Breeding density, fine-scale tracking, and large-scale modeling reveal the regional distribution of four seabird species

EWAN D. WAKEFIELD,1,2 ELLIE OWEN,3 JULIA BAER,4 MATTHEW J. CARROLL,1 FRANCIS DAUNT,5 STEPHEN G. DODD,1 JONATHAN A. GREEN,6 TIM GUILFORD,3 RODDY A. MAVOR,8 PETER I. MILLER,9 MARK A. NEWELL,5 STEPHEN F. NEWTON,4 GAIL S. ROBERTSON,2 AKIKO SHOJI,7 LOUISE M. SOANES,6,10 STEPHEN C. VOTIER,11 SARAH WANLESS,5 AND MARK BOLTON1,12

1RSPB Centre for Conservation Science, The Lodge, Sandy SG19 2DL United Kingdom
2Institute of Biodiversity, Animal Health and Comparative Medicine, University of Glasgow, Graham Kerr Building, Glasgow G12 8QQ United Kingdom
3RSPB Centre for Conservation Science, RSPB Scotland, Etive House, Beechwood Park, Inverness IV2 3BW United Kingdom
4BirdWatch Ireland, 20D Bullford Business Campus, Kilcoole, County Wicklow, Ireland
5Centre for Ecology & Hydrology, Bush Estate, Penicuik Midlothian EH26 0QB United Kingdom
6School of Environmental Sciences, University of Liverpool, Nicholson Building, Brownlow Street, Liverpool L69 3GP United Kingdom
7Oxford Navigation Group, Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS United Kingdom
8Joint Nature Conservation Committee, Inverdee House, Baxter Street, Aberdeen AB11 9QA United Kingdom
9Plymouth Marine Laboratory, Plymouth PL1 3DH United Kingdom
10Life Sciences Department, University of Roehampton, Whitelands College, London SW15 4JZ United Kingdom
11Environment & Sustainability Institute, University of Exeter, Falmouth TR10 9EZ United Kingdom

Abstract. Population-level estimates of species’ distributions can reveal fundamental ecological processes and facilitate conservation. However, these may be difficult to obtain for mobile species, especially colonial central-place foragers (CCPFs; e.g., bats, corvids, social insects), because it is often impractical to determine the provenance of individuals observed beyond breeding sites. Moreover, some CCPFs, especially in the marine realm (e.g., pinnipeds, turtles, and seabirds) are difficult to observe because they range tens to ten thousands of kilometers from their colonies. It is hypothesized that the distribution of CCPFs depends largely on habitat availability and intraspecific competition. Modeling these effects may therefore allow distributions to be estimated from samples of individual spatial usage. Such data can be obtained for an increasing number of species using tracking technology. However, techniques for estimating population-level distributions using the telemetry data are poorly developed. This is of concern because many marine CCPFs, such as seabirds, are threatened by anthropogenic activities. Here, we aim to estimate the distribution at sea of four seabird species, foraging from approximately 5,500 breeding sites in Britain and Ireland. To do so, we GPS-tracked a sample of 230 European Shags Phalacrocorax aristotelis, 464 Black-legged Kittiwakes Rissa tridactyla, 178 Common Murres Uria aalge, and 281 Razorbills Alca torda from 13, 20, 12, and 14 colonies, respectively. Using Poisson point process habitat use models, we show that distribution at sea is dependent on (1) density-dependent competition among sympatric conspecifics (all species) and parapatric conspecifics (Kittiwakes and Murres); (2) habitat accessibility and coastal geometry, such that birds travel further from colonies with limited access to the sea; and (3) regional habitat availability. Using these models, we predict space use by birds from unobserved colonies and thereby map the distribution at sea of each species at both the colony and regional level. Space use by all four species’ British breeding populations is concentrated in the coastal waters of Scotland, highlighting the need for robust conservation measures in this area. The techniques we present are applicable to any CCPF.

Key words: animal tracking; central-place foraging; coloniality; density dependence; habitat use; Poisson point process; species distribution models.

INTRODUCTION

Accurate distribution estimates are key to effective wildlife management yet many colonial central-place foragers (i.e., those that return regularly to a common breeding location or refuge) are difficult to observe because they range so widely. Innovations in telemetry are increasingly making it possible to track these species at the individual level (Wikelski et al. 2007, Hart and Hyrenbach 2010, O’Mara et al. 2014), but both theoretical and analytical advances are needed before unbiased, population-level, distribution estimates can be derived from the resulting data (Aarts et al. 2008, Hebblewhite and Haydon 2010). This is of concern because many
colonial central-place foragers are currently suffering unsustainable declines due to human activities (Mickleburgh et al. 2002, Williams and Osborne 2009, Hamann et al. 2010).

Seabirds are one of the world’s most endangered avian groups (Croxall et al. 2012). This is due to anthropogenic impacts including invasive species, fisheries bycatch, pollution, and direct exploitation. The distribution and size of seabird breeding colonies has been recorded directly in many regions. In contrast, the distribution of birds at sea is generally estimated from visual survey or more recently, tracking data. Systematic surveys from ships or planes began in earnest in the 1970s (Ainley et al. 2012). They provide coarse-scale (1–10 km) Eulerian data (i.e., observations at fixed points in space) but cannot reliably ascribe provenance or, in many cases, life history stage. Hence, colony-level distributions cannot be estimated using this technique. Since the 1990s, it has also been feasible to track the movements of seabirds using bird-borne devices, which are now becoming sufficiently small and cost-effective to obtain statistically robust sample-sizes for a wider range of species (Burger and Shaffer 2008). Devices are usually deployed at colonies so the origin and status of tracked birds are known. However, while GPS tracking is now providing a wealth of fine scale movement data, it is not clear how these aggregations function as groups or independently at different scales (Wakefield et al. 2014). Hence, we might predict birds breeding at colonies with restricted access to the sea travel further than those breeding on isolated islands. Moreover, although it is clear that seabirds breed in hierarchically nested aggregations (i.e., with increasing scale, nests within sub-colonies, within colonies, within islands, archipelagos, etc.) it is not clear how these aggregations function as groups or independently at different scales (Wakefield et al. 2014). Colonies, defined subjectively during censuses, may not therefore correspond to functional units.

Despite these uncertainties, it is clear that while some threats to seabirds are widespread (e.g., climate change) others, such as offshore windfarms, episodic pollution incidents, fisheries bycatch, and the depletion of fish stocks, may be localized, impacting colonies within wider metapopulations unequally (Furness and Tasker 2000, Inchausti and Wimberskirk 2002, Montevecchi et al. 2012). Hence, colony-level distribution estimates may be required in order to target and monitor conservation measures, such as Marine Protected Areas (MPAs) or fisheries closures, effectively (Lascelles et al. 2012, Russell et al. 2013).

Current barriers to estimating colony-level distributions via individual tracking are both logistical and analytical: for most species, it would be impractical to track birds from all colonies. In theory, distribution could be predicted from tracked birds from a sample of colonies by modelling space use as a function of habitat, foraging costs, competition, etc. (Aarts et al. 2008, Wakefield et al. 2009, 2011, Catry et al. 2013). However, statistical techniques for producing unbiased estimates of distribution using tracking data are still in development (Aarts et al. 2008, Patterson et al. 2008, Illian et al. 2012). This is partly because tracking data violate many of the assumptions inherent to conventional parametric models (reviewed by Turchin 1998, Aarts et al. 2008, Wakefield et al. 2009). Repeat observations on individuals (typically 10^2–10^4 locations/individual in seabird studies) tend to be spatiotemporally autocorrelated and the movements of individuals drawn from the same colony may be dependent on one another due to public information transfer and cultural and genetic divergence (Wakefield et al. 2013, Paredes et al. 2015). Furthermore, tracking data record the presence of animals but not their absence (Aarts et al. 2012). In order to account for these attributes, habitat use by tracked animals has been modelled using logistic mixed-effects models (Aarts et al. 2008, Wakefield et al. 2011). This entails the construction of a binary response variable, which comprises animal locations and randomly generated pseudo-absence points. However, the logistic model approximates an inhomogeneous Poisson point
process (IPP) model (Cressie 1993, Aarts et al. 2012), which may be fitted more directly and efficiently by using numerical quadrature to approximate the model’s pseudo-likelihood (Berman and Turner 1992, Baddeley and Turner 2000, Warton and Shepherd 2010; see Methods for details). This approach may therefore be more tractable for GPS tracking data sets, which typically comprise $10^3$–$10^4$ locations per individual. A further substantial problem is that habitat selection may vary between colonies due to differences in the relative availability of prey and habitats among those colonies (e.g., Chivers et al. 2012, Paredes et al. 2012), a phenomenon termed functional response in resource selection (Mysterud and Ims 1998). As such, habitat selection models fitted to data from one site may predict poorly for others (Torres et al. 2015). Matthiopoulos et al. (2011) show that Generalized Functional Response (GFR) models can interpolate usage to unsampled sites more accurately than conventional habitat selection models. GFR models require that usage is sampled under a range of availability regimes allowing the response to environmental covariates to be conditioned on the expected site-level availability of all environmental covariates in the model.

