Diversity & Distributions

Biodiversity Research

REVISED MANUSCRIPT: DDI-2015-0139

Identifying predictable foraging habitats for a wide-ranging marine predator using ensemble ecological niche models

Kylie L. Scales[*1,2,3], Peter I. Miller[1], Simon N. Ingram[4], Elliott L. Hazen[3], Steven J. Bograd[3], Richard A. Phillips[5]

1. Plymouth Marine Laboratory, Prospect Place, Plymouth PL1 3DH, UK
2. Institute of Marine Sciences, University of California Santa Cruz, CA 95064, USA
3. National Oceanic and Atmospheric Administration (NOAA) Southwest Fisheries Science Center, Environmental Research Division, 99 Pacific Street, Suite 255A, Monterey CA 93940, USA
4. Marine Biology and Ecology Research Centre, Plymouth University, Plymouth PL4 8AA, UK
5. British Antarctic Survey, Natural Environment Research Council, High Cross, Madingley Road, Cambridge CB3 0ET, UK

Running title: *Modelling habitat suitability for a marine predator*

ABSTRACT

Aim

Ecological niche modelling can provide valuable insight into species' environmental preferences, and aid in identification of key habitats for populations of conservation concern. Here, we integrate biologging, satellite remote-sensing and ensemble ecological niche models (EENM) to identify predictable foraging habitats for a globally important population of the grey-headed albatross (GHA) *Thalassarche chrysostoma*.

Location

Bird Island, South Georgia; Southern Atlantic Ocean

Methods

GPS and geolocation-immersion loggers were used to track at-sea movements and activity patterns of GHA over two breeding seasons (n=55; brood-guard). Immersion frequency (landings per 10-minute interval) was used to define foraging events. EENM combining Generalised Additive Models (GAM), MaxEnt, Random Forest (RF) and Boosted Regression Trees (BRT) identified the biophysical conditions characterising the locations of foraging events, using time-matched oceanographic predictors (Sea Surface Temperature, SST;

chlorophyll-*a*, chl-*a*; thermal front frequency, *TFreq*; depth). Model performance was assessed through iterative cross-validation, and extrapolative performance through cross-validation between years.

Results

Predictable foraging habitats identified by EENM spanned neritic (<500m), shelf-break and oceanic waters, coinciding with a set of persistent biophysical conditions characterised by particular thermal ranges (3-8°C, 12-13°C), elevated primary productivity (chl-*a* > 0.5mg m^{-3}) and frequent manifestation of mesoscale thermal fronts (*TFreq* > 25%). Our results confirmed previous suggestions that GHA utilise oceanic fronts, and objectively identified the Antarctic Polar Frontal Zone (APFZ) as suitable foraging habitat.

Over the spatial and temporal scales investigated here, performance of EENM was superior to that of single-algorithm models. In particular, MaxEnt performed poorly, resulting in highly variable predictions and exclusion from final EENM. Resultant EENM displayed good extrapolative performance between years.

Main Conclusions

EENM techniques are useful for integrating the predictions of several single-algorithm models, reducing potential bias. Our analysis highlights the value of EENM for use with movement data in identifying at-sea habitats of wide-ranging marine predators, with clear implications for conservation and management.

Keywords:

albatross; biologging; Boosted Regression Trees; front map; GAM; habitat model; ocean front; Random Forest; satellite remote sensing

(A) Introduction

Ecological niche modelling (also referred to as *species-habitat, predictive habitat, habitat-based* and *species distribution modelling*) provides a framework for understanding species' distributions as a function of their environmental preferences, and for identifying priority areas for conservation. Understanding the mechanisms that underlie environmental preference is particularly challenging for highly mobile species with complex life histories, especially in the marine realm where conditions are dynamic. Recent efforts to integrate animal tracking ('biologging'), satellite remote-sensing and ecological niche modelling have generated valuable insights into the interactions between highly mobile marine species and the oceanic environment (e.g. Torres *et al.*, 2015; Howell *et al.,* 2015; Raymond *et al.,* 2015). However, most studies utilise a single modelling framework with its specific biases, reducing the comparability of results and potentially limiting predictive capacity. An alternative is to adopt an ensemble ecological niche modelling approach (EENM; Araújo & New 2007), which combines the output of multiple model algorithms into one predictive surface and has been used successfully for identifying key habitats for marine predators, including sea turtles (Pikesley *et al.,* 2013) and seabirds (Oppel *et al.,* 2012).

