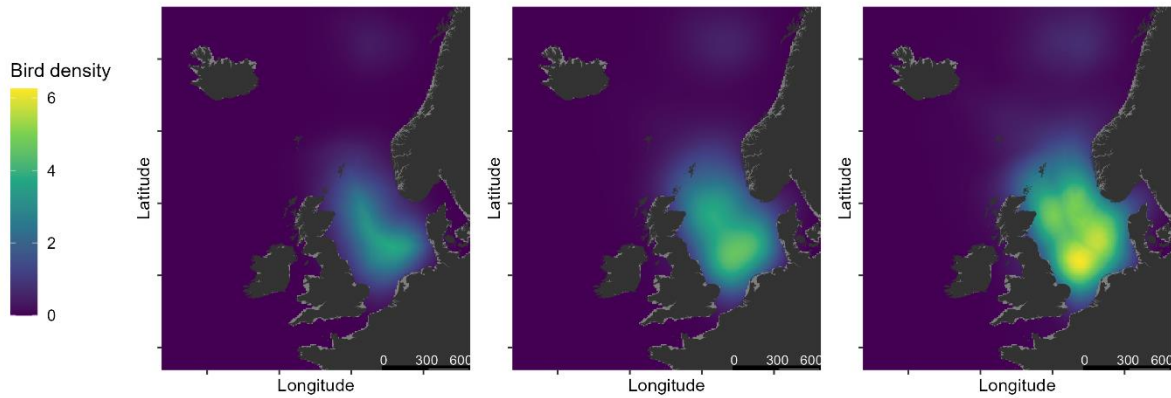


Figure 30. Estimated population density (L) lower 95% CI, (C) mean, (R) upper 95% CI of adult breeding razorbill in the Isle of May colony during the non-breeding winter seasons (Nov-Dec)

Razorbill at the Orkney colony in the winter

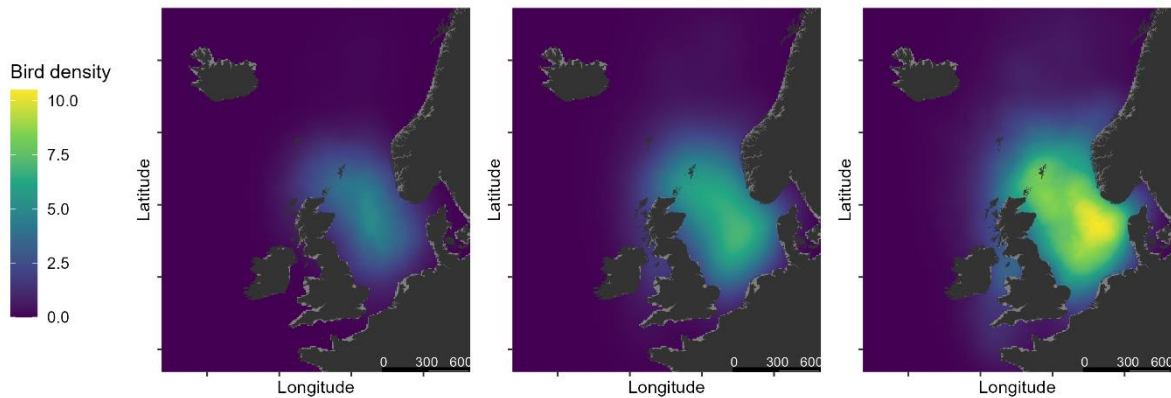


Figure 31. Estimated population density (L) lower 95% CI, (C) mean, (R) upper 95% CI of adult breeding razorbill in the Orkney colony during the non-breeding winter seasons (Nov-Dec)

Razorbill at the Shiantis colony in the winter

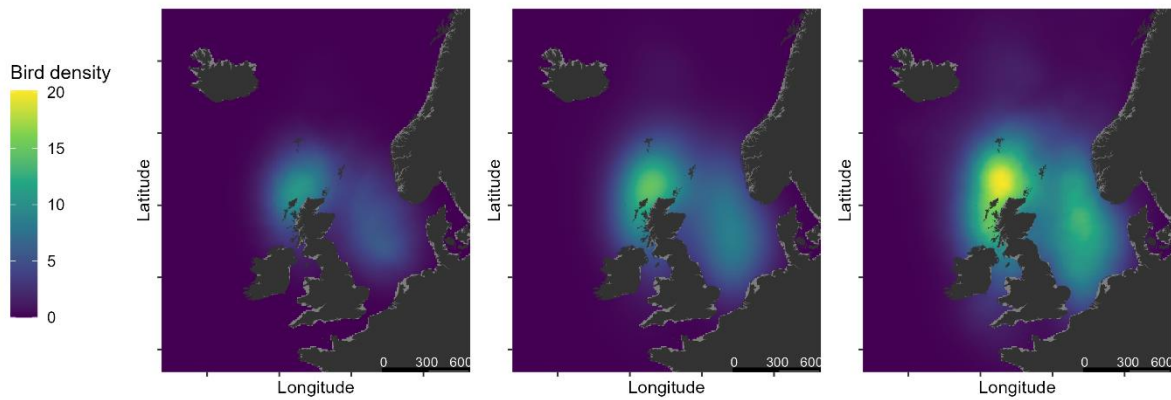


Figure 32. Estimated population density (L) lower 95% CI, (C) mean, (R) upper 95% CI of adult breeding razorbill in the Shiantis colony during the non-breeding winter seasons (Nov-Dec)

Razorbill at the Treshnish colony in the winter

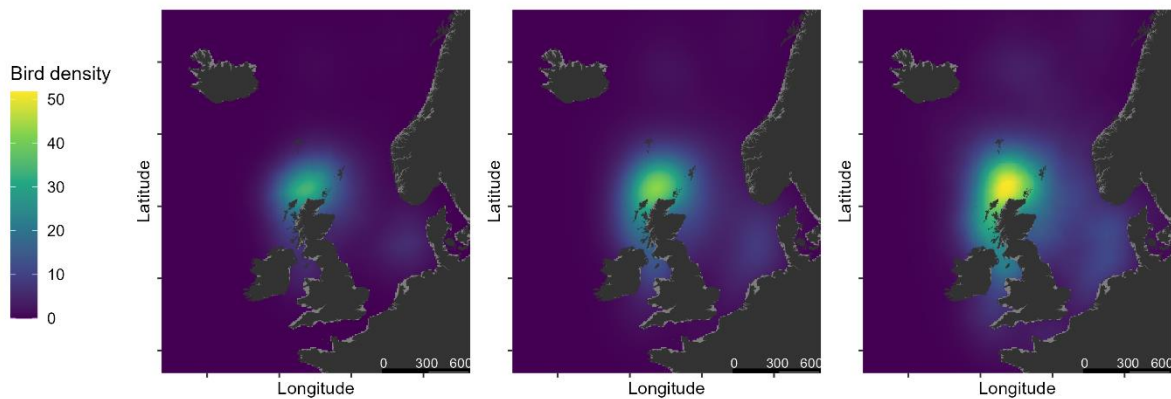


Figure 33. Estimated population density (L) lower 95% CI, (C) mean, (R) upper 95% CI of adult breeding razorbill in the Treshnish colony during the non-breeding winter seasons (Nov-Dec)

Razorbill at the Whinnyfold colony in the winter

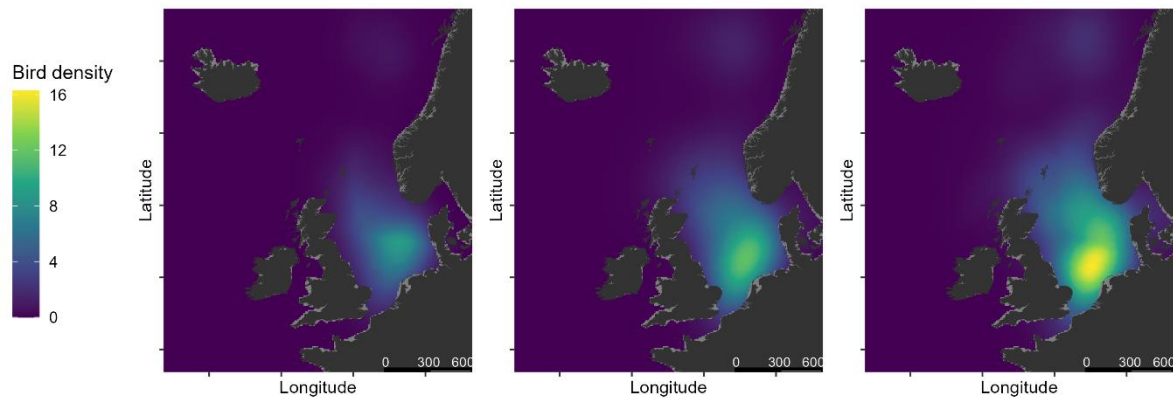


Figure 34. Estimated population density (L) Lower 95% CI, (C) Mean, (R) Upper 95% CI of adult breeding razorbill in the Whinnyfold colony during the non-breeding winter seasons (Nov-Dec)

Comparison against BDMPS tables

Table 7 shows the estimated proportion of adult breeding guillemot in the non-breeding season (August-February) that visit each region in each season. The estimated proportion of adult breeding razorbill are shown in migration seasons (August-October and January-March; Table 8) and winter season (Nov-Dec; Table 9) that visit each region in each season, based on the geolocator-based maps and on BDMPS tables. The estimated proportion of adult breeding birds are presented with the lower and upper 95% CIs. Region is defined as the two UK BDMPS areas (UK North Sea and Channel waters, and UK western waters) plus a third category including all estimated density outside of these areas (Outside of UK regions). The estimated mean proportions were aggregated over months (Aug-Feb for guillemot; Aug-Oct, Jan-Mar for razorbill in migrations; and Nov-Dec for razorbill in winter) and locational uncertainty in the GLS tags was captured in the lower and upper 95% CIs. The results from BDMPS calculations are shown as the BDMPS population (SPA or non-SPA) that incorporates the colony.

Concordance between the estimated proportion of adult breeding birds and the BDMPS proportion of adults, defined as the BDMPS proportion falling on or between the lower and upper 95%CI of the mean estimated proportion, is denoted by a green tick mark (✓) in the last column on the right in each Table (7, 8, 9). The results for guillemots (Table 7) and razorbill in the migration seasons (Table 8) suggest large differences between the GLS-based results and BDMPS results for almost all colonies, for both species, even with GLS locational uncertainty accounted for in the apportioning estimates. Table 9 suggests a higher level of concordance, with 14 from 27 of the BDMPS calculations falling on or between the lower and upper 95% CIs of the estimated mean proportion of razorbill adults in the winter season.

Table 7. Estimated proportion of adult breeding guillemot in the non-breeding season (all months Aug-Feb) by tracked colony BDMPS area estimated as mean proportion with 95%CI, calculated from the GLS UD, along with equivalent BDMPS values for the most relevant corresponding population in BDMPS. *Designated SPA but not for guillemot features, hence include with non-SPA populations in BDMPS tables for this species.  denotes instances where the BDMPS proportion of adults falls within the estimated 95%CI bounds of the proportion of adults from a colony.

GLS utilisation distributions			BDMPS calculations	
Colony	Region	Mean proportion of adults from this colony in each BDMPS region (95% CI low,upp)	BDMPS population relevant to colony	Proportion of adults from this population in each BDMPS region
Canna	Outside of UK regions	0.51 (0.45, 0.65)	Canna and Sanday SPA	0
	UK North Sea and Channel waters	0.11 (0.07, 0.16)		0.05
	UK western waters	0.38 (0.26, 0.47)		0.95
Colonsay	Outside of UK regions	0.48 (0.43, 0.51)	North Colonsay and western cliffs SPA	0
	UK North Sea and Channel waters	0.08 (0.05, 0.10)		0
	UK western waters	0.44 (0.41, 0.49)		1.00
East Caithness	Outside of UK regions	0.55 (0.41, 0.71)	East Caithness Cliffs SPA	0.30
	UK North Sea and Channel waters	0.40 (0.26, 0.53)		0.70
	UK western waters	0.04 (0.02, 0.07)		0
Fair Isle	Outside of UK regions	0.70 (0.60, 0.79)	Fair Isle SPA	0.28
	UK North Sea and Channel waters	0.28 (0.19, 0.36)		0.70
	UK western waters	0.02 (0.01, 0.04)		0.02 
Foula	Outside of UK regions	0.70 (0.54, 0.82)	Foula SPA	0.28
	UK North Sea and Channel waters	0.27 (0.15, 0.41)		0.70
	UK western waters	0.03 (0.02, 0.05)		0.02 
Isle of May	Outside of UK regions	0.45 (0.32, 0.65)	Forth Islands SPA	0.10
	UK North Sea and Channel waters	0.49 (0.32, 0.62)		0.90
	UK western waters	0.06 (0.03, 0.09)		0
Puffin Island	Outside of UK regions	0.53 (0.43, 0.61)		0.02

	UK North Sea and Channel waters	0.25 (0.09, 0.41)	UK West Coast non-SPA populations*	0.03
	UK western waters	0.22 (0.13, 0.33)		0.95
Treshnish	Outside of UK regions	0.54 (0.48, 0.57)	UK West Coast non-SPA populations*	0.02
	UK North Sea and Channel waters	0.08 (0.04, 0.11)		0.03
	UK western waters	0.39 (0.33, 0.46)		0.95
Whinnyfold	Outside of UK regions	0.51 (0.39, 0.65)	Buchan Ness to Collieston Coast	0.20
	UK North Sea and Channel waters	0.46 (0.33, 0.57)		0.80
	UK western waters	0.03 (0.02, 0.05)		0

Table 8. Estimated proportion of adult breeding razorbill in the non-breeding migration seasons (Aug-Oct; Jan-Mar) by tracked colony BDMPS area estimated as mean proportion with 95%CI, calculated from the GLS UDs, along with equivalent BDMPS values for the most relevant corresponding population in BDMPS. *Designated SPA but not for razorbill features, hence included with non-SPA populations in BDMPS tables for this species.  denotes instances where the BDMPS proportion of adults falls within the estimated 95%CI bounds of the proportion of adults from a colony.