Britain and Ireland are home to internationally important populations of breeding seabirds (Fig. 1). These include 34% of the world’s European Shags *Phalacrocorax aristotelis* (26,600 pairs), 20% of its Razorbills *Alca torda* (93,600 pairs), 13% of its Common Murres *Uria aalge* (708,200 pairs), and 8% of its Black-legged Kittiwakes *Rissa tridactyla* (378,800 pairs) (Mitchell et al. 2004). Our study focuses on these species, referred to hereafter as Shags, Razorbills, Murres, and Kittiwakes. Although the foraging niches of these species partially overlap, they are differentiated along several axes. In Britain and Ireland, all are almost exclusively neritic while breeding, feeding primarily on sandeels (*Ammodytes* spp.) and other small fish and crustaceans (Grémillet et al. 1998, Watanuki et al. 2008, Thaxter et al. 2010). Shags forage either benthically or pelagically (maximum dive depth ~60 m) in coastal waters, relatively close (≤~30 km) to their colonies (Grémillet et al. 1998, Watanuki et al. 2008, Bogdanova et al. 2014). Kittiwakes, Murres, and Razorbills are more wide ranging, foraging tens to hundreds of kilometers from their colonies. Kittiwakes are surface feeders; Murres make relatively long, deep, foraging dives to the pelagic and demersal zones; while Razorbills make more frequent, shallow, dives to the pelagic zone (Thaxter et al. 2010, Linnebjerg et al. 2013). There is some evidence that Kittiwakes from adjacent colonies segregate in space while foraging (Ainley et al. 2003, Paredes et al. 2012) but nothing is known about this phenomenon in the other species.

**Fig. 1.** Breeding distribution and individual movement data used to estimate the distribution at sea of seabirds foraging from UK colonies. Panels a, c, e, and g show numbers of apparently occupied nests recorded during the Seabird 2000 census (Mitchell et al. 2004; red indicates study colonies). Panels b, d, f, and h show tracks of individual birds (colors correspond to colonies). Places mentioned in the text are shown in the upper right panel: CS, Colonsay; DB, Dublin Bay; FH, Flamborough Head; GW, Galway Bay; IL, Islay; IS, Isles of Scilly.
In a recent assessment of conservation status in the UK, Shags and Kittiwakes were reclassified from amber to red due to 62% and 71% declines, respectively, over 25 years (Eaton et al. 2015). In the UK, Murres and Razorbillrs are amber listed due to their restricted range and international importance (Eaton et al. 2015), while internationally Razorbills have recently been reclassified from Least Concern to Globally Near-threatened (BirdLife International 2015). Current declines are thought to be due in part to falls in prey stocks (especially sandeels in the northern North Sea), due to over fishing and climate change (Frederiksen et al. 2007, Cook et al. 2014). Kittiwakes are also regarded to be particularly vulnerable to wind farm developments, which are burgeoning in UK waters (Furness et al. 2013). The diving species face ongoing threats from oil spills (Williams et al. 1995, Votier et al. 2005) and gill nets (Zydelis et al. 2013). Domestic and international legislation and agreements require countries to manage and conserve seabirds (Croxall et al. 2012). Two measures adopted by governments in UK and elsewhere in the European Union that contribute to seabird conservation are the extension of existing colony-based Special Protection Areas (SPAs) for seabirds to adjacent waters that are used for “maintenance activities” (e.g., foraging, courtship, etc.) and secondly, the establishment of marine SPAs around important foraging areas (Garthe et al. 2012, Perrow et al. 2015). However, both marine protected area identification and wider spatial planning at sea are being hampered by a lack of colony-specific distribution estimates (Perrow et al. 2015). In the absence of such information, policy-makers frequently make the unrealistic assumption that seabirds are uniformly distributed out to some threshold distance from their colonies, such as their putative maximum foraging range (Thaxter et al. 2012).

The main aim of our study is to estimate the coarse scale (1–10 km) metapopulation and colony-level utilization distributions of four species of seabirds breeding in Britain and Ireland during the late incubation and early chick-rearing periods. To do so, we tracked birds from a sample of colonies drawn from throughout the geographical, environmental, and colony size range of our study species in Britain and Ireland and modelled their distributions as functions of colony distance, sympatric and parapatric intraspecific competition, coastal morphology, and habitat availability. In so doing, we estimate the distribution of birds from >5,500 breeding sites. Further, we specifically explored the marine distributions of birds from all colonies designated as SPAs, in order to establish the extent, and intensity of usage, of the marine areas required by individuals from these protected breeding locations.

**Materials and Methods**

**Tracking data collection**

We carried out fieldwork at seabird colonies around the coast of Britain and Ireland during May-July 2010–2014, when the study species were either approaching the end of the incubation period or raising small chicks. We stratified sampling effort to reflect the northwards bias in the breeding distribution of seabirds in the region (Mitchell et al. 2004). We caught birds while they attended their nests, either by hand or using a wire noose or crook fitted to a pole, and temporarily attached a modified i-GotU GT-120 (Mobile Action Technology, Taipei, Taiwan) GPS logger to their backs (or rarely, in the case of Kittiwakes, to their tails) with Tesa tape (Tesa SE, Norderstedt, Germany). Total instrument mass was ≤3% body mass for all species, except Kittiwakes, for which it was ≤5% body mass and ≤3% if tail attachment was used. We programmed loggers to record one position every 100 s. Handling time during capture/recapture was ≤6 min. GPS deployments were carried out following the ethical guidelines of the British Trust for Ornithology, under license by Scottish Natural Heritage, Natural England, Natural Resources Wales, the Northern Ireland Environment Agency and the National Parks and Wildlife Service, Ireland.

**Data preparation**

Diving by tagged seabirds can result in short hiatuses in tracking data. To estimate missing locations, and to standardize sampling effort to exactly 100-s intervals, we resampled GPS tracks data by linear interpolation prior to further analysis. Due to the need to deploy and retrieve loggers at the nest, it is normal practice in tracking studies of breeding seabirds to record and analyze bursts of data from one or more complete foraging trip per individual. However, this usually results in individuals being observed for unequal amounts of time because trip duration typically varies widely among individual seabirds. To reduce this bias we subsampled tracking data by randomly selecting a 24-h burst of locations.

### Table 1. Summary of tracking data obtained during the study (see Appendix S2: Table S1 for full details).

<table>
<thead>
<tr>
<th>Species</th>
<th>No. sites</th>
<th>No. birds tracked</th>
<th>No. birds tracked ≥24 h</th>
<th>Median tracking duration (h)</th>
<th>Median trip length (h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shag</td>
<td>13</td>
<td>239</td>
<td>230</td>
<td>75 (55–94)</td>
<td>1.7 (1.0–2.6)</td>
</tr>
<tr>
<td>Kittiwake</td>
<td>20</td>
<td>583</td>
<td>464</td>
<td>42 (25–51)</td>
<td>4.0 (1.6–8.7)</td>
</tr>
<tr>
<td>Murre</td>
<td>12</td>
<td>192</td>
<td>178</td>
<td>54 (45–74)</td>
<td>7.5 (2.0–13.1)</td>
</tr>
<tr>
<td>Razorbill</td>
<td>14</td>
<td>299</td>
<td>281</td>
<td>70 (50–86)</td>
<td>6.3 (1.8–12.6)</td>
</tr>
</tbody>
</table>

*Note: Numbers in parentheses are Interquartile Range IQR.*
from each bird (Table 1). We omitted the small number of individuals that were tracked for <24 h from our analysis. We then selected locations recorded when birds were at sea, categorized according to distance and time from the nest (see Appendix S1 for details). Prior to analysis, we projected all spatial data in Lambert Azimuthal equal area (LAEA) coordinates.