Predicting the locations of suitable foraging habitats for wide-ranging pelagic species such as procellariiform seabirds (albatrosses, petrels and shearwaters) is non-trivial, given the complex and scale-dependent interactions between oceanographic processes and prey field dynamics, and the diverse aspects of bird physiology, energetics, reproductive and other constraints that govern foraging behaviour. The spatial ecology of pelagic seabirds appears to be influenced by processes both extrinsic and intrinsic to each individual. For example, habitat preferences of Southern Ocean seabirds vary among species (Commins *et al.,* 2014), populations (Nel *et al.,* 2001, Louzao *et al.,* 2011, Joiris & Dochy 2013), and individuals (Phillips *et al.,* 2006; Patrick & Weimerskirch 2014); between sexes (Phillips *et al.,* 2004); between life history stages (Phillips *et al.,* 2005); through the annual cycle (Phillips *et al.,* 2006, Wakefield *et al.,* 2011); and in response to changes in oceanographic conditions (Xavier *et al.,* 2013). Ecological niche modelling must be conducted with an awareness of the multi-faceted influences on habitat selection if it is to be informative for identifying and managing priority areas for conservation (Lascelles *et al.,* 2012).

The energetic demands of reproduction are known to strongly influence habitat selection by pelagic seabirds during breeding periods. The constraints of incubation and chick provisioning impose a central-place foraging mode, as trips are restricted to waters within an accessible range of the colony (Weimerskirch *et al.,* 1993). Individuals face trade-offs

between the costs of flight and the necessity for reliable acquisition of prey of sufficient quality to meet the demands of chick provisioning in addition to their own energetic requirements, including for self-maintenance (Weimerskirch *et al.,* 1997). These constraints are particularly pronounced during the brood-guard period, when chicks require continual attendance by a parent to avoid chilling, are at their most vulnerable to predation, and have a small stomach volume so require frequent meals (Weimerskirch *et al.,* 1988, Xavier *et al.,* 2003, Wakefield *et al.,* 2011).

Breeding success is therefore conditional upon the abilities of each bird to predict the locations of suitable foraging habitats within a commutable distance of the colony. The oceanic seascapes over which pelagic seabirds search for food are highly heterogeneous, with prey distributed within a *nested patch hierarchy* (Fauchald *et al.,* 2000, Weimerskirch 2007). Suitable foraging habitats that include prey of sufficient number and quality are accessible within the diving capabilities of the species, are formed by stochastic biophysical processes; hence, the locations of exploitable prey aggregations are usually unpredictable at small spatial scales (Hazen *et al.*, 2013). However, there is evidence to suggest that some species, particularly albatrosses, may target or track regions in which the availability of prey resources is related to persistent oceanographic conditions and hence predictable over broad- to meso-scales, thus optimising foraging success (Kappes *et al.,* 2010, Louzao *et al.,* 2011, Piatt *et al.,* 2006, Weimerskirch 2007).

Grey-headed albatrosses (GHA) *Thalassarche chrysostoma*, in common with many Southern Ocean predators, have been shown to exploit predictable and profitable foraging opportunities generated through bio-physical coupling along ocean fronts – physical interfaces between contrasting water masses (Bost *et al.,* 2009, Belkin *et al.,* 2009). The Antarctic Polar Frontal Zone (APFZ), an extensive, dynamic region that marks the northern boundary of the Antarctic Circumpolar Current (ACC), is known to be an important feature for seabirds and marine mammals in this sector of the Southern Ocean (Catry *et al.,* 2004, Scheffer *et al.*, 2012, Wakefield *et al.,* 2011). Within the broad-scale APFZ, intense oceanographic dynamics lead to the generation of chaotic eddies and the manifestation of mesoscale (10s -100s of kilometres) or sub-mesoscale (~1 kilometre) thermohaline fronts. Aggregations of prey, such as the mesopelagic fish and cephalopods often targeted by the grey-headed albatross, can be concentrated within this zone, both through processes of mechanical entrainment and bottom-up forcing of biophysical hotspots (Rodhouse & White 1995, Reid *et al.,* 1996, Catry *et al.,* 2004, Rodhouse & Boyle 2010). Areas of frequent or persistent frontal activity, such as the APFZ, may therefore constitute predictable foraging habitats for regional populations of pelagic seabirds.