GLS utilisation distributions			BDMPS calculations	
Colony	Region	Mean proportion of adults from this colony in each BDMPS region (95% CI low,upp)	BDMPS population relevant to colony	Proportion of adults from this population in each BDMPS region
Canna	Outside of UK regions	0.49 (0.34, 0.66)	UK Western non-SPA colonies*	0
	UK North Sea and Channel waters	0.26 (0.17, 0.37)		0.02
	UK western waters	0.25 (0.16, 0.43)		0.98
East Caithness	Outside of UK regions	0.53 (0.21, 0.77)	East Caithness Cliffs SPA	0
	UK North Sea and Channel waters	0.42 (0.20, 0.71)		1.00
	UK western waters	0.06 (0.02, 0.10)		0
Fair Isle	Outside of UK regions	0.60 (0.22, 0.76)	Fair Isle SPA	0
	UK North Sea and Channel waters	0.35 (0.20, 0.71)		0.95
	UK western waters	0.05 (0.02, 0.09)		0.05 
Farnes	Outside of UK regions	0.48 (0.17, 0.72)	UK North Sea non-SPA colonies*	0
	UK North Sea and Channel waters	0.48 (0.27, 0.75)		1.00
	UK western waters	0.04 (0, 0.14)		0.00
Isle of May	Outside of UK regions	0.45 (0.15, 0.76)	Forth Islands SPA	0
	UK North Sea and Channel waters	0.50 (0.22, 0.78)		1.00
	UK western waters	0.05 (0.01, 0.11)		0
Orkney	Outside of UK regions	0.51 (0.20, 0.72)	UK North Sea non-SPA colonies*	0
	UK North Sea and Channel waters	0.41 (0.23, 0.70)		1.00
	UK western waters	0.08 (0.04, 0.11)		0
Shiants	Outside of UK regions	0.45 (0.27, 0.63)		0

	UK North Sea and Channel waters	0.29 (0.15, 0.42)	Shiant Isles SPA	0.02
	UK western waters	0.26 (0.12, 0.40)		0.98
Treshnish	Outside of UK regions	0.47 (0.32, 0.68)	UK Western non-SPA colonies*	0
	UK North Sea and Channel waters	0.23 (0.15, 0.29)		0.02
	UK western waters	0.30 (0.16, 0.45)		0.98
Whinnyfold	Outside of UK regions	0.49 (0.17, 0.78)	UK North Sea non-SPA colonies*	0
	UK North Sea and Channel waters	0.46 (0.19, 0.74)		1.00
	UK western waters	0.05 (0.01, 0.11)		0

Table 9. Estimated proportion of adult breeding razorbill in the winter season (Nov-Dec) by tracked colony BDMPS area estimated as mean proportion with 95%CI, calculated from the GLS UD, along with equivalent BDMPS values for the most relevant corresponding population in BDMPS. *Designated SPA but not for razorbill features, hence included with non-SPA populations in BDMPS tables for this species.  denotes instances where the BDMPS proportion of adults falls within the estimated 95%CI bounds of the proportion of adults from a colony.

GLS utilisation distributions			BDMPS calculations	
Colony	Region	Mean proportion of adults from this colony in each BDMPS region (95% CI low,upp)	BDMPS population relevant to colony	Proportion of adults from this population in each BDMPS region
Canna	Outside of UK regions	0.48 (0.44, 0.51)	UK Western non-SPA colonies*	0.60
	UK North Sea and Channel waters	0.35 (0.31, 0.39)		0.10
	UK western waters	0.17 (0.14, 0.20)		0.30
East Caithness	Outside of UK regions	0.66 (0.60, 0.72)	East Caithness Cliffs SPA	0.69 
	UK North Sea and Channel waters	0.33 (0.28, 0.37)		0.30 
	UK western waters	0.02 (0.01, 0.03)		0.01 
Fair Isle	Outside of UK regions	0.63 (0.57, 0.69)	Fair Isle SPA	0.69
	UK North Sea and Channel waters	0.34 (0.27, 0.42)		0.30 
	UK western waters	0.03 (0.01, 0.05)		0.01 
Farnes	Outside of UK regions	0.59 (0.48, 0.70)	UK North Sea non-SPA colonies*	0.69 
	UK North Sea and Channel waters	0.40 (0.29, 0.52)		0.30 
	UK western waters	0 (0, 0.02)		0.01 
Isle of May	Outside of UK regions	0.55 (0.45, 0.65)	Forth Islands SPA	0.69
	UK North Sea and Channel waters	0.43 (0.34, 0.52)		0.30
	UK western waters	0.02 (0.01, 0.03)		0.01 
Orkney	Outside of UK regions	0.53 (0.45, 0.61)	UK North Sea non-SPA colonies*	0.69
	UK North Sea and Channel waters	0.39 (0.32, 0.46)		0.30
	UK western waters	0.08 (0.05, 0.11)		0.01
Shiant	Outside of UK regions	0.46 (0.41, 0.52)	Shiant Isles SPA	0.50 
	UK North Sea and Channel waters	0.29 (0.24, 0.32)		0.10

	UK western waters	0.25 (0.22, 0.29)		0.40
Treshnish	Outside of UK regions	0.39 (0.34, 0.45)	UK Western non-SPA colonies*	0.60
	UK North Sea and Channel waters	0.27 (0.23, 0.31)		0.10
	UK western waters	0.34 (0.29, 0.38)		0.30 
Whinnyfold	Outside of UK regions	0.64 (0.57, 0.71)	UK North Sea non-SPA colonies*	0.69 
	UK North Sea and Channel waters	0.35 (0.28, 0.41)		0.30 
	UK western waters	0.01 (0, 0.03)		0.01 

Visualizing density using space time cubes

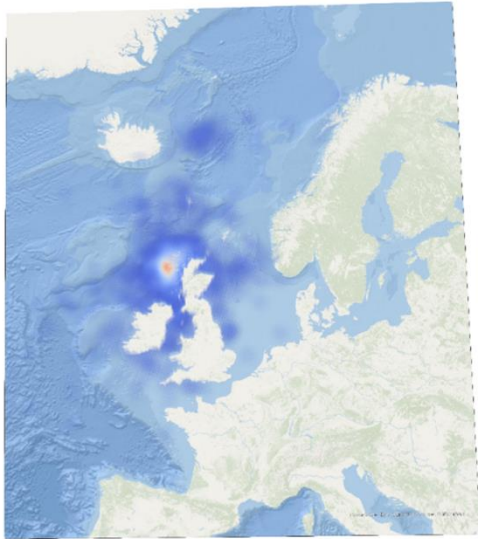
Figs. 35 and 36 show 3D (longitude, latitude, day of month) mean monthly densities of guillemot from the Canna and Isle of May colonies from July – March. The figures are also available as GIFs in the accompanying files (Canna_Guillemots_GIF and IoM_Guillemots_GIF).

There is a noticeable difference between the guillemots from the Isle of May and those from Canna in terms of the consistency and stability of their estimated distribution throughout the non-breeding season. The guillemots from the Isle of May are predominantly restricted to the North Sea and the Norwegian Sea for the majority of the non-breeding season with a gradual drift out from the coast with increased time since the breeding season and a tendency to move slightly northwards through time – though a small group of birds do appear to travel to the English channel from December to February.

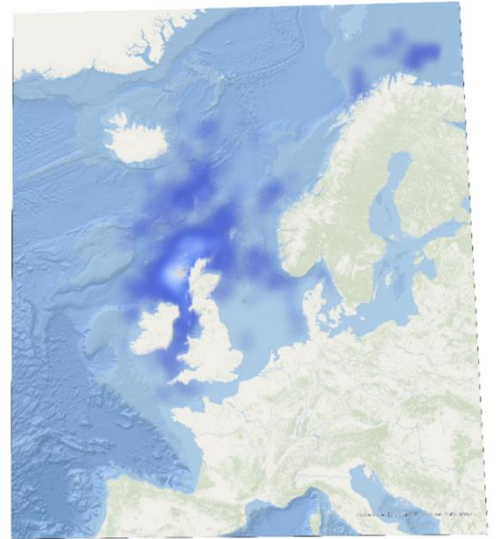
The Canna guillemots are similarly condensed around the colony early in the non-breeding season before gradually dispersing outwards; however, a small proportion were observed to travel as far as the north coast of Norway by as early as August. These birds which have travelled to Norway's north coast appear to return before the end of the breeding season, at which point a relatively large group travel south to the Bay of Biscay, predominantly during December and January. By March the majority of the population is congregated in a relatively small area off the north coast of Scotland.

Thus, it appears that the assumption of a static distribution throughout the non-breeding season may be more appropriate for some colonies than others, with guillemots from Canna acting as a good example of a population which shows substantial variability in its distribution throughout the non-breeding months.

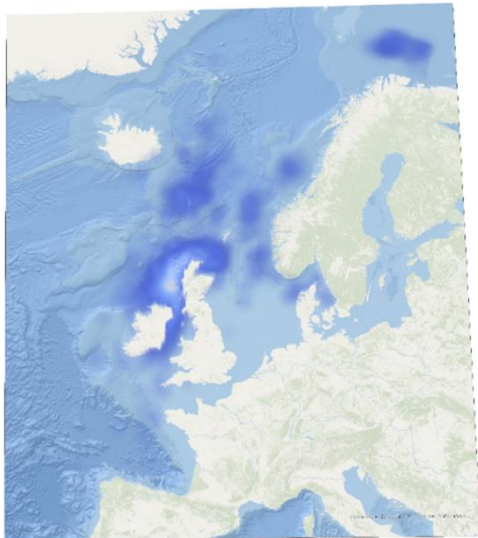
Species: Guillemots
Colony: Canna
Month: July



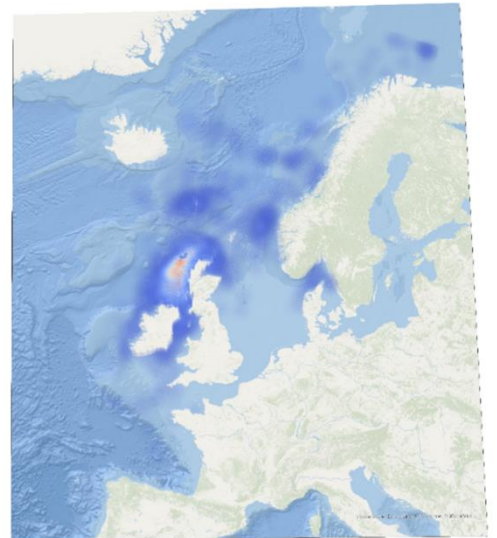
Species: Guillemots
Colony: Canna
Month: August



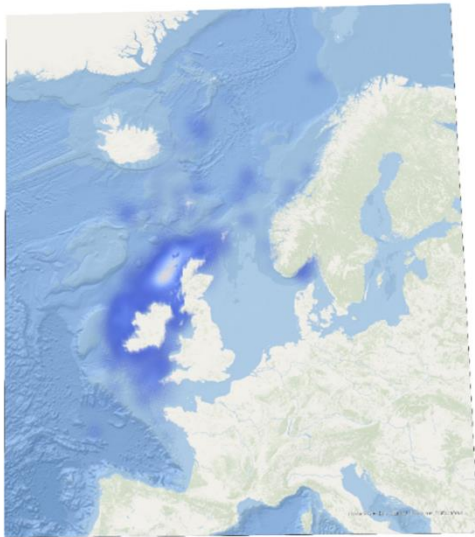
Species: Guillemots
Colony: Canna
Month: September



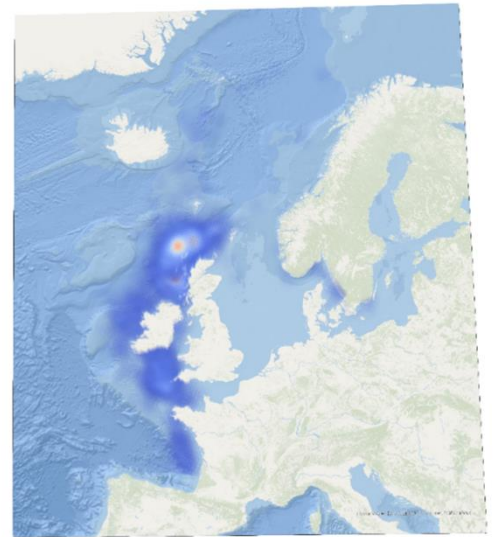
Species: Guillemots
Colony: Canna
Month: October