**Modelling approach**

We modeled habitat use as a function of habitat availability, accessibility and proxies of intraspecific competition. In view of the size of the data set (55,000–210,000 locations per species), we fitted IPP models by numerical quadrature (Berman and Turner 1992; Baddeley and Turner 2000, Warpton and Shepherd 2010) rather than approximating them using logistic regression (Aarts et al. 2012). Following Warpton and Shepherd (2010), we modeled the intensity of tracking locations \( l(y_i) \) at the point \( i \) in space as a function of \( n \) explanatory variables:

\[
\log(l_i) = \beta_0 + \sum_{j=1}^{n} c_j \beta_j
\]  

where \( c \) is a vector of covariates and \( \beta = (\beta_0, \beta_1, \ldots, \beta_n) \) the corresponding parameters. The pseudo likelihood of IPP models can be estimated by numerical quadrature (Berman and Turner 1992) as

\[
l_{\text{IPP}}(\beta; y_0, y_0, w) \approx \sum_{i=1}^{m} w_i(s_i \log l_i - \lambda_i)
\]  

where \( y_0 = \{y_{a+1}, \ldots, y_m\} \) are quadrature points (i.e., both data and dummy points), \( w = (w_1, \ldots, w_m) \) is a vector of weights,

\[
s_i = z_i / w_i \quad \text{and} \quad z_i = \begin{cases} 1 & \text{if } y_i \text{ is a data point} \\ 0 & \text{if } y_i \text{ is a dummy point} \end{cases}
\]

The right-hand side of Eq. 2 is equivalent to the likelihood of a weighted log-linear Poisson model, which can readily be estimated using conventional GLM software (Baddeley and Turner 2000). We assigned the centroids of the cells of a regular LAEA grid as dummy points, a quadrature scheme that ensures even distribution across the study area (Warpton and Shepherd 2010). We then assigned weights \( w_i = \alpha n_i \) to each quadrature point, where \( n_i \) is the number of points (data or dummy) in the same cell as the \( i \)th point and \( \alpha \) is the area of that cell (Baddeley and Turner 2000). Note that dummy points are not equivalent to the “pseudo-absence” points used in some case-control models fitted to tracking data (see Aarts et al. [2012]).

In order to account for the highest level of grouping in the tracking data (i.e., breeding colony) we structured models as mixed-effects GLMs

\[
\lambda_{k,i} \sim \text{Poisson}(\mu_{k,i}) \Rightarrow E(\lambda_{k,i}) = \mu_{k,i}
\]

\[
\log(\mu_{k,i}) = \text{offset}(n_k) + \beta_0 + \sum_{j=1}^{m} x_{ij} \beta_j + u_k
\]

where \( \lambda_{k,i} \) is the intensity of locations of birds from the \( k \)th colony and \( u_k \) is a random, colony-level, intercept. The offset term is included to standardize model predictions because the number of birds tracked \( n_k \) varied across colonies. Each bird was tracked for a period of 24 h so the response \( \mu_{k,i} \) is the expected number of tracking locations at sea per bird per day per unit area from the \( k \)th colony. Normalized to sum to unity over all grid cells this approximates the colony-level utilization distribution \( U_D_k \). The inclusion of the colony-level random intercept necessitates a separate set of dummy points for each colony: for the \( k \)th colony, we therefore generated dummy points and weights within the sea area accessible from each colony, which we define as that lying \(<d_{\text{max}}\) from that colony, where \( d_{\text{max}} \) is 1.1 \times the maximum foraging range observed across colonies in our study (Shags 35 km, Kittiwakes 300 km, Murres 340 km, Razorbills 305 km). In the absence of theoretical estimates of the maximum foraging ranges for our study species, we used the maximum observed foraging range. We apply the multiplier 1.1 to ensure that the quadrature grid encompasses the areas bounded by the putative maximum foraging range. Models were fitted using the R package lme4 (Bates et al. 2015).

Warpton and Shepherd (2010) show that the accuracy of the quadrature approximation method described above increases as the ratio of dummy points to data increases. During model development we investigated this effect by fitting single covariate models to data sets generated using quadrature grids of varying resolutions. We found that, within the computationally practicable range of scales, parameter estimates did not converge with increasingly finer scale (Appendix S1: Fig. S1). Hence, following Warpton and Shepherd (2010), we conducted our analysis at the finest resolution practicable. This was 0.5 km for Shags (55,436 tracking locations; 150,557 dummy points) and 2 km for the other species (range 82,741–206,413 tracking locations; 417,578–806,384 dummy points).

**Model selection**

Eq. 3 assumes independence among data (Baddeley and Turner 2000) yet animal tracking locations are repeated measures on individuals and tend to be serially and spatially autocorrelated (Aarts et al. 2008). Hence, the standard errors provided should be treated as relative rather than absolute. The full likelihood of Eq. 3 is unknown, precluding the provision of \( P \) values or model selection using conventional information criteria. Rather, we used \( k \)-folds cross-validation to compare the out-of-sample predictive performance of models based on the similarity between the observed and predicted
utilization distributions (Fewster and Buckland 2001). To do so, we calculated the observed UD of tracked birds from the kth colony (i.e., the proportion of all locations of birds tracked from that colony falling in each cell in the regular grids mentioned above). We then fitted the model under consideration to data from the remaining colonies, predicted the UD of the kth colony from this model and calculated the Bhattacharyya affinity between the observed and predicted UDs

\[
BA_k = \sum_{x,y} \sqrt{UD_{obs,k}(x,y)UD_{pred,k}(x,y)}
\]  

(4)

BA has previously been used in the contexts of UD comparison and model selection (Thacker et al. 1997, Fieberg and Kochanny 2005). It ranges from 0 (no similarity) to 1 (identical UDs). We calculated the weighted mean similarity across colonies

\[
BA = \frac{\sum_{k} n_k BA_k}{\sum_{k} n_k}
\]  

(5)

where \( n_k \) is the number of birds tracked from the kth colony. The contribution to \( BA \) of colonies from which larger numbers of birds were tracked is upweighted because the UDs of colonies with small samples of tracked birds are likely to underestimate the area used by the entire colony (Soanes et al. 2013, Bogdanova et al. 2014).

In order to estimate space use from all colonies in the study area, we aimed to select the best model from a field of biologically plausible alternatives. Previous studies suggest that seabird space use may depend on numerous covariates, including colony distance, density-dependent competition, and habitat (Wakefield et al. 2009, 2011). The number of plausible alternative models is therefore large. This, combined with the time taken for models to converge, precluded backward model selection. Rather, we built usage models using a stepwise forward selection procedure, adding candidate explanatory covariates to the intercept-only model in order of their expected effects sizes. We retained covariates if \( \Delta BA \) was positive, selecting the most parsimonious model if \( \Delta BA \) was tied. In order to compare effect sizes using standardized partial regression coefficients we standardized covariates prior to analysis (Schielzeth 2010).

**Candidate explanatory covariates**

In the absence of other factors, central-place foraging theory suggests that breeding seabirds should seek prey as close to their nest sites as possible (Orians and Pearson 1979). First, therefore, we added distance to colony \( d \) to the model, with the expectation that usage would decline with distance (Dukas and Edelstein-Keshet 1998, Matthiopoulos 2003). Our study species generally avoid crossing extensive land barriers when commuting (Fig. 1) so we defined \( d_{k,i} \) to be the minimum distance by sea between the kth breeding site and the ith location, which we calculated on a 0.5 km (Shags) or 1 km (Kittiwakes, Murres, and Razorbills) LAEA grid using the R package gdistance (van Etten 2012, Wakefield et al. 2013). Space use by breeding seabirds is further modulated by density-dependent competition among sympatrically breeding conspecifics (Wakefield et al. 2013, Iovani et al. 2015). Given that competition is proportional to the density of animals we next considered whether the area of sea available to birds from each breeding site, which varies with coastal geometry, affects usage (Birkhead and Furness 1985). We hypothesize that density-dependent competition would be higher among birds foraging from colonies with restricted access to the open sea such that they would forage further from their colonies than birds from colonies surrounded by open water. To model this effect, we considered the addition of \( A_{k,i} \), the cumulative area at the ith location relative to the kth breeding site, to our models, where

\[
A_{k,i} = \sum_{x} \delta_{x,i} \delta_{k,x} = \begin{cases} 1 & \text{if } d_{k,x} \leq d_{k,i} \\ 0 & \text{otherwise} \end{cases}
\]  

(6)

and \( \delta_i \) is the area of the ith cell of the LAEA grids mentioned above. Exploratory analysis indicated that log-transforming \( A_k \) reduced colinearity with \( d_k \), improving model stability.