Here, a novel application of EENM is developed, using high-resolution data tracking the movements and activity patterns of GHA from the largest global colony, to identify persistent oceanographic conditions that characterise predictable foraging habitats within the area accessible to breeding birds. We use a suite of remotely-sensed oceanographic data, including the first regional application of a thermal front frequency index, in an iterative presence-availability model framework, with the following aims: i) to identify the biophysical conditions that characterise the locations of observed foraging events during the brood-guard phase; ii) to model the spatial distribution of predictable foraging habitats, iii) to explore the comparative utility of EENM and single-algorithm models in the context of using movement data to define foraging habitats of wide-ranging species over broad- to meso-scales and iv) to evaluate the extrapolative performance of EENM through time, and hence its usefulness for conservation and management applications.

(A) Methods

(B) Device deployment

Birds were tracked from Colony B at Bird Island, South Georgia (54°00'S 38°03'W) over December-January of two austral breeding seasons, during the brood-guard phase (total n=55 birds; n=25 in 2009/10; n=30 in 2011/12; Fig. 1). GPS loggers used were i-gotU (MobileAction Technology; http://www.i-gotu.com; 25g mass), earth & Ocean Technology (e&O-Tec) MiniGPSlog (25g) or e&O-Tec MicroGPSlog (10g) and were attached using Tesa® marine cloth tape (total 5g) to mantle feathers. Devices were programmed to record fixes at 10 or 15 minute intervals and were recovered after one complete foraging trip. Birds were also equipped with geolocation-immersion loggers (British Antarctic Survey; Mk 13; ~1.5g mass), attached to a standard British Trust for Ornithology metal or plastic ring. Birds were restrained on the nest only during device deployment, and handling time during deployment and retrieval was minimised (5-10 mins).

(B) Behavioural classification

Landing rate (number of landings per 10-minute interval) derived from the immersion data was used to identify foraging bouts (following Dias *et al.,* 2010). Take-off from the water surface is energetically costly for albatrosses, so we assumed that immersion events indicated prey capture attempts (following Wakefield *et al.,* 2011). Empirical evidence from previous

work on this population shows that birds frequently catch prey in rapid directed flight without any obvious area-restricted search (ARS) behaviour (Catry *et al.,* 2004), so we used landing rate as an indicator of foraging behaviour in preference to identifying ARS.

Locations of immersion events were derived through temporal matching of GPS and immersion data. As birds rest on the water surface overnight (Catry *et al.,* 2004), and night-time foraging could not be differentiated from resting, only those locations recorded in daylight hours were used (bounded by civil dawn and dusk; solar zenith angle of -6°). All locations within a 50km radius of the colony were excluded from analysis to remove rafting behaviour. All GPS tracks were interpolated to regular 10 minute intervals. Landing rate was derived using a sliding window that summed the number of immersion events and total time spent immersed in the 10 minute track section preceding each GPS point location. Interpolated point locations along each track were then classified as either foraging – associated with at least one immersion event within ten minutes – or transit – not associated with immersion.

The study area was defined as the area enclosing a radius corresponding to the absolute maximum displacement from the colony by any tracked bird (1185km). To obtain an indication of the spatial distribution of foraging events over the tracking period, a 2-dimensional regular grid of the study area (71°S to 32°S; 55°W to 21°W) was created at 0.5° resolution. A binary classification index of grid cell usage was used to identify foraging areas - grid cells in which foraging events were recorded over the course of the tracking period were designated as 1, and grid cells that contained transit locations, or no bird presence, were designated as 0. All analyses were conducted in R version 3.1.

(B) Oceanographic data

Remotely-sensed oceanographic data were obtained for a matching timespan (late December – end January) for each tracking period (2009/10; 2011/12). Daily NASA Multi-Sensor Merged Ultra-High Resolution (MUR) Sea Surface Temperature (SST) imagery was downloaded via OpenDAP, and daily chlorophyll-*a* (chl-*a*) imagery was processed from MODIS-Aqua data; both were mapped to the study area in geographic projection at 1.2km resolution. Daily images were used to generate monthly median SST and chl-*a* (log-scaling) composites. Bathymetric data were obtained for a matching spatial extent from the General Bathymetric Chart of the Oceans (GEBCO_08 grid; http://www/gebco.net), and used to derive depth at 30 arc-second resolution.