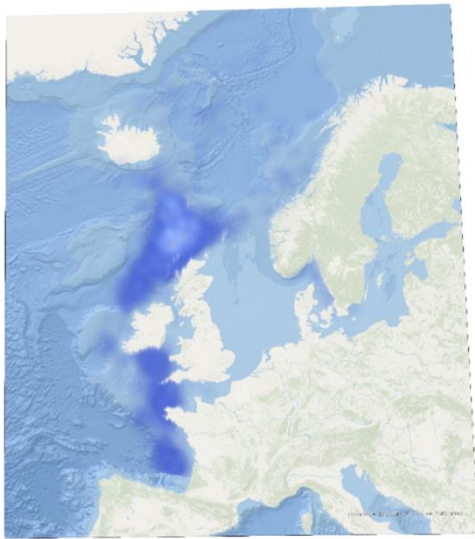
Species: Guillemots
Colony: Canna
Month: November



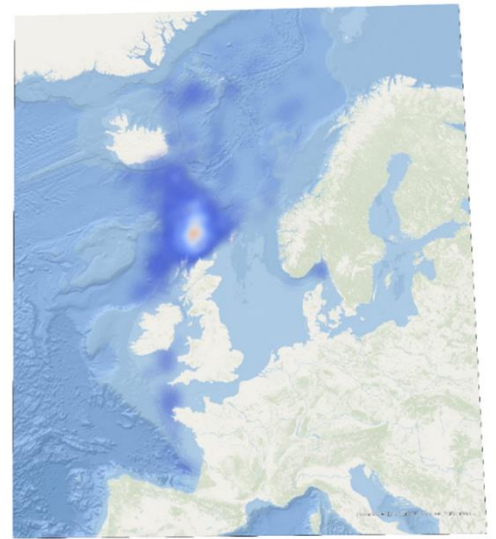
Species: Guillemots
Colony: Canna
Month: December



Species: Guillemots
Colony: Canna
Month: January



Species: Guillemots
Colony: Canna
Month: February



Species: Guillemots
Colony: Canna
Month: March

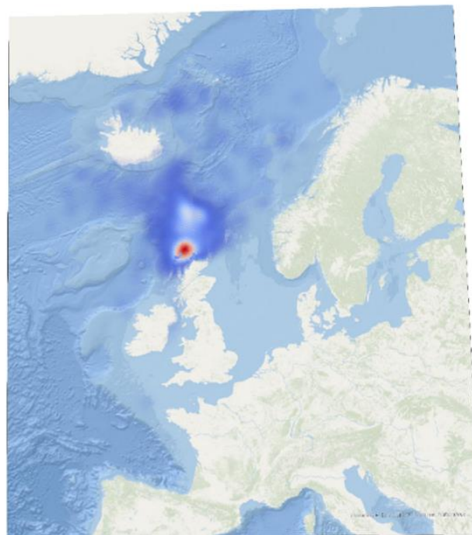
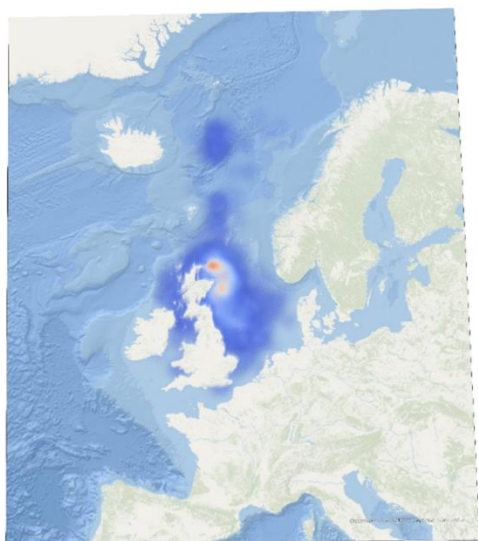
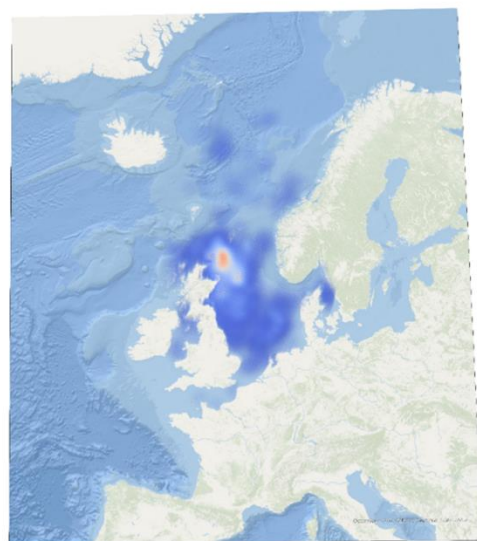


Figure 35. Space time cube visualisations showing guillemot mean usage by month from the Canna colony. The 3D usage represents longitude and latitude in the x and y axis and day of the month in the z-axis.

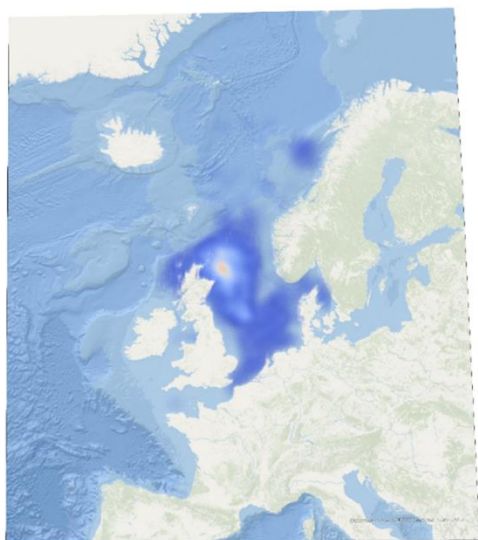
Species: Guillemots
Colony: Isle of May
Month: July



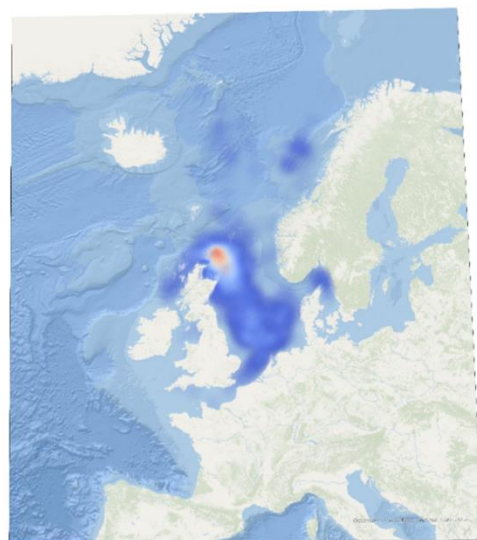
Species: Guillemots
Colony: Isle of May
Month: August



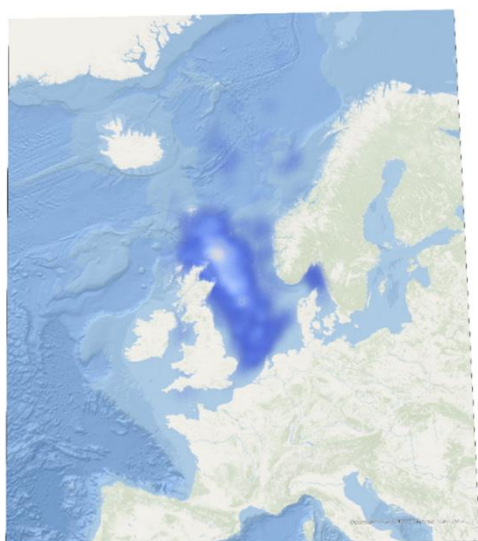
Species: Guillemots
Colony: Isle of May
Month: September



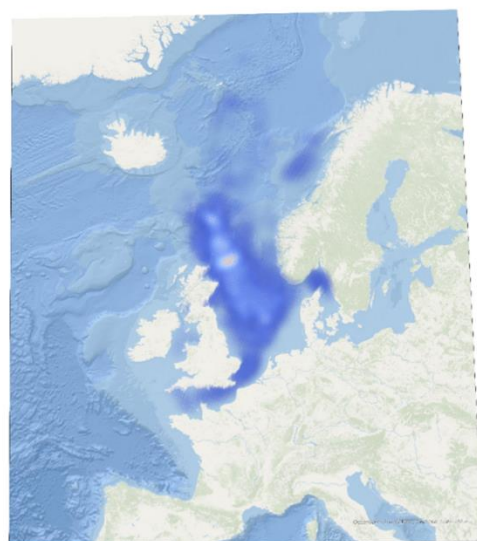
Species: Guillemots
Colony: Isle of May
Month: October



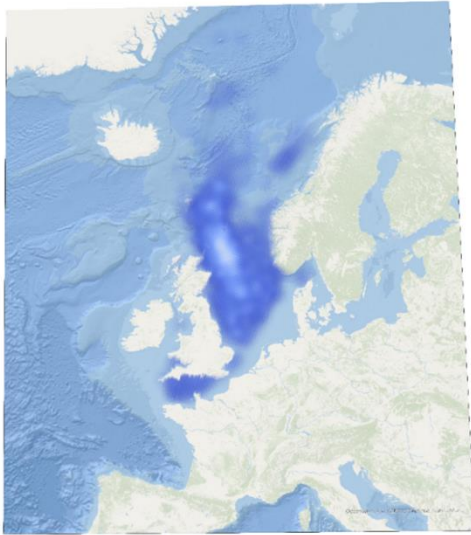
Species: Guillemots
Colony: Isle of May
Month: November



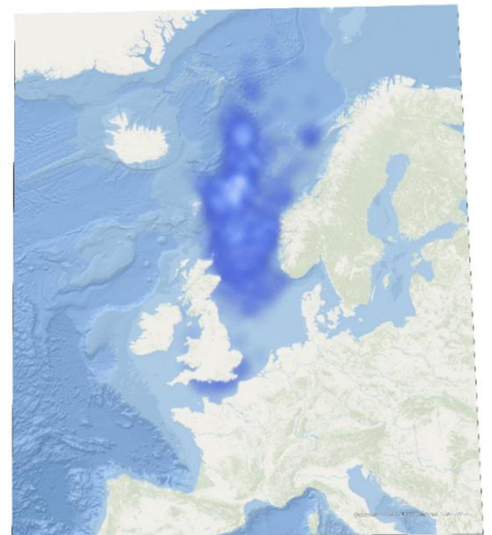
Species: Guillemots
Colony: Isle of May
Month: December



Species: Guillemots
Colony: Isle of May
Month: January



Species: Guillemots
Colony: Isle of May
Month: February



Species: Guillemots
Colony: Isle of May
Month: March

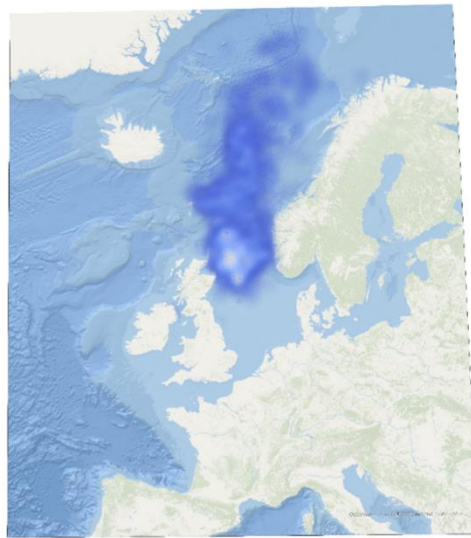


Figure 36. Space time cube visualisations showing guillemot mean usage by month from the Isle of May colony. The 3D usage represents longitude and latitude in the x and y axis and day of the month in the z-axis.

Potential to produce distributions for untracked colonies

The results of summarizing the distribution of distances between tracked and untracked colonies suggest that a reasonable compromise may be to assume that untracked colonies that are less than 270km from a tracked colony can be linked to GLS-based distributions from tracked colonies (Figures 37 and 38): the balance here is between ensuring that the maximum number of colonies can be capturing using the GLS-based maps (since they have substantial advantages over the existing apportioning approach) and not extrapolating these maps unreasonably far beyond the colonies that were used to create them. 94% and 92% of all UK adult breeding guillemots and razorbills, respectively, are within colonies that lie within 270km of a tracked colony (Fig. 6; **Error! Reference source not found.**Table 10). The intention is that the threshold of 270km may be used as a default threshold in the apportioning tool (beyond which BDMPS will be used in place of the GLS-based maps), but that users will be able to modify this threshold within the tool.