We next considered the number of sympatric breeders, the other determinant of density at sea. We extracted numbers of apparently occupied nests (AON) recorded during the most recent complete census of seabird colonies in Britain and Ireland (Seabird 2000, carried out between 1998 and 2002; see Mitchell et al. [2004] for methods) from the Seabird Monitoring Programme (SMP) Database. Defining seabird colonies objectively can be problematic because the degree to which breeding seabird nests are clustered in space varies with scale (Wakefield et al. 2014). During the Seabird 2000 census, AON were recorded by “subsite” (for clarity, simply referred to as “sites” hereafter). These Mitchell et al. (2004) nominally defined as segments of coastline \( <1 \) km long, containing clusters of breeding seabirds. However, for practical reasons fieldworkers were allowed some scope to deviate from this definition. In practice, sites sometimes therefore comprise isolated islands or segments of coastline \( >1 \) km long. In the latter cases, we reassigned sites by splitting the coastline into the minimum possible number of segments \( \leq 1 \) km long, dividing AON equally between each. During model selection we considered several potential proxies of competition from sympatric breeders. First, the number, \( N \), of conspecific AONs at the home site. Second, because arbitrary census divisions may not correspond to ecologically functional units (Wakefield et al. 2014) we considered proxies that include conspecifics breeding in the vicinity of the home
site of tracked birds. These were the inverse-distance weighted number of breeding conspecifics

$$\theta = \sum_{A\in k} \frac{N_k}{d_{h,k} + 1}$$  

(7)

where $N_k$ is the number of conspecific AON at the $k$th site of the set of all breeding sites (including the home site) within the species’ maximum foraging range, and $d_{h,k}$ is the distance from the home breeding site to the $k$th breeding site. Finally, based on exploratory analyses, we also considered the square-roots of these indices, as well as Eq. 3 the inverse-distance weighted square-root number of conspecific breeders

$$\theta' = \sum_{A\in k} \frac{\sqrt{N_k}}{d_{h,k} + 1}$$  

(8)

We considered each of the indices of sympatric competition as a main effect and interaction with $A$, selecting that which resulted in the best improvement in model performance (step 3).

In addition to sympatric competition, breeding seabirds may be subject to competition from conspecifics breeding at neighboring colonies (Furness and Birkhead 1984; hereafter, parapatric competition [Wakefield et al. 2011]). As with sympatric competition, this is thought to be density dependent (Wakefield et al. 2013). Our expectation is therefore that birds avoid locations at which the null density of conspecifics from other colonies is high (Wakefield et al. 2011, 2013, Catry et al. 2013). In some systems, this leads to striking patterns of among-colony spatial segregation (Masello et al. 2010, Wakefield et al. 2013). It has been hypothesized that these are mediated by individual-level information transfer and cultural divergence during colony growth (Wakefield et al. 2013). Current uncertainty about these mechanisms makes this phenomenon difficult to model satisfactorily but as a first approximation we considered whether birds avoided areas in which the null density of conspecifics from other colonies is greater than that from their own (Catry et al. 2013). Taking the best models from previous steps (hereafter models I–IV for Shags, Kittiwakes, Murres, and Razorbills, respectively), we predicted $\rho_{h,i}$, the ratio of the expected intensity of locations $x_{h,i}$ from the focal breeding site $h$ to the sum of those from all other sites in the region

$$\rho_{h,i} = \frac{\lambda_{h,i}N_k}{\sum_{k\neq h} \lambda_{k,i}N_k}$$  

(9)

We then determined whether adding this covariate to the usage models improved their performance.

We next considered whether the addition of environmental indices describing habitat improved model performance (Wakefield et al. 2009). We identified candidate biophysical covariates meeting two criteria: first, that

data coverage was sufficient to allow seabird distributions to be estimated throughout the waters of Britain and Ireland, and second, that previous studies had established links between the covariate (or the phenomenon it quantifies) and the foraging behavior or distribution of the study species or their prey. As noted above, each model level requires a separate set of quadrature points. Hence, although we considered both static and dynamic covariates, we averaged monthly mean dynamic covariates over the study period (May–July 2010–2014; Appendix S1: Fig. S2) to maintain the number of data, and thereby computing time, within tractable limits. We then determined the value of environmental covariates at each quadrature point. We considered (1) depth (ETOPO2 Global Relief 2v2, provided by the U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Geophysical Data Center, 2006); (2) seabed slope, calculated from the latter in ArcGIS 10 (ArcGIS10 manufactured by ESRI based in Redlands, California, USA); (3) minimum distance to the coast, calculated in ArcGIS 10; (4) the proportion of gravel; and (5) the ratio of sand to mud in seabed sediments, derived from British Geological Survey 1:250,000 maps [(available online; see Appendix S1)];[15] (6) the potential energy anomaly (PEA), which quantifies the intensity of thermohaline stratification; and (7) the proportion of time during which the water column was stratified, both estimated using UK Met Office FOAM AMM reanalysis data [(available online, see Appendix S1)];[15] (8) AVHRR sea surface temperature (SST), supplied by the Natural Environment Research Council Earth Observation Data Acquisition and Analysis Service (NEODAAS); (9) standardized sea surface temperature (sSST), calculated on a monthly basis by subtracting the mean SST in the study area and dividing by its standard deviation, which is an alternative index of stratification (Wakefield et al. 2015); (10) thermal front gradient density (TFGD), estimated following (Scales et al. 2014) using AVHRR SST to provide an index of the mean intensity and persistence of thermal fronts (Miller and Christodoulou 2014); and (11) net primary production (NPP) estimated and supplied by NEODAAS using MODIS chlorophyll and photosynthetically available radiation data. For further details of candidate covariates and our rationale for their consideration, see Appendix S1: Table S2 and reviews by Hunt (1997), Mann and Lazier (2006), and Wakefield et al. (2009). In brief, the phenomena described by these covariates may affect our study species’ distributions either by modulating lower trophic level production (depth, seabed slope, indices of stratification, SST, TFGD, NPP [Begg and Reid 1997, Mann and Lazier 2006, Scott et al. 2010, Carroll et al. 2015]); by physically aggregating prey (indices of stratification, TFGD, and indirectly SST and depth; Lefèvre 1986, Begg and Reid 1997, Mann and Lazier 2006, Embling et al. 2012); or

14 http://digimap.edina.ac.uk
15 http://marine.copernicus.eu/
due to the habitat preference of prey species, especially Ammodopsidae and Clupeidae (depth, coast distance, sediment, indices of stratification [Whitehead 1986, Holland et al. 2005, van der Kooij et al. 2008]).

In order to establish in what order to add environmental covariates to models, we first determined the improvement in performance afforded by adding each singly to the best model resulting from the previous steps. Based on previous work (Wakefield et al. 2011, 2015) and exploratory analyses we considered log and square-root transformations of some covariates (Appendix S1: Table S2). In order to model potential variation in habitat selection in response to among-colony variability in habitat availability, we also considered interactions between each covariate and its expected value at each colony. This we define as the covariate’s mean (hereafter denoted by an overbar) in waters accessible from that colony (i.e., the sea area within $d_{\text{max}}$). This partially implements the GFR model proposed by Matthisopoulos et al. (2011). The full GFR model, in which variables interact not only with their own colony-level expectations but those of all other environmental covariates, proved computationally unfeasible with our data set (see Appendix S1). We ranked environmental covariates in order of $\Delta BA$ afforded by the addition of each covariate (transformed or otherwise) and its GFR equivalent. We then added these terms sequentially to the model, retaining them if $\Delta BA$ was positive (step 5). If two covariates were considered proxies of the same phenomenon (e.g., stratification) or were otherwise colinear, we considered only that ranked highest. Finally, because relationships between space use and environmental covariates may be nonlinear, we also considered their second degree polynomials, retaining them if their addition resulted in an increase in $\Delta BA$ (step 6).

Estimating usage

For each species, we estimated $\lambda$ and thereby the UD for birds from each Seabird 2000 site $s$ using the fixed-effects part of the best models (hereafter models V–VIII for Shags, Kittiwakes, Murres, and Razorbills, respectively). We then calculated the population-level UD across the study area

$$UD_{P,s} = \sum_{A} UD_{A,s} N_s$$

where $N_s$ is the number of AON at the $s$th site. Notwithstanding the comments on standard errors above, we quantified spatial variation in the relative uncertainty of our model estimates by plotting the coefficient of variation (CV) of $UD_{P,s}$ which we calculated using parametric resampling adapted from Bolker (2008) and Lande et al. (2003). Assuming that the sampling distribution of $\beta$ is multivariate normal, we generated 100 random sets of fixed-effects parameters for each model, predicted the $UD_{P,s}$ using each set of parameters, and then calculated its CV.