Thermal composite front maps (Miller 2009) were generated from MUR SST data, over rolling 7-day periods spanning the tracking period. Thermal fronts were detected in each MUR SST scene using Single-Image Edge Detection (SIED; Cayula & Cornillon 1992; front detection threshold = 0.4°C). Successive 7-day composites were used to prepare monthly front frequency (*TFreq*) layers, which quantify the frequency in which a front is detected in each pixel as a ratio of the number of positive detections to the number of cloud-free observations. All environmental data layers were standardised at 0.5 degree resolution through bilinear interpolation ('raster' package for R; Hijmans & van Etten 2012; Fig. 2). Oceanographic data layers were selected on the basis of availability, coverage and previously demonstrated influence on habitat selection by GHA and sympatric seabird species (e.g. Xavier *et al.,* 2003, Phillips *et al.,* 2006, Wakefield *et al.,* 2011, Ballard *et al.,* 2012).

(B) Ensemble Ecological Niche Modelling (EENM)

Previous work comparing the efficacy of various modelling algorithms for predicting habitat preferences in seabirds concluded that an ensemble approach can be preferable to the use of a single-algorithm models (Oppel *et al.,* 2012). However, the technique has not to our knowledge been used previously to identify predictable foraging habitats for seabirds using movement data. We used EENM to identifying the biophysical conditions characterising the locations of observed albatross foraging events. Ecological niche models (ENM) were fitted using the Generalised Additive Modelling (GAM), Maximum Entropy (MaxEnt), Random Forest (RF) and Boosted Regression Tree (BRT) algorithms within the biomod2 package for R (Thuiller *et al.,* 2009, 2014).

The package ‘biomod2’ uses a presence-availability framework to model preferred conditions. As grid cells in which no foraging events were detected cannot be classified as true absences, control locations (‘pseudo-absences’) were iteratively resampled from within the accessible radius of the breeding colony. Five iterations of 1000 randomly-selected control locations were used over successive model runs (Barbet-Massin *et al.*, 2012). Each model run involved 10-fold cross-validation, with data randomly apportioned to a 75% / 25% split for model calibration and testing phases.

Relative importance of environmental variables was determined using the built-in method in biomod2, which overcomes difficulties associated with comparing model-specific outcomes through a randomisation procedure (Thuiller *et al.,* 2009, 2014), which fits a Pearson correlation between the fitted values and predictions, where each variable has been randomly permutated. If the two predictions are similar, i.e. highly correlated, the variable is

186 considered of little importance. This procedure was repeated 10 times for each variable
187 within each model run. The relative importance of each environmental variable (Relative
188 Importance of the Contribution to the model Coefficients, RICC) was then scaled by
189 subtracting the mean correlation coefficient from 1. The overall explanatory power of the
190 environmental variables was derived using the mean-of-means of standardised variable
191 importance over all iterations per algorithm (Table S1).
192
193 The EENM combines predictions from single-algorithm model runs. Outputs of each single-
194 algorithm model were evaluated over both model calibration and testing datasets for each
195 model iteration. A triad of model performance metrics (AUC, TSS, Boyce Index) was
196 generated for each iteration per algorithm, and the mean of each of these metrics over each
197 iteration of control locations was calculated. The mean of each performance metric over all
198 models fit per algorithm was then calculated (n=50; 10-fold cross-validation for each of 5
199 iterations of control locations; Tables S3, S4). Only those with a True Skill Statistic (TSS)
200 equal to or greater than 0.7 were included in the final ensemble, to minimise inclusion of
201 poorly-performing models. The ensemble projections were created using a weighted average
202 across all included single-algorithm models, based on TSS, and accounting for differences in
203 algorithm performance. EENM projections were based on a habitat suitability index (HSI),
204 scaled between 0 and 1, where 1 represents greatest suitability.
205
206 Resultant EENMs were then evaluated, using AUC, TSS and Boyce Index (Boyce *et al.*,
207 2002; Hirzel *et al.*, 2006). We calculated all performance metrics for each EENM fitted to the
208 full dataset from each year. AUC and TSS were calculated using in-built biomod2
209 functionality. Boyce Index was calculated through projection of each model on to the full
210 dataset for each year ('ecospat' package for R; Broenniman *et al*., 2014) to obtain a value
211 comparing model predictions of HSI with the input presence dataset in each case.
212
213 In preference to specifying a threshold of HSI to calculate the extent of suitable foraging
214 habitat within the area accessible to the population during this breeding phase, we derived a
215 measure of the proportion of this accessible area in which suitable foraging conditions were
216 predicted over a continuum of HSI from 0 to 1.
217
218 (B) EENM Extrapolative Performance
219
220 EENM extrapolative performance was assessed through cross-validation between the two
221 years for which we had data. We projected each model on to the combined synoptic
222 environmental data surfaces for the years following (2009/10 model onto 2011/12