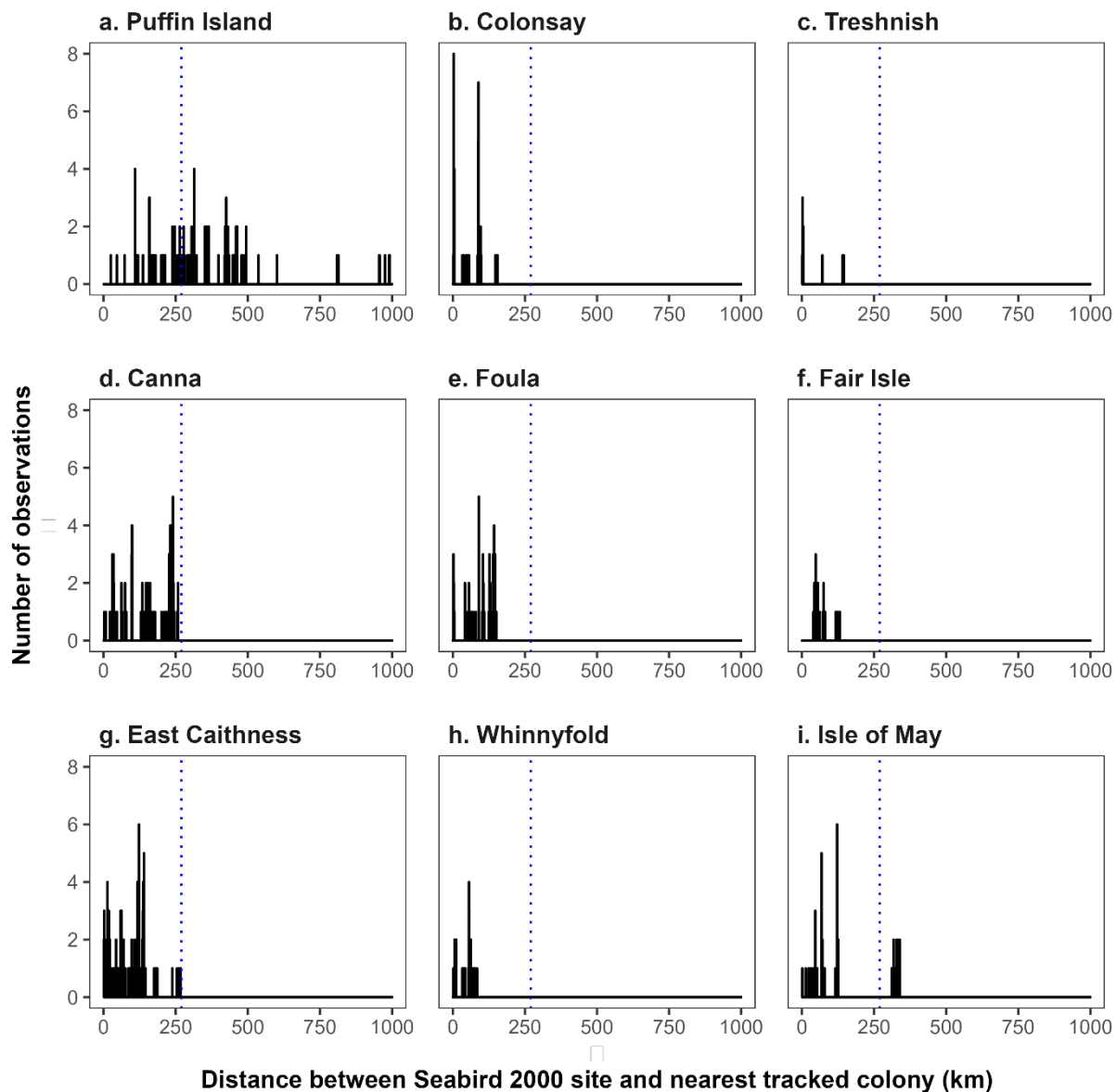


Figure 37. Histograms showing the distance between each Seabird 2000 site and the nearest tracked colony for common guillemots. The vertical blue dotted line indicates 270km.

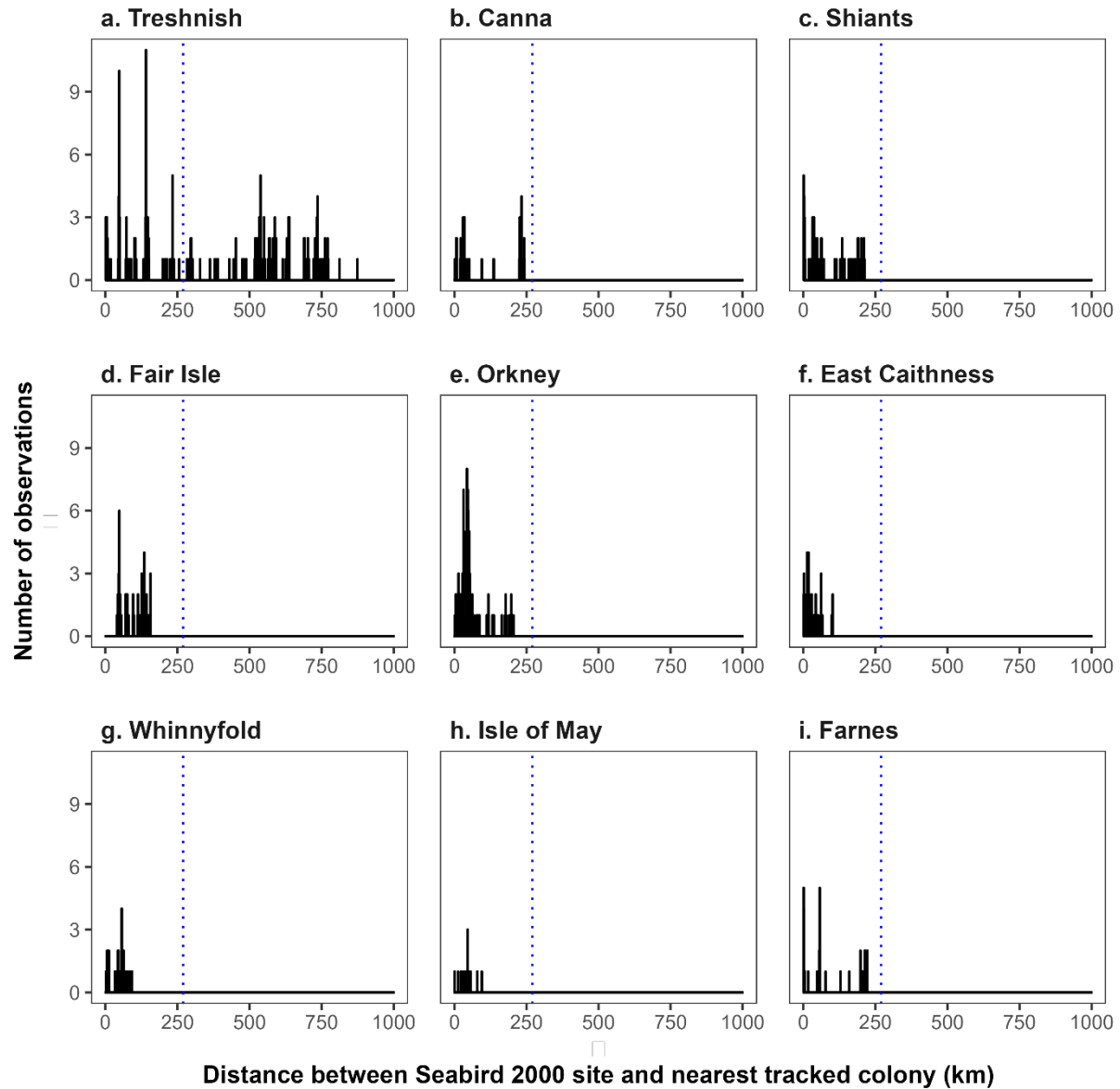


Figure 38. Histograms showing the distance between each Seabird 2000 site and the nearest tracked colony for razorbills. The vertical blue dotted line indicates 270km.

Table 10. The number of breeding common guillemots and razorbills represented by each tracked colony. Untracked colonies were assigned to the nearest colony with tracking data within 270km for each analysis. Population sizes were based on Seabird 2000 count data for each colony, multiplied by 0.67 to estimate the number of pairs and multiplied by 2 to achieve number breeding adults (Mitchell *et al.* 2004).

Tracked colony	Number of individuals represented	
	Guillemot	Razorbill
Canna	521215	8276
Colonsay	278200	-
East Caithness	696277	24450
Fair Isle	213949	18994
Farnes	-	18610
Foula	262704	-
Isle of May	227203	16977
Orkney	-	38022
Puffin Island	86438	-
Shiants	-	66953
Treshnish	88058	115914
Whinnyfold	299787	36058
Total	2673831	344254
Percentage of adult UK breeding population (%)	94.06	91.82

Discussion

The use of geolocator-based models in apportioning

The geolocator-based models are colony-specific, and produced estimated spatial distributions associated with colonies that have GLS tracking data, together with a quantification of the uncertainty associated with these distributions. However, non-breeding season spatial distributions of seabirds are much less straightforward to model in relation to explanatory variables than breeding season distributions, because the central place foraging constraint in the breeding season means that the distribution will necessarily be closely related to distance to colony. This, and the fact that the much higher levels of uncertainty in GLS tracking data than in GPS tracking data would also make it difficult to detect

habitat associations, explain why colony-specific models have been used for the analysis of GLS tracking data, whereas habitat association models that can be generalized beyond the colonies whose data was used in fitting them have been used in analysis of breeding season GPS tracking data (e.g. analysis of lesser black-backed gulls in this project, and Wakefield et al., 2017).

Apportioning, however, relies upon knowing the spatial distributions associated with every colony, not only those colonies with tracking data. In order for the geolocator-based models to be used for apportioning it was therefore necessary to either be able to use these models to produce spatial distributions for colonies without GLS tracking data, or else to use an alternative approach for colonies without GLS tracking data. The only obvious alternative approach to use in this context was BDMPS. Since there are limitations to BDMPS, the approach that we use is to extrapolate the geolocator-based models out to colonies without GLS tracking data in situations where these colonies are within a certain distance threshold of a colony with GLS tracking data, and to use BDMPS in situations where the distance to the nearest colony with GLS tracking data exceeded this threshold. This will therefore provide a hierarchical approach in the apportioning tool being developed in [WP4](#), in which the best available method is used for each colony. We propose that the tool will allow users to specify this distance threshold, allowing users to investigate the sensitivity of the results to the choice of distance threshold. The choice of the distance threshold represents a trade-off between the caveats associated with extrapolating the GLS maps to untracked colonies and the caveats associated with BDMPS. The tool will show a map that illustrates the impact of varying the distance threshold on the set of colonies for which GLS-based distributions will be used.

The spatial distributions of each untracked colony will be based on a weighted average of the spatial distributions associated with tracked colonies that lie within the distance threshold of this colony; we propose to use a simple linear weighting, so that:

$$\text{Spatial distribution of untracked colony} = \text{SUM over all tracked colonies of (weight of tracked colony * spatial distribution of tracked colony)}$$

[Equation 1]

where the weight of each tracked colony relative to the untracked colony is assumed to be proportional to

$$(\text{distance threshold} - \text{distance from untracked colony to tracked colony}) / (\text{distance threshold})$$

[Equation 2]

Tracked colonies that are closer to the untracked colony will therefore be given more weight than those that are further away.

We anticipate that for the apportioning tool colonies will be based on Seabird 2000 sub-sites, and the size of these colonies based on Seabird 2000 counts, for consistency with the apportioning approaches used in the breeding season. The use of Seabird 2000 ensures that all colonies are included, and that counts for them were taken in a comparable period. One challenge, however, is that BDMPS uses SPAs, and aggregates non-SPA colonies into just two very broad regional groups for guillemot and razorbill: North Sea and Western / West Coast. In areas where the tool defaults to BDMPS it will therefore need to use these definitions rather than the Seabird 2000 colony definitions. This will require some adjustments – e.g. the population size of the non-SPA populations in BDMPS will need to be reduced to remove those non-SPA populations that lie in the areas where the geolocator-based maps are being used for apportioning.

A key advantage of the approach of using GLS-based distributions in the areas for which data are available, and using BDMPS elsewhere, is that this hybrid approach provides a straightforward mechanism for non-UK populations to be incorporated into the apportioning calculations: these will simply be accounted for as they are currently are within BDMPS, using the same estimates for non-UK populations as are currently used in BDMPS. There would be potential in future to exploit GLS data from outside the UK to provide new empirical estimates of non-breeding season distributions for non-UK colonies, but until such estimates are available the hybrid approach will ensure that non-UK populations continue to be accounted for within apportioning.

The tool will account for uncertainty in the GLS maps, for uncertainty in the location of the boundary between BDMPS regions (capturing the idea that the location of the boundary was never intended to be precise, but to represent a broad separation of two distinct regions), and for uncertainty in the allocation of tracked colonies to untracked colonies (e.g. by using the weights from Equation 2 as probabilities within a simulation-based approach).

Given the variations in spatial distribution between months (e.g. as shown by the space-time cube visualisations), and the challenges in producing biological definitions of seasons, the tool will allow users a high level of flexibility in specifying the time period over which apportioning will be performed. In particular, we intend that users will be able to use the tool to either apportion for seasons, whose start and end months they can specify, or for individual months. This will futureproof the tool against any possible future changes in the way in which assessments define seasons, and against any potential future switch from apportioning at a seasonal level to apportioning at a monthly level.