In order to illustrate how one might use these UDs to identify marine areas whose statutory protection would facilitate the functional protection of the existing suite of colony SPAs, following Eq. 10, we also calculated the mean UD of birds breeding at sites within each UK SPA. We identified breeding sites falling within existing colony SPAs using boundaries downloaded from the Joint Nature Conservation Committee (available online). For each SPA, we then determined polygons encompassing the core 50%, 75%, and 90% of estimated usage as well as the maximum curvature boundaries (MCBs, see Appendix S1). While MCBs have no ecological basis (Perrow et al. 2015), it has been suggested that they balance the proportion of a population protected against the extent of the protected area and have been used by statutory bodies to define boundaries for delimiting avian marine protected areas in the UK (O’Brien et al. 2012). We then overlaid percentage UDs and MCBs of all species in order to estimate the overall extent of sea area thus encompassed.

RESULTS

Seabird tracking

We tracked 1,313 birds from 29 colonies for a median of 2–3 d/bird (Fig. 1, Table 1; Appendix S2: Table S1). Following resampling to standardize the observation period to 24 h/bird, data from 80% of Kittiwakes and 93–96% of the remaining species were retained for analysis, totaling 1,153 individuals. Full data are available for download from the BirdLife Seabird Tracking Database (available online). The duration of deployment was set by the need to recapture birds before tags became detached from feathers. Recapture was attempted after 24 h (Kittiwake, where the mantle feathers are relatively weak) to 48 h (other species). Median foraging trip length was <24 h in all species (Table 1) so the 24-h observation window generally spanned >½ trip/individual. Differences in foraging ranges were marked among species (Fig. 1): Shags remained relatively close to their nest sites (median 3.4 km, IQR 1.6–7.5), whereas Kittiwakes (11.9 km, IQR 4.2–30.9), Murres (10.5 km, IQR 3.2–19.1), and Razorbills (13.2 km, IQR 5.1–26.2) travelled further from their colonies during foraging trips.

Explanatory covariates

The addition of distance to colony $d$ improved the performance of usage models for all species (Appendix S2: Tables S2 and S3) and its effect, which was negative, was relatively large (Table 2). The addition of $A$ (the cumulative area at $d$), interacting with indices of sympatric...
Table 2. Summary of fixed effects in inhomogeneous Poisson point process models of the density of seabird tracking locations as functions of colony distance, coastal geometry, intra-specific competition, and habitat.

<table>
<thead>
<tr>
<th>Model and covariate</th>
<th>Estimate</th>
<th>SE ‡</th>
<th>z</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>VI. Shag</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>−6.092</td>
<td>0.240</td>
<td>−25.43</td>
</tr>
<tr>
<td>(d)</td>
<td>−1.254</td>
<td>0.018</td>
<td>−71.52</td>
</tr>
<tr>
<td>(\log(A))</td>
<td>−1.239</td>
<td>0.010</td>
<td>−128.17</td>
</tr>
<tr>
<td>(\theta^*)</td>
<td>0.353</td>
<td>0.250</td>
<td>1.41</td>
</tr>
<tr>
<td>Gravel</td>
<td>0.512</td>
<td>0.012</td>
<td>41.92</td>
</tr>
<tr>
<td>(\sqrt{\text{PEA}})</td>
<td>−1.613</td>
<td>0.028</td>
<td>−58.64</td>
</tr>
<tr>
<td>NPP</td>
<td>0.048</td>
<td>0.011</td>
<td>4.22</td>
</tr>
<tr>
<td>Coast distance</td>
<td>−1.187</td>
<td>0.034</td>
<td>−35.28</td>
</tr>
<tr>
<td>SST</td>
<td>0.797</td>
<td>0.046</td>
<td>17.37</td>
</tr>
<tr>
<td>SST^2</td>
<td>0.474</td>
<td>0.026</td>
<td>18.14</td>
</tr>
<tr>
<td>(\log(A) \times \theta^*)</td>
<td>0.110</td>
<td>0.005</td>
<td>23.02</td>
</tr>
<tr>
<td>gravel \times gravel</td>
<td>−0.627</td>
<td>0.020</td>
<td>−30.78</td>
</tr>
<tr>
<td><strong>VI. Kittiwake</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>−6.375</td>
<td>0.175</td>
<td>−36.39</td>
</tr>
<tr>
<td>(d)</td>
<td>−1.338</td>
<td>0.010</td>
<td>−140.65</td>
</tr>
<tr>
<td>(\log(A))</td>
<td>−0.486</td>
<td>0.005</td>
<td>−91.12</td>
</tr>
<tr>
<td>(\theta)</td>
<td>−0.388</td>
<td>0.189</td>
<td>−2.06</td>
</tr>
<tr>
<td>(\log(\rho))</td>
<td>1.669</td>
<td>0.014</td>
<td>118.75</td>
</tr>
<tr>
<td>(\log(\text{seabed slope}))</td>
<td>−0.019</td>
<td>0.005</td>
<td>−4.15</td>
</tr>
<tr>
<td>(\log(\text{seabed slope}))^2</td>
<td>−1.381</td>
<td>0.261</td>
<td>−5.29</td>
</tr>
<tr>
<td>(\text{sSST})</td>
<td>−1.006</td>
<td>0.007</td>
<td>−143.32</td>
</tr>
<tr>
<td>stratification</td>
<td>0.033</td>
<td>0.004</td>
<td>9.21</td>
</tr>
<tr>
<td>stratification</td>
<td>0.969</td>
<td>0.308</td>
<td>3.15</td>
</tr>
<tr>
<td>(\log(A) \times \theta)</td>
<td>0.167</td>
<td>0.004</td>
<td>46.31</td>
</tr>
<tr>
<td>(\log(\text{seabed slope}) \times \log(\text{seabed slope}))</td>
<td>0.979</td>
<td>0.009</td>
<td>104.99</td>
</tr>
<tr>
<td>stratification \times stratification</td>
<td>0.942</td>
<td>0.011</td>
<td>87.33</td>
</tr>
<tr>
<td><strong>VI. Murre</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>−7.294</td>
<td>0.177</td>
<td>−41.19</td>
</tr>
<tr>
<td>(d)</td>
<td>−1.627</td>
<td>0.028</td>
<td>−57.56</td>
</tr>
<tr>
<td>(\log(A))</td>
<td>−0.862</td>
<td>0.007</td>
<td>−124.54</td>
</tr>
<tr>
<td>(\sqrt{\theta})</td>
<td>0.206</td>
<td>0.171</td>
<td>1.21</td>
</tr>
<tr>
<td>(\log(\rho))</td>
<td>0.929</td>
<td>0.029</td>
<td>32.07</td>
</tr>
<tr>
<td>Gravel</td>
<td>−0.223</td>
<td>0.005</td>
<td>−46.71</td>
</tr>
<tr>
<td>(\sqrt{\text{sand:mud}})</td>
<td>−0.184</td>
<td>0.011</td>
<td>−16.42</td>
</tr>
<tr>
<td>((\sqrt{\text{sand:mud}})^2)</td>
<td>−0.196</td>
<td>0.010</td>
<td>−18.80</td>
</tr>
<tr>
<td>(\sqrt{\text{sand:mud}})^2</td>
<td>−2.037</td>
<td>0.543</td>
<td>−3.75</td>
</tr>
<tr>
<td>TFGD</td>
<td>0.331</td>
<td>0.004</td>
<td>77.63</td>
</tr>
<tr>
<td>Coast distance</td>
<td>−1.709</td>
<td>0.032</td>
<td>−53.81</td>
</tr>
<tr>
<td>(\text{coast distance})</td>
<td>3.098</td>
<td>0.370</td>
<td>8.38</td>
</tr>
<tr>
<td>(\log(A) \times \sqrt{\theta})</td>
<td>0.273</td>
<td>0.005</td>
<td>54.23</td>
</tr>
<tr>
<td>(\sqrt{\text{sand:mud}} \times \sqrt{\text{sand:mud}})</td>
<td>−0.481</td>
<td>0.034</td>
<td>−14.16</td>
</tr>
<tr>
<td>(\text{coast distance} \times \text{coast distance})</td>
<td>1.760</td>
<td>0.057</td>
<td>30.68</td>
</tr>
<tr>
<td><strong>VII. Razorbill</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>−4.623</td>
<td>0.105</td>
<td>−43.84</td>
</tr>
<tr>
<td>(d)</td>
<td>−1.066</td>
<td>0.009</td>
<td>−119.85</td>
</tr>
<tr>
<td>(\log(A))</td>
<td>−1.106</td>
<td>0.004</td>
<td>−255.08</td>
</tr>
<tr>
<td>(\sqrt{N})</td>
<td>0.552</td>
<td>0.106</td>
<td>5.23</td>
</tr>
<tr>
<td>SST</td>
<td>−0.083</td>
<td>0.008</td>
<td>−10.60</td>
</tr>
<tr>
<td>SST</td>
<td>0.336</td>
<td>0.130</td>
<td>2.58</td>
</tr>
<tr>
<td>sand : mud</td>
<td>−0.290</td>
<td>0.006</td>
<td>−47.53</td>
</tr>
<tr>
<td>((\text{sand: mud})^2)</td>
<td>−0.266</td>
<td>0.005</td>
<td>−53.46</td>
</tr>
<tr>
<td>(\log(\text{seabed slope}))</td>
<td>0.027</td>
<td>0.005</td>
<td>5.30</td>
</tr>
<tr>
<td>(\log(\text{seabed slope}))</td>
<td>−0.306</td>
<td>0.221</td>
<td>−1.38</td>
</tr>
</tbody>
</table>
competition, further improved model performance (Appendix S2: Table S3). In the case of Razorbills, the square-root of the number of breeding pairs in the home site was the best index of sympatric competition. For the other species, competition indices based on the summation of numbers of breeders inversely weighted by distance from the focal breeding site best improved model performance (Appendix S2: Table S2). Models I–IV suggest that, in all species, the rate of decline in usage with $A$ lessens with increasing sympatric competition (Table 2; Appendix S2: Tables S2 and S3, Fig. S1). The inclusion of relative parapatric competition improved the performance of Kittiwake and Murre usage models but not those of Shags and Razorbills (Appendix S2: Table S2). The former species tended to avoid areas in which the potential density of conspecifics from other colonies was higher than that from their own colony. The addition of environmental covariates improved the performance of all models (cf. Fig. 2; Appendix S2: Fig. S3) and conditioning some but not all covariates on their regional means improved performance further (Appendix S2: Table S4). Cross-validation shows that the final models for Shags, Kittiwakes, and Murres all performed similarly well ($BA/C6SD = 0.52/0.13, 0.53/0.13,$ and $0.53/0.22$), respectively but the performance of the Razorbill model was somewhat poorer ($BA/C6SD = 0.34/0.11$). Spatial plots confirm our expectation that the similarity between observed and predicted utilization distributions was greatest for colonies from where more birds were tracked (Appendix S2: Fig. S3).

The effects of many environmental covariates were comparable in magnitude to those of colony distance, cumulative area, and competition (Table 2). Taking the environmental covariates retained during model selection in order of their effect sizes, these suggest that Shags tend to use relatively mixed waters (i.e., low PEA) close to the coast. In areas where gravel is scarce, they use relatively gravelly substrates but this is reversed in more gravelly areas (Appendix S2: Fig. S5). Shags’ usage with respect to SST was quadratic, with a tendency to visit areas where SST was either warmer or cooler than the average (Table 2; Appendix S2: Fig. S5). Shags also manifested a weak preference for areas of high NPP.

Usage by Kittiwakes with respect to seabed slope and stratification was complex: in areas where the mean seabed slope was low, they tended to avoid steep bathymetric relief but this preference was reversed somewhat in areas where the mean slope was high (Appendix S2: Fig. S5). Similarly, in areas where the mean occurrence of stratification was low, Kittiwakes avoided stratified

### Table 2. (Continued)

<table>
<thead>
<tr>
<th>Model and covariate†</th>
<th>Estimate</th>
<th>SE‡</th>
<th>$z$</th>
</tr>
</thead>
<tbody>
<tr>
<td>log($A$) × $\sqrt{N}$</td>
<td>0.331</td>
<td>0.003</td>
<td>123.37</td>
</tr>
<tr>
<td>$SST \times SST$</td>
<td>$-0.882$</td>
<td>0.010</td>
<td>$-90.40$</td>
</tr>
<tr>
<td>log(seabed slope) × log(seabed slope)</td>
<td>$-0.525$</td>
<td>0.015</td>
<td>$-34.22$</td>
</tr>
</tbody>
</table>

**Notes:** Numbers in parentheses after model name are the numbers of sites and birds.

†Covariates standardized prior to model fitting; $d$, distance by sea from the colony; $A$, cumulative area at distance $d$; $h$, inverse-distance-weighted number of conspecific breeders; $h_0$, inverse-distance-weighted square-root number of conspecific breeders; $N$, number of conspecific breeders at the home site; $p$, density of birds from the home site relative to those from all other sites; NPP, net primary production; PEA, mean potential energy anomaly; SST, mean sea surface temperature; sSST, mean standardized SST; TFGD, thermal front gradient density. Overbars indicate the mean of the covariate in water accessible from each colony.

‡Relative standard errors.
waters, whereas in more frequently stratified areas, they tended to avoid mixed waters.

In areas with low regional mean coastal distance (i.e., archipelagos) Murres used areas close to the coast, whereas in areas with less complex coastlines they tended to forage further from land (Table 2; Appendix S2: Fig. S5). In regions with a relatively high proportion of sand in the substrate, Murres preferred sandy areas but this preference reversed in less sandy regions. Murres also showed a weak preference for frontal regions and substrates containing a relatively low proportion of gravel (Table 2). Razorbills used areas with higher SSTs in regions with relatively cool surface waters, whereas in warmer regions the opposite was true (Appendix S2: Fig. S5). In regions with relatively low seabed relief they tended to select areas with steep relief and vice versa. Razorbills’ habitat preference with respect to the sand: mud ratio of the substrate was quadratic, peaking just below intermediate values (Appendix S2: Fig. S5).

Estimated population-level distributions

Raster files of space use during late incubation and early chick-rearing from all of the region’s colonies estimated using models V–VIII are available for download from the Data Archive for Marine Species and Habitats (DASSH; available online).18 Composite usage maps predict that breeding Shags, Kittiwakes, Murres, and Razorbills forage mainly within 100 km of the coast of Scotland, primarily to the north and east of the mainland in the North Sea, and around the Northern Isles (Fig. 3a; Appendix S2: Fig. S7). For all species, 90% of the UK regional population’s UDs also included waters in the southern North Sea; Dublin Bay and the North Channel of the Irish Sea; as well as waters surrounding Islay; the northern Minch; and isolated islands north-west of Scotland (Appendix S2: Fig. S7). The estimated distributions of Shags, which is the least wide-ranging of the study species, largely reflects that of its colonies (cf. Figs. 1 and 2). In contrast, that of Kittiwakes is more pelagic, with activity more patchily distributed offshore (Fig. 2). In addition to core areas mentioned above, usage hotspots included a large area southeast of Flamborough Head and the northern Norfolk Banks; the central Irish Sea; and Galway Bay, west of Ireland. Of the two Auks, our models suggest that Murres forage closer on average to their colonies (Fig. 2), outnumbering Razorbills in many coastal areas and in the vicinity of the Celtic Sea front. In contrast, Razorbills predominate in the North Channel and much of the Minch (Appendix S2: Fig. S8).

DISCUSSION

Several recent studies have assimilated tracking data from multiple colonies in order to map and understand seabirds distributions (BirdLife International 2004, Block et al. 2011, Wakefield et al. 2011, 2013, Ramos et al. 2013). However, this is the first to model how colony-level distributions vary due to the combined effects of sympatric and parapatric conspecific interactions, coastal geomorphology, and regional habitat availability. By tracking and modelling the space use of Shags, Kittiwakes, Murres, and Razorbills from a sample of colonies around Britain and Ireland, we estimated the coarse-scale (tens of kilometers) distribution of these species at sea from all of colonies in the region. Moreover, by combining these results, we were able to map the at-sea distribution of each species’ breeding population across a study area extending over ~1.5 million km² (Fig. 2). Until recently, it was only practicable to attempt to estimate the distributions of seabirds over such wide areas at comparable resolutions by surveying birds from boats or planes (Stone et al. 1995, Bradbury et al. 2014). However, these methods generally fail to discriminate among birds from different colonies or life history stages (e.g., breeders vs. non-breeders). Our results therefore provide unprecedented insights into marine distributions of breeding seabirds.

We modelled the occurrence of tracking locations as an inhomogeneous Poisson point (IPP) process (Cressie 1993), which is a computationally efficient and, it has been argued, natural method of treating presence-only data (Warton and Shepherd 2010, Aarts et al. 2012, Renner et al. 2015). We discuss our approach in more detail in Appendices S1 and S3. However, it is pertinent to highlight two caveats to our results. First, due to the

---

18 www.dassh.ac.uk
large volume of data involved in our analysis, only relatively simple models were computationally tractable and therefore not all correlation structures inherent to the data (e.g., serial autocorrelation within individuals; Aarts et al. 2008) were modelled. Hence, although we presume that our parameter and usage estimates are unbiased their associated uncertainty is likely to be underestimated. Second, the likelihood estimation technique we used is approximate (Berman and Turner 1992). We therefore opted to select among models by k-folds cross validation, rather than using penalized information criteria, such as AIC. The k-folds cross validation technique is robust to over-fitting when the number of data is large, and the field of candidate models relatively small (Arlot 2010). However, our models are optimized for prediction, rather than parsimony, so the biological inferences drawn from them here are tentative.