The structure of the tool is proposed to be similar to that used for current breeding season apportioning tools: users will have an option to either select “map mode” (in which case apportioning probabilities will be mapped, with uncertainty, for a selected colony), or “footprint mode” (in which case users will upload a footprint, and the apportioning probabilities will be estimated, with uncertainty, for all colonies within this footprint). Unlike in breeding season tools, however, the intention is that users will be able to select the month or season that they wish to consider, since the GLS-based maps are monthly. It is anticipated that users will also be able to select the distance threshold, beyond which BDMPS should be used, and to select the level of uncertainty they assume exists in the boundary between BDMPS regions.

Wider interpretations

The difference between BDMPS and GLS-derived distributions could have arisen from the different scales at which they have been estimated, with BDMPS presented at very broad scales. Another possibility is that information on colony provenance in BDMPS was derived primarily from ringing recoveries, which is the largely coastal distribution of birds found dead, in the limited availability of relevant geolocation data at the time that the work was undertaken. The marked difference in the comparison between the three categories (guillemot non-breeding, razorbill migration and razorbill winter) was instructive – the greater accordance between methods in the latter may be due to the high level of aggregation of different razorbills in midwinter, in contrast to guillemots (Buckingham et al. 2022). As such, the reliability of BDMPS as a tool for winter apportioning is likely to be context dependent. It should also be noted that the hard boundaries that were placed between BDMPS in order to undertake the calculations required in the comparison, as outlined in the last section, may not be appropriate if those boundaries were originally considered to be more diffuse.

The GLS data were based on a study period of 2017-21, comprising four non-breeding seasons but with a particular emphasis on 2017-18 and 2018-9 for both species and 2019-20 for guillemots. As such, the

data are unlikely to fully capture interannual variation in winter distribution. Further data collection at the study colonies and elsewhere would be useful in factoring in interannual variation more completely, and the level of variation could also be explored more comprehensively in longer term studies, though most of these have taken place outside the UK, with the exception of the Isle of May. The representativeness of the study colonies of the larger suite of untracked colonies is untested, and warrants further investigation. However, a more urgent requirement is to plug gaps in data from regions in the UK with no GLS data, such as areas in western Scotland and Wales and the large colony at Flamborough Head.

Our analysis identified a particular challenge with geolocation data during the post-breeding and pre-breeding periods, when a large proportion of birds are in close association with colonies. At these times, the locations are likely to be more greatly affected by shading at dawn and dusk associated with colony attendance, which can affect the estimated time of dawn and/or dusk and, in turn, location estimates. As such, we would advise that distributions at this time are treated with more caution than at other times of the non-breeding season. However, it is important to clarify that guillemots may attend the colony in all months of the non-breeding season to a greater or lesser extent after completing their moult i.e. any month from October through to the breeding season (Bennett et al. 2022). As such, further work is required that considers the potential effect of colony attendance outside the breeding season on geolocation data.

4. Conclusions

This workpackage has focused upon providing new apportioning methods, in two key situations in which there was previously only one method available – summer apportioning for lesser-black backed gulls and winter apportioning for guillemot and razorbill - and in evaluating the spatial distributions that underpin the new methods against those that underpin the existing methods.

The results suggest that there are some broad similarities between the results obtained using a simple inverse distance squared decay rule (as used in the SNH/NatureScot Apportioning Tool) and those obtained using summer GPS-based maps for lesser black-backed gulls, reflecting similar results in Butler et al. (2020) for the other four species for which summer GPS-based maps have been used for apportioning. This broad similarity presumably reflects the dominance of distance to colony in describing the summer distributions of breeding seabirds, given the central place foraging constraint. The results did reveal differences in the spatial distributions obtained using the two methods, however, and identified properties of the SNH/NatureScot Apportioning Tool that will be explored further in [WP4](#).

In contrast, the GLS-based maps for the non-breeding season for guillemots and razorbill imply proportions of time spent in BDMPS regions that differ substantially from those used in BDMPS. The larger differences between methods in the non-breeding season are to be expected, given that there is less likely to be any single, readily quantifiable, dominant factor in determining non-breeding season distributions. Further evaluation is more challenging here, since BDMPS only considers broad-scale regions.

Within [WP4](#) the spatial distributions that have been produced and evaluated here will be converted into apportioning estimates. The conversion into apportioning estimates will require some additional, relatively simple, steps: for lesser black-backed gulls the application of a colony-specific adjustment to account for proportion of time spent foraging on land, and for guillemot and razorbill the use of a hybrid approach that combines GLS-based spatial distributions (where appropriate) with BDMPS (for colonies that lie far from the nearest colony with GLS data). The apportioning estimates obtained using different

methods will also be evaluated against each other, complementing the evaluation here of the colony-specific spatial distributions that underpin these estimates, and the different apportioning methods will be integrated into a user-friendly tool. This tool will also incorporate a RAG dashboard, that provides a straightforward way to identify the appropriateness of each tool.

The [WP4](#) report will also outline the process that will be needed to update the tool as and when new data and model outputs become available in future. In general, statistical modelling of GPS and GLS tracking data involves relatively sophisticated modelling, and requires both biological and statistical judgement, and this process cannot therefore usefully be completely automated, so future updating of the models to incorporate new data will be a relatively substantial piece of analytical work. The process for updating the tool to incorporate new model outputs will, however, be made as clear and streamlined as possible, minimizing the amount of work required to integrate new model outputs into the tool.

5. Acknowledgements

Analysis of GPS data for lesser black-backed gulls

Data sources

MoveBank Studies, available on agreement and request from: www.movebank.org:

Booth-Jones, K. (2018) '2018 BTO NI Belfast gulls' Movebank study ID: 466426501

Clewley, G.C., Thaxter, C.B., Scragg, E. & Burton, N. (2017) 'BTO – Ribble Estuary 2017 - Lesser Black-backed Gull'. Movebank study ID: 277841852

Clewley, G.C., Thaxter, C.B., Scragg, E. & Burton, N. (2018) 'BTO – Ribble Estuary 2018 - Lesser Black-backed Gull'. Movebank study ID: 482136485

Clewley, G.C., Thaxter, C.B., Scragg, E. & Burton, N. (2019) 'BTO – Ribble Estuary 2019 - Lesser Black-backed Gull'. Movebank study ID: 849740134

Humphreys, E.M., Clewley, G.C., Johnston, D.T., Thaxter, C.B. & Burton, N. (2018) 'BTO – Firth of Forth 2019 - Herring Gull'. Movebank study ID: 834346398

Humphreys, E.M., Clewley, G.C., Johnston, D., Thaxter, C.B. & Burton, N. (2018) 'BTO – Firth of Forth 2019 - Lesser Black-backed Gull'. Movebank study ID: 834344857

Thaxter, C.B., Clewley, G.C., Scragg, E. & Burton, N. (2016) 'BTO – North West England 2016 - Lesser Black-backed Gull'. Movebank study ID: 167983392

Thaxter, C.B., Clewley, G.C., Scragg, E. & Burton, N. (2017) 'BTO – Barrow 2017 - Lesser Black-backed Gull'. Movebank study ID: 277843654

Thaxter, C.B., Clewley, G.C., Scragg, E. & Burton, N. (2018) 'BTO – Barrow 2017 - Lesser Black-backed Gull'. Movebank study ID: 482136669

UvA-Studies, available with agreement on request from www.uva-bits.nl :

Clewley, G.C., Thaxter, C.B., Scragg, E., Bouten, W. & Burton, N. (2017) Lesser Black-backed Gull – Ribble Estuary.

Humphreys, E.M., Clewley, G.C., Johnston, D., Thaxter, C.B., Bouten, W. & Burton, N. (2017) Lesser Black-backed Gull – Isle of May.

Thaxter, C.B., Ross-Smith, V.H., Bouten, W. and Burton, N.H.K. (2011) Lesser Black-backed Gull – Orford Ness.

Thaxter, C.B., Ross-Smith, V.H., Bouten, W. and Burton, N.H.K. (2014) Lesser Black-backed Gull – Walney.

Thaxter, C.B., Ross-Smith, V.H., Bouten, W. and Burton, N.H.K. (2014) Lesser Black-backed Gull – Skokholm.

The UvA-BiTS tracking studies are facilitated by infrastructures for e-Science, developed with support of the NLeSC (<http://www.esciencecenter.com/>) and LifeWatch, carried out on the Dutch national e-infrastructure with support from the SURF Foundation. The data are held jointly by the BTO, University of Amsterdam and the funders of the project, and can be made available through their agreement.

General acknowledgements

Studies at Walney (2010-14), Skokholm, Orford Ness and the Firth of Forth were funded by the Department for Business, Energy and Industrial Strategy (BEIS) Offshore Energy Strategic Environmental Assessment (OESEA) research programme – and we are particular grateful to John Hartley of Hartley Anderson for support and initiation of this work – and supported at Walney by the Environmental Research Institute (North Highland College, University of the Highlands and Islands) through funding from the Marine Renewable Energy and the Environment (MaREE) project (funded by Highlands and Islands Enterprise, the European Regional Development Fund, and the Scottish Funding Council). Further work at Walney and Barrow (2014-16) was funded by Ørsted and supported by Natural England. Natural England also funded further studies, with BAE Systems, at the Ribble Estuary. Work at Bangor was funded by Natural Resources Wales and the study in Belfast was funded by the Northern Ireland Environment Agency, Northern Ireland Water, and the Agri-Food and Biosciences Institute.

We thank all landowners, together with Natural England, Natural Resources Wales, NatureScot and the Northern Ireland Environment Agency for permissions to work on the sites concerned. We also thank the many people who have assisted in fieldwork, notably Liz Humphreys, Gary Clewley, Viola Ross-Smith, Emily Scragg, Nigel Clark, Katherine Booth-Jones, and Shane Wolsey as field leads. Niall Burton provided strategic lead of the gull tracking programme at BTO.

Special thanks to Willem Bouten, Edwin Baaij and Judy Shamoun-Baranes for their invaluable assistance and collaboration through projects using the University of Amsterdam Bird-Tracking System; Elizabeth Masden for supply of some UvA-BiTS tags and tracking equipment for Walney; and Phil Atkinson and Marta Acacio from Movetech Telemetry.

Analysis of geolocator data for guillemot and razorbill

This analysis was underpinned by data collection and processing funded by a PhD studentship to Lila Buckingham as part of a funding package from Vattenfall to MacArthur Green. We are grateful for the supervision and analytical discussion during Lila's PhD provided by Jonathan Green and Robert Furness, including their insights into seabird winter foraging ecology and the use of geolocation data to inform Offshore Wind Farm effects outside the breeding seasons. Data collection was funded by Vattenfall, Equinor (as part of Hywind Scotland's Environmental Monitoring Programme) and SEATRACK. We are very grateful to all individuals and groups who contributed to data collection: Sophie Bennett, Richard Bevan, Andrew Call, Michael Canham, Colin Corse, Steve Dodd, Michael Harris, Christopher Heward, David Jardine, Mark Newell, Jim Lennon, David Parnaby, Chris Redfern, Liz Scott, Robert Swann, Robin Ward, Sarah Wanless, Ewan Weston, Chris Andrews, Phil Bloor, Calum Campbell, Martin Davison, Hayley Douglas, Raymond Duncan, Sarah Fenn, Alexander Gilliland, Kenny Graham, Robin Gray, Carrie Gunn, Daniel Johnston, Anne Middleton, Tim Morley, Robert Rae, Stuart Rae, Moray Souter, Caitlin Tarvet, Jenny Weston, Alastair Young, Fair Isle Bird Observatory, Highland Ringing Group, Newcastle University, Orkney Ringing Group, SCAN Ringing Group, Seabird Ecology Group at the University of Liverpool (SEGUL), Shiant Auk Ringing Group, Treshnish Isles Auk Ringing Group. We thank Ruth Dunn for discussions regarding geolocator analysis. We thank Hallvard Strøm and Vegard Bråthen for their help and field support with guillemots on the Isle of May as part of the SEATRACK project (<http://www.seapop.no/en/seatrack/>). We thank all landowners for access and NatureScot for logistical support on the Isle of May.