Distribution with respect to colony distance and competition

Space use by all four study species declined with distance from the colony (Table 2), supporting the hypothesis that central-place foragers minimize distance-dependent travel costs (Orians and Pearson 1979). Our results also support the hypothesis that colonial central-place foragers seek to minimize density-dependent intraspecific competition (Ashmole 1963, Lewis et al. 2001, Wakefield et al. 2013): in all species, the rate of decline in usage with cumulative area at distance decreased as the number of sympatrically breeding conspecifics increased (Appendix S2: Table S3, Fig. S1). Although this echoes the observation that foraging range is positively dependent on colony size in many seabirds (Lewis et al. 2001, Wakefield et al. 2013), it also demonstrates that conspecific density is dependent not only upon numbers of birds but the availability of suitable habitat (most simply, open sea). In short, models V–VIII show that birds foraging from a colony with limited access to the sea (e.g., those located in inlets) travel further on average than those from a colony of the same size surrounded by open water (i.e., on isolated islands; Appendix S2: Figs. S4 and S5). For the purposes of our analysis, we recognized that colonies as defined in the Seabird 2000 census (Mitchell et al. 2004) might not correspond to functional units. Our results suggest that, in all species except Razorbills, this is indeed the case (Appendix S2: Table S2). For the other three species, we found that sympatric competition was better quantified by the sum of the inverse distance-weighted number of conspecifics breeding in the area. We hypothesize that this is because the intensity of potential competition from any one conspecific declines as a function of distance to its nest.

It has been hypothesized that seabirds foraging from adjacent colonies segregate in space if potential density of competing conspecifics is high (Wakefield et al. 2013). Segregation among the UDs of colonies has been observed in several species (Masello et al. 2010, Wakefield et al. 2013) but evidence for this phenomenon in our system was mixed: in accordance with the density-dependence hypothesis (Wakefield et al. 2013). Kittiwakes and Murres avoided the areas at which the null ratio of the density of birds from the home colony to those from other colonies was low but Shags and Razorbills did not. Among-colony segregation is also evident in Kittiwake populations geographically disparate from the UK (Ainley et al. 2003, Paredes et al. 2012) and may therefore be widespread in this species but this is the first time that the phenomenon has been reported in Murres. Given the close taxonomic and functional affinities between Razorbills and the latter species it is perhaps surprising that terms describing among-colony segregation were not retained during model selection for Razorbills. This may be because a relatively large proportion (48%) of the Razorbills in our study were tracked from the Northern Isles (Fig. 1; Appendix S2: Table S1), where populations of this and other seabird species have been in decline for the past decade (JNCC 2014) due to declines in forage fish availability (Cook et al. 2014). Razorbills from this region travelled much further (median range 62.7 km, IQR 39–87) than those from other areas (median 20 km, IQR 11–28), possibly due to local food shortages. It is hypothesized that patterns of spatial segregation are, in part, culturally perpetuated via information transfer among conspecifics (Wakefield et al. 2013). If so they may become unstable in a declining population. The apparent lack of spatial segregation among Shags from different breeding sites is notable given that this phenomenon occurs in several other members of the Phalacrocoracidae, such as Phalacrocorax atriceps, P. magellanicus, and P. georgianus (Wanless and Harris 1993, Sapoznikow and Quintana 2003). However, in comparison to these species, European Shags breed in relatively dispersed colonies throughout much of their range in Britain and Ireland so density-dependent competition among breeding aggregations may be insufficient to cause segregation of foraging areas. This could be viewed as an extreme form of segregation, where inter-colony spacing generally exceeds the species’ maximum foraging range. Additionally, in Britain and Ireland, Shag colonies tend to be small, further reducing inter-colony competition. For example, in the Isles of Scilly, where Shags breed at very low densities, birds from different breeding sites forage in common areas (Evans et al. 2015), as suggested by model V (Appendix S2: Fig. S3). Notwithstanding these comments it is possible that our analysis could not detect among-colony foraging segregation in Razorbills and Shags, for two reasons. First, we were unable to track these species from multiple large and closely adjacent breeding sites, where theory suggests segregation is most likely to occur (Wakefield et al. 2013). Second, the census data we used to estimate intraspecific competition was collected 8–16 years before our tracking campaign. Populations of all species in our study are in a state of...
flux: over the past 15 years, Shags have declined by ~30% throughout the region, while Razorbill have declined in the Northern Isles (JNCC 2015). Further tracking from pairs of large, closely adjacent and recently censused colonies would be required to conclusively establish the degree to which spatial segregation occurs among colonies of Shags and Razorbills.

In modelling competition, we made the assumption that seabirds avoid areas of high conspecific density. This is consistent with established foraging theories (the ideal free distribution, optimal foraging, etc.) and is supported by empirical evidence at scales of tens of kilometers and above (Ford et al. 2007, Wakefield et al. 2013). However, at finer scales, local enhancement (when individuals searching for prey are attracted to feeding conspecifics) may cause seabirds to cluster (Fauvel 2009). In our modelling framework, this would manifest as unexplained spatial autocorrelation. Similarly, memory-based foraging or site fidelity, which cause individuals to return repeatedly to the same area (Irons 1998, Wakefield et al. 2015), would result in unexplained temporal, as well as spatial, autocorrelation within individuals. Techniques have been developed for modelling some of these sources of autocorrelation (Marzluff et al. 2004, Aarts et al. 2008, Johnson et al. 2013) but as far as we are aware, no study on a colonial central-place forager to date has been able to model all of these correlation structures simultaneously. This is not only because of the complexity of the task but because the underlying mechanisms are still poorly understood. Conversely however, modelling these dependencies in a hierarchical framework would provide important insights into the foraging strategies employed by seabirds and similar taxa. Recent methodological advances, especially in Integrated Nested Laplace Approximation, may soon make this possible and we look forward to further development of these techniques (Blangiardo et al. 2013).

**Distribution with respect to habitat**

Our principal aim was to estimate usage at sea, irrespective of behavior. Had we modelled foraging locations only, stronger associations than we report might be expected between habitat and distribution (Wakefield et al. 2009). Similarly, considering time-averaged environmental covariates, though expedient, may have reduced the ability of our models to resolve dynamic environmental drivers of distribution if seabirds closely track spatiotemporally unpredictable prey. However, there is increasing evidence that, at the coarse scale, temperate neritic seabirds forage in individually consistent locations, both within and across breeding years (Irons 1998, Weimerskirch 2007, Woo et al. 2008, Wakefield et al. 2015). This may be because shelf sea oceanography is predictably structured by seasonal insolation and tidal stirring (Simpson et al. 1978), suggesting that time-averaged environmental covariates may be reasonable proxies for prey distribution.

The effects of habitat on spatial usage in our models were comparable in magnitude to those of foraging costs and competition (Table 2). Moreover, the habitat preferences indicated by models V–VIII accord with current understanding of the foraging ecology of the study species. For example, covariates describing substrate were retained only in models of habitat use for the three diving species (Shags, Murres, and Razorbills). Shags and Murres forage both at or near the seabed and in the water column so substrate type may affect prey availability directly (Watanuki et al. 2008, Thaxter et al. 2010). Razorbills forage at shallower depths but in common with all species in the study, prey primarily on sandeels, whose distribution varies with sediment coarseness and silt content (Wright et al. 2000, Holland et al. 2005). Previous studies suggest that sympatrically breeding Razorbills and Murres, which are closely related, do not segregate in space (Thaxter et al. 2010, Linnebjerg et al. 2013). However, our results suggest some landscape-scale niche partitioning: Murres outnumber Razorbills in inshore waters of the North Sea, the Northern Isles, and the Irish Sea, whereas Razorbills predominate in the Western Isles. Notably, our models also suggest a preponderance of Murres in the vicinity of the Celtic Sea front, which may reflect divergent foraging adaptations in these species (Appendix S2: Fig. S8).

Covariates best describing the distribution of Kittiwakes, which are obligate surface feeders, either described properties of the water column (stratification and relative sSST) or the morphology of the seabed (slope), which affects turbulent mixing. Presumably, these covariates were retained because they describe physical mechanisms that affect prey availability indirectly, either by enhancing production at lower trophic levels (e.g., tidal stirring resupplies nutrients to the photic zone; Scott et al. 2010, Carroll et al. 2015) or by advecting prey to the surface (Embling et al. 2012, Cox et al. 2013).