6. References

- Ahrens, J., Geveci, B., Law, C., (2005) ParaView: An End-User Tool for Large Data Visualization, Visualization Handbook, Elsevier, ISBN-13: 978-0123875822
- Anderson, G.B., Bell, M.L., Peng, R.D. (2013) Methods to calculate the heat index as an exposure metric in environmental health research. *Environmental Health Perspectives* 121(10), 1111-1119. R package version 1.2.2.
- Bates D, Mächler M, Bolker B, Walker S (2015) Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1–48. doi:10.18637/jss.v067.i01. R package version 1.1-29.
- Bennett, S., Harris, M.P., Wanless, S., Green, J.A., Newell, M.A., Searle, K.R. & Daunt, F. Earlier and more frequent occupation of breeding sites during the non-breeding season increases breeding success in a colonial seabird. *Ecology & Evolution* <https://doi.org/10.1002/ece3.9213>
- Bivand, R.S., Pebesma, E. & Gomez-Rubio, V. (2013) *Applied spatial data analysis with R*, Second edition. Springer, NY. <https://asdar-book.org/>
- Bivand, R. & Rundel, C. (2021) rgeos: Interface to Geometry Engine - Open Source ('GEOS'). R package version 0.5-9. <https://CRAN.R-project.org/package=rgeos>
- Bivand, R., Keitt, T. & Rowlingson, B. (2022) rgdal: Bindings for the 'Geospatial' Data Abstraction Library. R package version 1.5-32. <https://CRAN.R-project.org/package=rgdal>
- Bogdanova, M.I., Wanless, S., Harris, M.P., Lindström, J., Butler, A., Newell, M.A., Sato, K., Watanuki, Y., Parsons, M. and Daunt, F. (2014) Among-year and within-population variation in foraging distribution of European shags *Phalacrocorax aristotelis* over two decades: implications for marine spatial planning. *Biological Conservation*, 170, 292-299.
- Booth Jones, K., Thaxter, C., Clewley, G., Wolsey, S., Calbrade, N., Atkinson, P., Calladine, J. & Burton, N. (2021) *Belfast's Urban Gulls: An Assessment of Breeding Populations, Breeding Season Movements and Winter Population*. BTO Research Report carried out on behalf of the Northern Ireland Environment Agency and Agri-Food and Biosciences Institute. March 2021.
- Bouten, W., Baaij, E.W., Shamoun-Baranes, J. & Camphuysen, K.C.J. (2013) A flexible GPS tracking system for studying bird behaviour at multiple scales. *Journal of Ornithology*, 154, 571-580.
- Brabant, R., Vanermen, N., Stienen, E. W. M., & Degraer, S. (2015) Towards a cumulative collision risk assessment of local and migrating birds in North Sea offshore wind farms. *Hydrobiologia*, 75, 63–74.
- Buckingham, L., Bogdanova, M.I., Green, J.A., Dunn, R.E., Wanless, S., Bennett, S., Bevan, R.M., Call, A., Canham, M., Corse, C.J., Harris, M.P., Heward, C.J., Jardine, D.C., Lennon, J., Parnaby, D., Redfern, C.P.F., Scott, L., Swann, R.L., Ward, R.M., Weston, E.D., Furness, R.W. & Daunt, F. (2022) Interspecific variation in non-breeding aggregation: a multi-colony tracking study of two sympatric seabirds. *Marine Ecology Progress Series*, 684, 181–197.
- Burnell, D. (2021a) Urban nesting Herring Gull *Larus argentatus* and Lesser Black-backed Gull *Larus fuscus* population estimates: devising species-specific correction models for ground-based survey data. Natural England publication ref: JNCC21_01

- Burnell, D. (2021b) Population estimates for urban and natural nesting Herring Gull *Larus argentatus* and Lesser Black-backed Gull *Larus fuscus* in England. Natural England publication ref: JNCC21_02
- Calenge, C. (2006) The package “adehabitat” for the R software: A tool for the analysis of space and habitat use by animals. *Ecological Modelling*, 197, 516–519.
- Carroll, M.J., Butler, A., Owen, E., Ewing, S.R., Cole, T., Green, J.A., Soanes, L.M., Arnould, J.P.Y., Newton, S.F., Baer, J., Daunt, F., Wanless, S., Newell, M.A., Robertson, G.S., Mavor, R.A. & Bolton, M. (2016) Effects of sea temperature and stratification changes on seabird breeding success. *Climate Research*, 66 (1), 75-89.
- Clewley, G.D., Johnston, D., Thaxter, C.B., Green, R.M., Bouten, W., Burton, N.H.K. & Humpreys, E.M. (2020) Measuring the interaction between marine features of Special Protection Areas with offshore wind farm development zones through telemetry - Lesser Black-backed Gulls in the Forth Islands Special Protection Area. First Year Report on behalf of Department for Business, Energy & Industrial Strategy.
- Clewley, G.D., Clark, N. A., Thaxter, C.B., Green, R.M., Scragg, E.S & Burton, N.H.K. (2022) Development of a weak-link wing harness for use on large gulls (Laridae): methodology, evaluation and recommendations. *Seabird*.
- D'Agostino McGowan, L. & Bryan, J. 2021. googledrive: An Interface to Google Drive. R package version 2.0.0. <https://CRAN.R-project.org/package=googledrive>
- Demšar, U. (2015) Code for calculation of space time densities from a set of movement trajectories. <https://github.com/udemsar/SpaceTimeDensities>
- Demšar, U., Buchin, K., van Loon, E.E. et al. (2015). Stacked space-time densities: a geovisualisation approach to explore dynamics of space use over time. *Geoinformatica*, 19, 85–115. <https://doi.org/10.1007/s10707-014-0207-5>
- Dunn, R.E., Wanless, S., Daunt, F., Harris, M.P. & Green, J.A. (2020) A year in the life of a North Atlantic seabird: behavioural and energetic adjustments during the annual cycle. *Scientific Reports*, 10, 5993.
- Eaton, M. A., Aebischer, N. J., Brown, A. F., Hearn, R. D., Lock, L., Musgrove, A. J., ... Gregory, R. D. (2015). Birds of Conservation Concern 4: the population status of birds in the United Kingdom, Channel Islands and Isle of Man. *British Birds*, 108, 708–746.
- EU Copernicus Marine Service Information (2021) Atlantic- European North West Shelf-Ocean Physics Reanalysis, NWSHELF_MULTIYEAR_PHY_004_009. https://resources.marine.copernicus.eu/product-detail/NWSHELF_MULTIYEAR_PHY_004_009/INFORMATION
- Evans, J.S. (2021) _spatialEco_. R package version 1.3-6, <https://github.com/jeffreyevans/spatialEco>.
- Fauchald, P., Tarroux, A., Amélineau, F., Bråthen, V.S., Descamps, S., Ekker, M., Helgason, H.H., Johansen, M.K., Merkel, B., Moe, B., Åström, J., Anker-Nilssen, T., Bjørnstad, O., Chastel, O., Christensen-Dalsgaard, S., Danielsen, J., Daunt, F., Dehnhard, N., Erikstad, K.E., Ezhov, A., Gavrilov, M., Hallgrímsson, G.T., Hansen, E.S., Harris, M., Helberg, M., Jónsson, J.E., Kolbeinnsson, Y., Krasnov, Y., Langset, M., Lorentsen, S.H., Lorentzen, E., Newell, M., Olsen, B., Reiertsen, T.K., Systad, G.H., Thompson, P., Thórarinnsson, T.L., Wanless, S., Wojczulanis-Jakubas, K. & Strøm, H. (2021) Year-round distribution of Northeast Atlantic seabird populations: applications for population management and marine spatial planning. *Marine Ecology Progress Series*, 676, 255–276.

- Feistel, R. (2004) A new extended Gibbs thermodynamic potential of seawater, *Prog. Oceanogr.*, 58, 43-115, <http://authors.elsevier.com/sd/article/S0079661103000880> ; corrigendum 61, 99, 2003.
- Fort, J., Pettex, E., Tremblay, Y., Lorentsen, S.-H., Garthe, S., Votier, S., ... Gremillet, D. (2012). Meta-population evidence of oriented chain migration in northern gannets (*Morus bassanus*). *Frontiers in Ecology and Environment*, 10(5), 237–242.
- Frederiksen, M., Moe, B., Daunt, F., Phillips, R. A., Barrett, R. T., Bogdanova, M. I., ... Anker-Nilssen, T. (2012). Multicolony tracking reveals the winter distribution of a pelagic seabird on an ocean basin scale. *Diversity and Distributions*, 18(6), 530–542.
- Furness, R.W. (2015). Non-breeding season populations of seabirds in UK waters: Population sizes for Biologically Defined Minimum Population Scales (BDMPS). Natural England Commissioned Reports, Number 164
- Furness, R.W., Wade, H.M. and Masden, E.A. (2013) Assessing vulnerability of marine bird populations to offshore wind farms. *Journal of Environmental Management*, 119, 56–66.
- Garthe, S., & Huppopp, O. (2004). Scaling possible adverse effects of marine wind farms on seabirds: Developing and applying a vulnerability index. *Journal of Applied Ecology*, 41(4), 724–734.
- GEBCO (2009) The GEBCO_08 Grid, version 20091120, General Bathymetry Chart of the Oceans.
- Geen, G. R., Robinson, R. A., & Baillie, S. R. (2019) Effects of tracking devices on individual birds – a review of the evidence. *Journal of Avian Biology* (doi.org/10.1111/jav.01823).
- Glew, K. St J., Wanless, S., Harris, M. P., Daunt, F., Erikstad, K. E., Strøm, H., & Trueman, C. N. (2018) Molt location and diet of auks in the North Sea inferred from coupled light-based and isotope-based geolocation. *Marine Ecology Progress Series*, 599, 239–251.
- Halpin, L.R., Ross, J.D., Ramos, R., Mott, R., Carlile, N., Golding, N., Reyes-González, J.M., Militão, T., De Felipe, F., Zajková, Z., Cruz-Flores, M., Saldanha, S., Morera-Pujol, V., Navarro-Herrero, L., Zango, L., González-Solís, J. & Clarke, R.H. (2021) Double-tagging scores of seabirds reveals that light-level geolocator accuracy is limited by species idiosyncrasies and equatorial solar profiles. *Methods in Ecology and Evolution*, 12, 2243–2255.
- Hijmans (2022a) raster: Geographic Data Analysis and Modeling. R package version 3.5-21. <https://CRAN.R-project.org/package=raster>
- Hijmans (2022b) terra: Spatial Data Analysis. R package version 1.5-34. <https://CRAN.R-project.org/package=terra>
- Holt, J., Wakelin, S., Lowe, J. & Tinker, J. (2010) The potential impacts of climate change on the hydrography of the northwest European continental shelf. *Progress in Oceanography*, 86, 361–379.
- Jackett, D.R., McDougall, T.J., Feistel, R., Wright, D.G. & Griffies, S.M. (2005) Algorithms for density, potential temperature, conservative temperature, and the freezing temperature of seawater. *Journal of Atmospheric and Oceanic Technology*, 23, 1709-1728
- Johnston, D.T., Thaxter, C.B., Boersch-Supan, P.H., Humphreys, E.M., Bouten, W., Clewley, G.D., Scragg, E.S., Masden, E.A., Barber, L., Conway, G.J., Clark, N.A., Burton, N.H.K. & Cook, A.S.C.P. (2022)

Investigating avoidance and attraction responses in lesser black-backed gulls *Larus fuscus* to offshore wind farms. *Mar Ecol Prog Ser*, 686, 187-200. <https://doi.org/10.3354/meps13964>.

JNCC (2021) <https://jncc.gov.uk/our-work/lesser-black-backed-gull-larus-fuscus/>

Kara, A.B., Rochford, P.A. & Hurlburt, H.E. (2003) Mixed layer depth variability over the global ocean. *Journal of Geophysical Research*, 108, No. C3, 3079, doi:10.1029/2000JC000736

Langley, L.P., Bearhop, S., Burton, N.H.K., Banks, A.N., Frayling, T., Thaxter, C.B., Clewley, G.D., Scragg, E.S. & Votier, S.C. (2021) GPS tracking reveals landfill closures induce higher foraging effort and habitat switching in gulls. *Movement Ecology*, 9, 56, <https://doi.org/10.1186/s40462-021-00278-2>.

Langley, L.P., Bearhop, S., Burton, N.H.K., Banks, A.N., Frayling, T., Thaxter, C.B., Clewley, G.D., Scragg, E., Votier, S.C. (2022) Urban and coastal breeding lesser black-backed gulls (*Larus fuscus*) segregate by foraging habitat. *Ibis* early online: <https://doi.org/10.1111/ibi.13109>.

Merkel, B., Phillips, R.A., Descamps, S., Yoccoz, N.G., Moe, B. & Strøm, H. (2016) A probabilistic algorithm to process geolocation data. *Movement Ecology*, 4, 26.

Merkel, B. (2018) Probabilistic algorithm for geolocation data. <https://github.com/benjamin-merkel/probGLS>

Miller, P.I. & Christodoulou, S. (2014) Frequent locations of oceanic fronts as an indicator of pelagic diversity: Application to marine protected areas and renewables. *Marine Policy*, 45, 318–329

Mitchell, P.I., Newton, S., Ratcliffe, N. & Dunn, T.E. (eds) (2004) *Seabird Populations of Britain and Ireland*. T & AD Poyser, London, UK

Mitchell, P.I., Newton, S., Ratcliffe, N. & Dunn, T.E. (2004) *Seabird Populations of Britain and Ireland: results of the Seabird 2000 census*. Joint Nature Conservation Committee, Peterborough.

Monterey, G. & Levitus, S. (1997) Seasonal variability of mixed layer depth for the World Ocean. NOAA Atlas, NESDIS 14, Washington, D.C.

NOAA (2021) National Centers for Environmental Information, National Oceanic and Atmospheric Administration ETOPO1 dataset: doi:10.7289/V5C8276M, <https://www.ngdc.noaa.gov/mgg/global/global.html>.

Pebesma, E.J., Bivand, R.S. (2005) Classes and methods for spatial data in R. *R News*, 5 (2), <https://cran.r-project.org/doc/Rnews/>.