Species distribution models fitted to data collected in one area may predict usage poorly in another where habitat availability differs. To account for this effect we considered models in which the response of birds to candidate environmental covariates was conditioned on their regional means (i.e., a partial implementation of a Generalized Functional Response [GFR] to resource availability; Matthiopoulos et al. 2011). GFRs with respect to some but not all covariates improved model performance, indicating that seabirds responded non-linearly to changes in the availability of some environmental covariates. This is perhaps unsurprising, given the oceanographically complex nature of the study area (Appendix S1: Fig. S2). For example, Murres tend to forage far from the coast in areas where the mean distance to the coast was high, such as the North Sea, which has a relatively simple geometry. In areas where the mean distance to the coast was low, such as the geometrically complex Northern and Western Isles, this relationship was reversed (Table 2; Appendix S2: Fig. S5). Presumably, this reflects differences in the
dominant physical drivers of prey distribution or the type of prey available to Murres in these areas.

Conservation implications

For conservation measures to be effective they must be evidence-based so there is an urgent need to map the distributions of seabirds at sea and to understand how these are shaped by intrinsic and extrinsic factors (Lewison et al. 2012). We estimated seabird distribution using data on the size and location of all known colonies in Britain and Ireland. However, missing substrate data meant that we did not estimate usage by Shags, Murres, and Razorbills outside the UK Exclusive Economic Zone (EEZ) or for parts of the Northern and Western Isles (Fig. 2). Moreover, we did not have access to contemporaneous data on conspecific colonies in countries bordering the study area. Although these may interact with colonies in Britain and Ireland, their relatively small size and large distance from Britain and Ireland suggest that any density-dependent competition from these colonies is likely to be negligible. Notwithstanding these caveats, the performance of our time-invariant models suggest that the factors determining the marine distribution of breeding seabirds in Britain and Ireland are sufficiently consistent across time to permit reliable estimation of area usage from biotelemetry, environmental covariates, and central-place foraging theory, which has important consequences for identification of priority areas for conservation measures. To date, potential offshore SPAs for European seabirds have been identified largely using at-sea transect survey data (Skov et al. 1995, Kober et al. 2012) and progress to designate offshore protected areas has been slow (BirdLife International 2010). Moreover, because it is impossible to derive colony-specific distribution estimates from at-sea observations, tracking is increasingly used to obtain the colony-level seabird distributions (Wakefield et al. 2011, Raymond et al. 2015) that are required for the assessment of impacts of marine industries on protected breeding colonies. Unfortunately, it is neither practicable to track widespread species from all their colonies, nor clear how usage can be interpolated from surveyed to unsurveyed colonies (Aarts et al. 2008, Matthiopoulos et al. 2011, Torres et al. 2015). Thaxter et al. (2012) suggested that, until better information becomes available, a pragmatic approach (the “radius” method) is to assume that seabirds are distributed uniformly out to some putative maximum range from their colonies. However, as our analysis and others confirm (e.g., Wakefield et al. 2011, 2013, Catry et al. 2013, Dean et al. 2015), seabird density declines with distance from the colony. Moreover, density-dependent competition, coastal morphology, and habitat preference result in highly non-uniform distributions. We show that these effects can be estimated by tracking birds from a sample of colonies and fitting IPP models, structured as partial GFRs (Matthiopoulos et al. 2011), to the resulting data. The ability of these models to estimate seabird distributions at unsampled colonies is a major innovation. Moreover, an advantage of IPP models over the logistic presence/pseudo-absence models latterly applied to tracking data is their interpretability (Aarts et al. 2012, Renner et al. 2015). Our models predict “occurrences at sea per day per individual” (i.e., incorporating information on both activity budget and space use), which is directly proportional to the average amount of time birds are expected to spend at a location and therefore of direct utility to conservation managers. The areas of intensive usage we identified, especially those used by birds from SPA breeding colonies, may warrant consideration for statutory protection following the principles recently outlined by Wilson et al. (2014). Moreover, the provision of colony-level predictions allows the potential impacts of anthropogenic and natural processes to be apportioned to specific colonies much more accurately than is possible using the radius method. This will be of particular importance in assessing potential impacts from offshore windfarms, which are projected to increase 10-fold in European shelf seas in the next decade, with the majority being constructed in UK waters (Infield 2013). Current assessments of the potential barrier, displacement, and collision impacts, both at the individual windfarm level and the region-wide level, rely either on data from boat or aerial surveys (Furness et al. 2013, Maclean et al. 2013), tracking from very few colonies (Perrow et al. 2006) or the radius method (Thaxter et al. 2012, Bradbury et al. 2014). As such, potential impacts cannot be reliably apportioned to breeding colonies, hampering attempts to predict their demographic consequences (Bailey et al. 2014). Similarly, the impacts of oil pollution and bycatch may be highly localized (Williams et al. 1995, Zydelis et al. 2013) so colony-level distribution estimates will facilitate spatial planning decisions that more effectively balance seabird conservation with competing interests, by linking marine aggregations of seabirds to specific colonies. The methods presented here demonstrate the utility of tracking data to estimate seabird distribution at national scales and further data are now required to allow the application of this modelling approach to other breeding seabird species. Moreover, by combining our results across species, potential areas of high conservation priority are revealed (Fig. 3; Appendix S2: Fig. S7). It is clear that, within Britain and Ireland, the core areas of usage of all four study species overlap within most of the coastal waters in Scotland. Areas of high multi-species usage may warrant particular attention, since both the vulnerability to threats and the potential benefits of conservation measures, are likely to be highest there. The regions identified as supporting the core 90% UD of at least three of the species considered here (Appendix S2: Fig. S7b) correspond well to those areas identified as of greatest international importance for 30 seabird species in the North Sea across all seasons (Skov et al. 1995), indicating the likely importance of these areas for a broad range of avian taxa.
Inclusion of density-dependent competition in our models increased their predictive performance. However, this improvement over previous similar analyses (Wakefield et al. 2011, Raymond et al. 2015) was only possible because the sizes of most seabird colonies in Britain and Ireland are known (Mitchell et al. 2004). In contrast, seabird colonies in many regions have not been censused (Croxall et al. 2012). Obtaining accurate estimates of colony size should be a priority for wildlife managers intending to use tracking data to estimate the distribution of seabirds from unsampled colonies. Moreover, our results suggest that distribution will change if colony sizes alter. Updating colony counts periodically would allow model-based distribution estimates to be revised without necessarily needing to collect more tracking data.

Acknowledgments

We gratefully acknowledge the many fieldworkers who assisted with data collection (S. Adlard, D. Aitken, G. Anderson, C. Bell, A. Bellamy, R. Brown, R. Bufton, M. Chimienti, T. Coledale, K. Colhoun, P. Collins, B. Dean, D. Evans, D. Fox, C. Gunn, R. Hughes, J. Lamb, R. Langton, L. Mackley, A. Macmillan, T. Newman, M. Nydegger, L. Quinn, N. Richardson, Y. Satgé, E. Scragg, J. Sturgeon, K. Smell, C. Taylor, J. Taylor, and others); landowners who permitted access to study sites; Jason Matthiopoulos and John Fieberg for statistical advice; and Peter Cornelius for access to computing facilities; Mitohiro Ito for access to tracking data; and NERC Earth Observation Data Acquisition and Analysis Service for supplying SST and NPP data. We thank John Marzluff, Norman Ratcliffe, Jeremy Wilson, Linda Wilson, Alex Kinninmonth, Rosie Sutherland, and Kate Jennings for valuable comments on earlier drafts of this paper. Seabird counts extracted from the Seabird Monitoring Programme (SMP) database were provided by the generous contributions of nature conservation and research organizations, and many volunteers throughout Britain and Ireland (the SMP Database Host, data provider, original recorder and the SMP Partnership bear no responsibility for any further analysis or interpretation of these data). This study was conducted using MyOcean products and we thank all organizations involved. Figures depicting results from models V, VII, and VIII were derived using Geographical Map Data © NERC 2015. Funding was provided by the European Regional Development Fund through its Atlantic Area Programme, Marine Scotland, Scottish Natural Heritage, Natural England, the Natural Environment Research Council, the Joint Nature Conservation Committee, Natural Resources Wales, Environment Wales, Argyll Bird Club, Fair Isle Bird Observatory Trust, and the RSPB.

Literature Cited


frequency coded data. Page 11 TIPR’97. University of Manchester, Prague, Czech Republic.


**Supporting Information**

Additional supporting information may be found online at: http://onlinelibrary.wiley.com/doi/10.1002/eap.1591/full

**Data Availability**