Pebesma, E. (2018) Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal*, 10 (1), 439-446, <https://doi.org/10.32614/RJ-2018-009>

Phillips, R.A., Silk, J.R.D., Croxall, J.P., Afanasyev, V. & Briggs, D.R. (2004) Accuracy of geolocation estimates for flying seabirds. *Marine Ecology Progress Series*, 266, 265–272.

R Core Team (2021) *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <http://www.r-project.org/>.

R Core Team (2022) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

Robinson, R.A. (2005) *BirdFacts: profiles of birds occurring in Britain & Ireland*. BTO, Thetford <http://www.bto.org/birdfacts> (last accessed on 15 Sep 2022).

- Ross, K.E., Burton, N.H.K., Balmer, D.E., Humphreys, E.M., Austin, G.E., Goddard, B., Schindler-Dite, H. & Rehfish, M.M. (2016) Urban breeding gulls: A review of methods and options for design. BTO Research Report 680.
- Ross, N. (2020). fasterize: Fast Polygon to Raster Conversion. R package version 1.0.3, <https://CRAN.R-project.org/package=fasterize>.
- Ross-Smith, V.H. (2014) The Lesser Black-backed Gull *Larus fuscus* in England: how to resolve a conservation conundrum. *Seabird*, 27.
- Scales, K.L., Miller, P.I., Embling, C.B., Ingram, S.N., Pirotta, E. & Votier, S.C. (2014) Mesoscale fronts as foraging habitats: composite front mapping reveals oceanographic drivers of habitat use for a pelagic seabird. *J. R. Soc. Interface*, 11, 20140679. <http://dx.doi.org/10.1098/rsif.2014.0679>
- Scragg, E.S., Thaxter, C.B., Clewley, G.D., & Burton, N.H.K. (2016) Assessing behaviour of Lesser Black-backed Gulls from the Ribble and Alt Estuaries SPA using GPS tracking devices. BTO Research Report No. 689. Report of work carried out by the British Trust for Ornithology on behalf of Natural England.
- Searle, K.R.; Butler, A.; Mobbs, D.; Trinder, M.; Waggitt, J.; Evans, P.; Daunt, F. 2020. Scottish Waters East Region Regional Sectoral Marine Plan Strategic Ornithology Study: final report (Report No. NEC07184). Report by Natural Environment Research Council (NERC). Report for Marine Scotland Science.
- Shamoun-Baranes, J., Bouten, W., van Loon, E., Meijer, C. & Camphuysen, C.J. (2016) Flap or soar? How a flight generalist responds to its aerial environment. *Philosophical Transactions of the Royal Society of London B*, 371, 20150395. (<http://dx.doi.org/10.1098/rstb.2015.0395>).
- Shamoun-Baranes, J., Burant, J.B., van Loon, E.E., Bouten, W. & Camhuysen, C.J. (2017) Short distance migrants travel as far as long distance migrants in lesser black-backed gulls *Larus fuscus*. *Journal of Avian Biology*, 48, 49-57.
- Sumner (2020) tidync: A Tidy Approach to 'NetCDF' Data Exploration and Extraction. R package version 0.2.4. <https://CRAN.R-project.org/package=tidync>.
- Thaxter, C.B., Ross-Smith, V.H., Clark, J.A., Clark, N.A., Conway, G.J., Marsh, M., Leat, E.H.K. & Burton, N.H.K. (2014) A trial of three harness attachment methods and their suitability for long-term use on Lesser Black-backed Gulls and Great Skuas. *Ringling & Migration*, 29, 65-76.
- Thaxter, C.B., Ross-Smith, V.H., Bouten, W., Masden, E.A., Clark, N.A., Conway, G.J., Barber, L., Clewley, G.D. & Burton, N.H.K. (2018) Dodging the blades: new insights into three-dimensional space use of offshore wind farms by lesser black-backed gulls *Larus fuscus*. *Marine Ecology Progress Series*, 587, 247-253.
- Thaxter, C.B., Ross-Smith, V.H., Bouten, W., Rehfish, M.M., Clark, N.A., Conway, G.J. & Burton, N.H.K. (2015) Seabird–wind farm interactions during the breeding season vary within and between years: A case study of lesser black-backed gull *Larus fuscus* in the UK. *Biological Conservation*, 186, 347-358.
- Thaxter, C. B., Clark, N.A., Ross-Smith, V.H., Conway, G.J., Bouten, W. & Burton, N.H.K. (2017) *Journal of Wildlife Management*, 819, 1098-1109

- Thaxter, C.B., Ross-Smith, V.H., Bouten, W., Clark, N.A., Conway, G.J., Masden, E.A., Clewley, G.D., Barber, L.J. & Burton, N.H.K. (2019) Avian vulnerability to wind farm collision through the year: insights from Lesser Black-backed Gulls (*Larus fuscus*) tracked from multiple breeding colonies. *Journal of Applied Ecology*, DOI: 10.1111/1365-2664.13488.
- Thaxter, B.C., Quinn, L.R., & Humphreys, E.M. (2021) Terrestrial and marine foraging ranges of Lesser Black-backed Gulls and Herring Gulls. *NatureScot Research Report No.1287*.
- Wakefield, E.D., Owen, E., Baer, J., Carroll, M.J., Daunt, F., Dodd, S.G., Green, J.A., Guilford, T., Mavor, R.A., Miller, P.I., Newell, M.A., Newton, S.F., Robertson, G.S., Shoji, A., Soanes, L.M., Votier, S.C., Wanless, S. & Bolton, M. (2017) Breeding density, fine-scale tracking, and large-scale spatial modeling reveal the regional distribution of four seabird species. *Ecological Applications*, 27, 7, 2074-2091.
- Wickham H (2019) stringr: Simple, Consistent Wrappers for Common String Operations. R package version 1.4.0, <<https://CRAN.R-project.org/package=stringr>>.
- Woodward, I.D., Austin, G.E., Balmer, D.E., Boersch-Supan, P., Irwin, C.G., Scott, M.S., Thaxter, C.B., Webb, A., Burton, N.H.K. (2020) Improving Methods for Counting Dispersed Bird Populations: Urban Gull Case Study. Report of work carried out by the British Trust for Ornithology and HiDef Aerial Surveying Ltd on behalf of the Department for Environment, Food and Rural Affairs

7. Appendices

Table S1. Total number of fixes available per colony (split by ‘onshore’ / ‘offshore’) after filtering GPS data to common rates of 5, 30, 60 and 180 minutes combined across years, to assess potential sample sizes for inclusion in the modelling.

Colony / filter rate	5 minute	30 minute	60 minute	180 minutes
BA (Barrow)	3376 / 37	18676 / 189	23939 / 261	16772 / 200
BE (Belfast)			11653 / 29	6162 / 11
CR (Craigleith)		2497 / 1212	1726 / 1073	948 / 847
FI (Fidra)		1846 / 14	4094 / 215	2523 / 107
MA (Isle of May)	285698 / 35743	99123 / 7109	51335 / 3724	27158 / 1621
ON (Orford Ness)	96122 / 24013	169642 / 13058	87221 / 6766	30599 / 2349
RI (Ribble)	27843 / 71	41358 / 35	42046 / 48	29658 / 44
SK (Skokholm)	394955 / 75953	236871 / 19165	124097 / 11076	42951 / 3746
WA (Walney)	690225 / 34107	353998 / 8780	200494 / 5732	79685 / 2119
Total no. onshore	1498219	924011	546605	236451
Total no. offshore	169926	49562	28924	11044

Table S2. Results of the process used to determine the order to add in environmental covariates to the model. The best-fitting model from the first stage of modelling had each environmental covariate added singly to examine the increase in weighted Bhattacharyya affinity ($\overline{BA}$) score. Shown below are each of the 10 environmental covariates considered, ordered by the size of improvement in $\overline{BA}$ score for models with each environmental covariate singly. Models which have an increase in weighted BA score ($\Delta\overline{BA}$ is positive) are highlighted in bold.

	Name of environmental covariate	$\Delta\overline{BA}$	$\overline{BA}$
-	Model without environmental covariates	-	0.370
1	Depth	+ 0.012	0.382
10	Net primary production	+ 0.011	0.381
6	Potential energy anomaly (PEA)	+ 0.011	0.381
9	Thermal front gradient density (TFGD)	+ 0.006	0.376
3	Distance to coast	+ 0.003	0.373
8	Sea surface temperature (SST)	- 0.006	0.364
7	Proportion of time during which water column stratified	-0.007	0.363
2	Seabed slope	-0.043	0.327
4	Proportion of gravel in sediment	-0.105	0.265
5	Ratio of sand:mud in sediment	-0.121	0.249

Examples of extracted environmental covariates

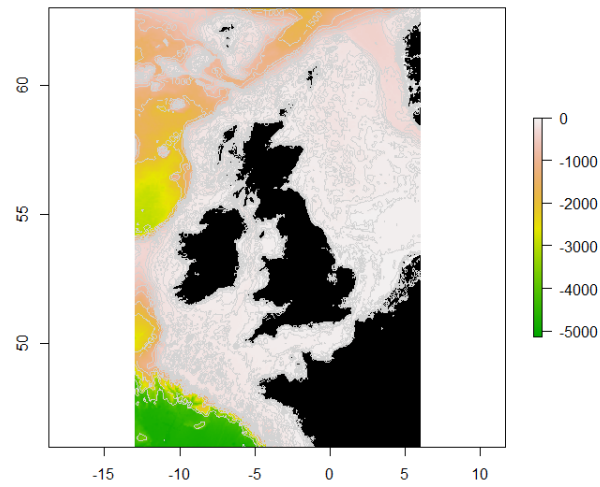


Figure S1. Seabed topography (depth) ETOP01

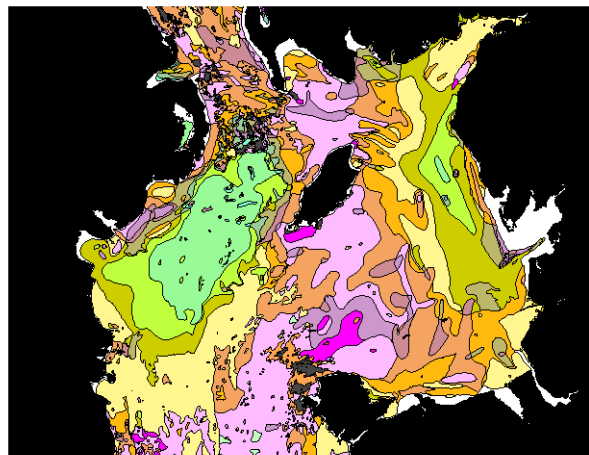
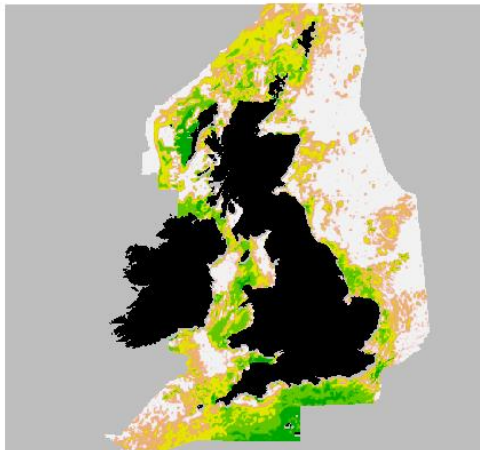


Figure S2. Original sediment shapes with local view in the Irish Sea from the Digimap BGS service

(a) Gravel



(b) Sand:mud ratio

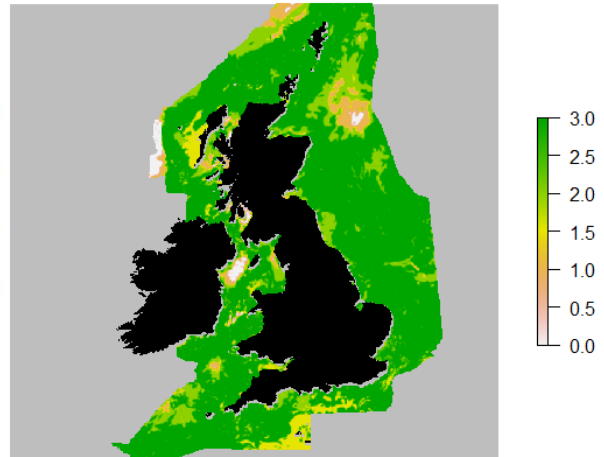


Figure S3. Rescaled and rasterised seabed sediment maps for (a) gravel proportion and (b) sand:mud ratio, using the methods and scaling approach as described in Wakefield et al. (2017)

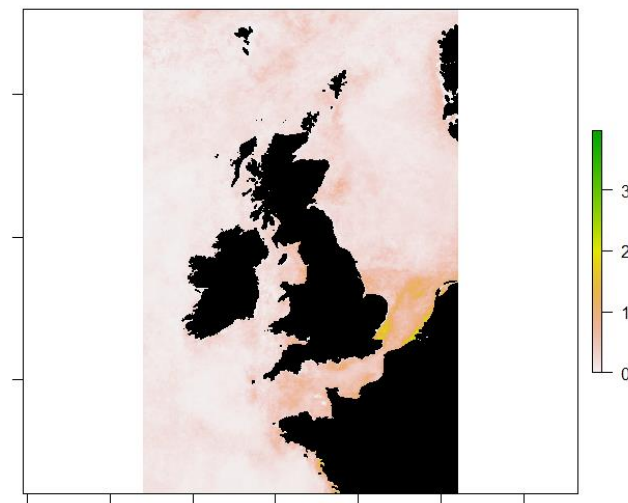
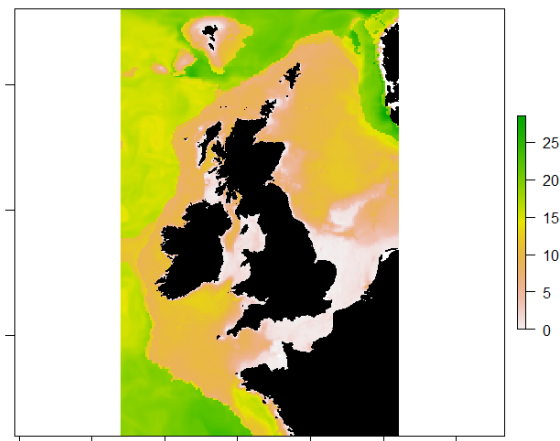


Figure S4. Example Chlorophyll data for August 2020 daily average on the original milligram scale for August 2020 alpha-chlorophyll as a composite image across daily datasets

(a) 23rd June 2020 example one day PEA



(b) Composite PEA $J/m^3 \times 0.5$ for Mar-June 2020

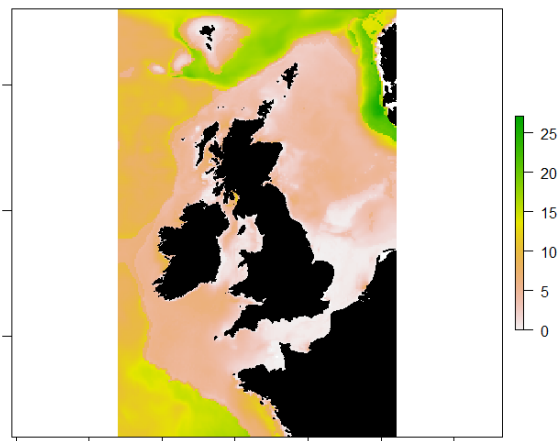
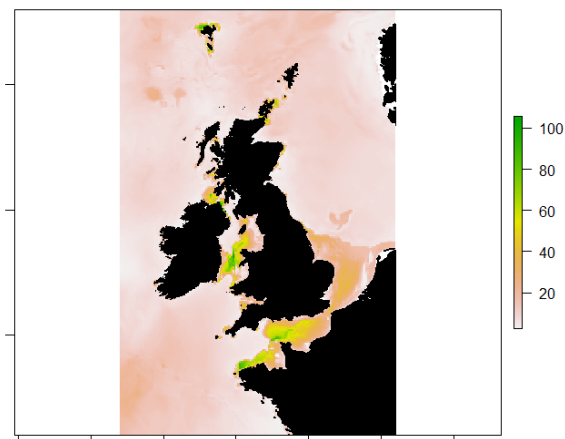


Figure S5. Example of one day slice in 2020, and a further composite for the March-June 2020 period, depending on how the rasters will be amalgamated for analysis

(a) 23rd June 2020 example one day (m)



(b) Composite proportional time stratified

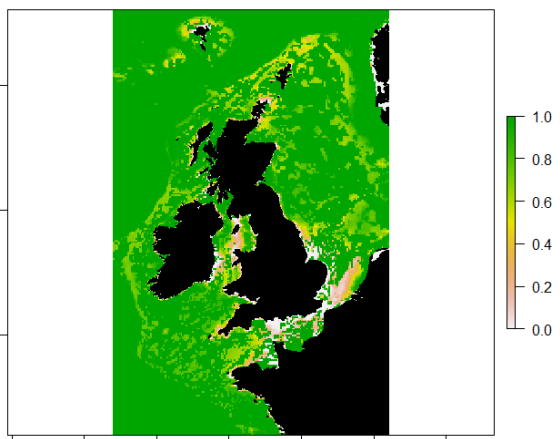


Figure S6. Example of the stratification layer, showing a single day slice of the raster for 23rd June 2020 and then a composite as per Wakefield et al. (2017) for a March-June 2020 proportional of days stratified ($MLD < \text{max water depth}$); for the latter the sea relief dataset was used.

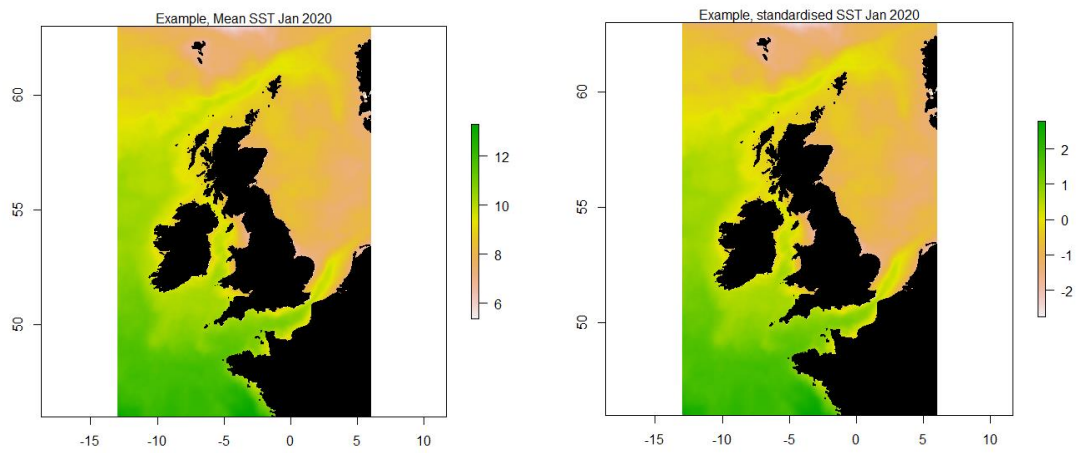


Figure S7. Example SST datasets for 2020, here plotted for January 2020 as a composite image across daily datasets

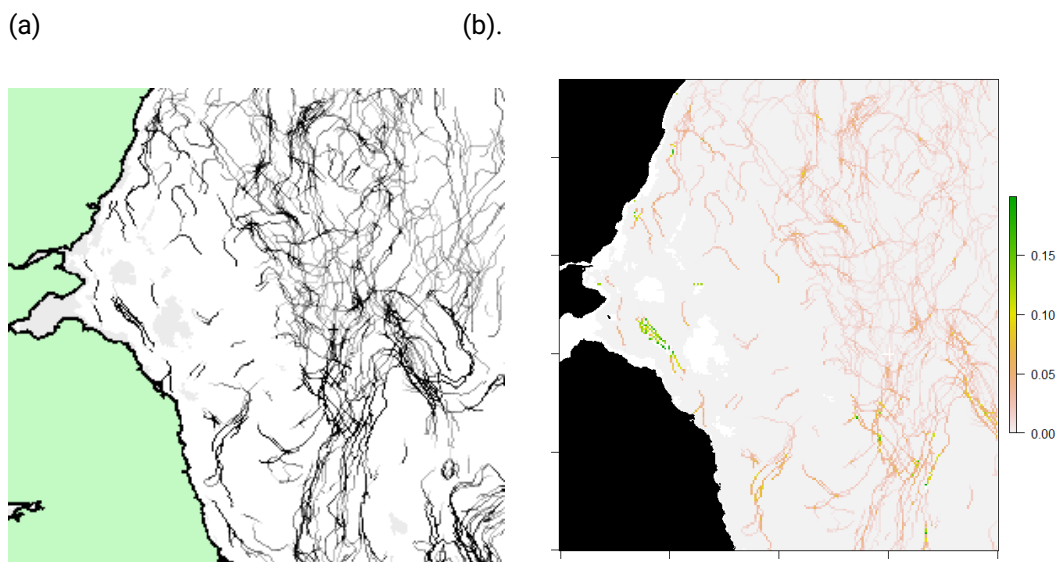


Figure S8. Example front data processing png front data; (a) example single front map for 2020-06-04 to 2020-06-10; (b) translated scale (F_{comp}) following Miller and Christodoulou (2014)

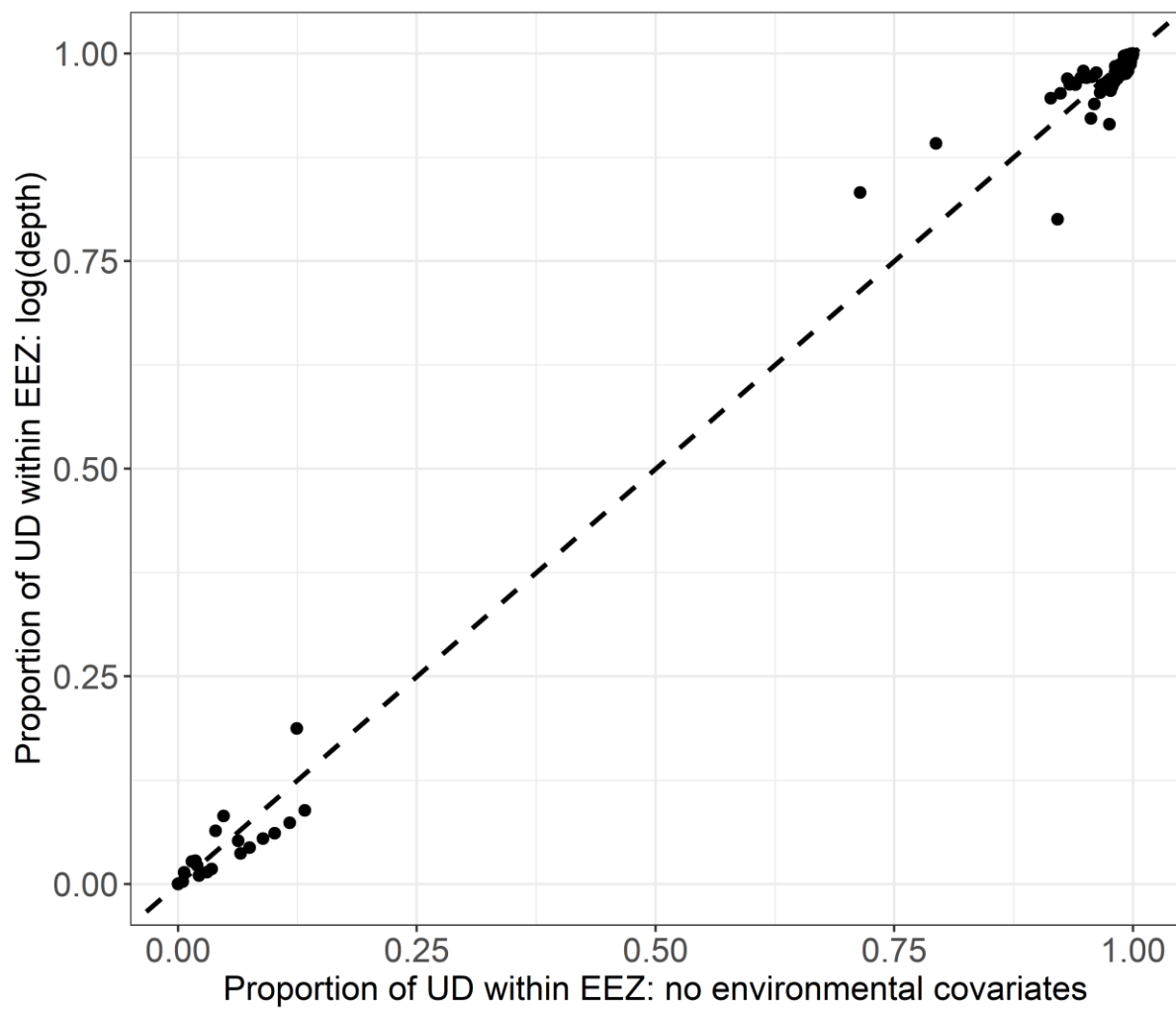


Figure S9. The proportion of the utilisation distribution (UD) predicted to be within the EEZ for each colony (each dot is one of the modelled colonies). The proportion of the UD within the EEZ is shown for both of the models considered: a model with log(depth) as a covariate (y-axis), and a model with no environmental covariates (x-axis).

carbontrust.com

+44 (0) 20 7170 7000

Whilst reasonable steps have been taken to ensure that the information contained within this publication is correct, the authors, the Carbon Trust, its agents, contractors and sub-contractors give no warranty and make no representation as to its accuracy and accept no liability for any errors or omissions. Any trademarks, service marks or logos used in this publication, and copyright in it, are the property of the Carbon Trust. Nothing in this publication shall be construed as granting any licence or right to use or reproduce any of the trademarks, service marks, logos, copyright or any proprietary information in any way without the Carbon Trust's prior written permission. The Carbon Trust enforces infringements of its intellectual property rights to the full extent permitted by law.

The Carbon Trust is a company limited by guarantee and registered in England and Wales under Company number 4190230 with its Registered Office at: Level 5, Arbor, 255 Blackfriars road, London SE1 9AX.

© The Carbon Trust 2024. All rights reserved.

Published in the UK: 2024