environmental data) or preceding (2011/12 model onto 2009/10 environmental data) that upon which the model was constructed. Performance metrics (AUC, TSS, Boyce Index) were calculated for each of these projected models, following methods described above. Spatial concordance between predictions of models extrapolated across time and year-specific models was quantitatively compared using Mantel tests (ade4 package for R; Dray & Dufour, 2004).

(A) Results

(B) Foraging Trips

Maximum displacement from the colony ranged between 153km and 1185km, with a mean ± SD of 744 ± 249km. Trip duration ranged between 0.6 and 6.1 days, with a mean of 2.9 ± 1.3 days. All trips involved at least one foraging event (based on landing rate derived from the immersion data), with a mean of 6.1 ± 3.7 foraging events per trip (range 2 – 17).

Sex was available for a small sub-sample of tracked birds (n=8, 2009/10; n=5, 2011/12), in which no differences in foraging trips between sexes were detected (Fig. 1). Given the small sample of known sex, sex effects were not included in further population-level analyses.

(B) Predictable foraging habitats

Median SST and chl-*a* concentration were important contributory variables to EENMs constructed for both years of the study, suggesting these biophysical variables strongly influence albatross foraging over the scales investigated by our models (Table 1). However, the overall explanatory contribution of chl-*a* to the 2011/12 EENM (RICC=0.150) was lower than its contribution to the 2009/10 EENM (RICC=0.585), and the inverse was observed for the contribution of SST to each EENM (RICC, 2009/10=0.577; RICC, 2011/12=0.744). The relative contributions of water depth and the frequency of mesoscale thermal front manifestation (*Tfreq*) to the explanatory capabilities of the EENM were lower than that of SST and chl-*a* across both years, although *TFreq* and depth were more important to the 2011/12 model set (RICC, *TFreq*=0.155, RICC, depth=0.100) than for 2009/10 (RICC, *TFreq*=0.037; RICC depth=0.086).

Spatial predictions of EENMs identified suitable foraging conditions across neritic (<500m depth), shelf-break and oceanic regions, reflecting the variety of foraging locations used by birds tracked in both the 2009/10 and 2011/12 breeding seasons (Fig. 3). EENM-derived spatial predictions of habitat suitability across the accessible area were very similar in extent and direction among years (Fig. 3a,b). Regions of high habitat suitability were associated with particular SST ranges (3-8°C, 12-13°C) and productive regions (median chl-*a* >0.5 mg m^{-3}) of

the area accessible to foraging birds. The APFZ (Fig. 2e,f) was also identified as an area highly suitable for foraging in both years (Fig. 3), although this zone lies at the extremes of the area accessible to birds during this breeding stage (Fig. 1).

(B) EENM vs. single-algorithm models

(C) Model Predictions

The ranking of the environmental variables in terms of explanatory contribution (mean over 50 runs per algorithm) was broadly comparable among single-algorithm models, although we observed some variability (Table 1). For example, ranking of environmental variable importance was similar among GAM, RF and BRT models in both years. EENM variable rankings smoothed out the variability evident in estimated variable importance among model sets. However, explanatory contributions of environmental variables were ranked differently by year-specific EENMs (Table 1).

Model response curves for each environmental variable were comparable among algorithms. GAM, RF and BRT in particular generated model sets with very similar response curves for SST (Fig. 4), *TFreq* and depth, although less consistency among algorithms is evident in chl-*a* response curves. MaxEnt models were subject to greater inconsistency in predicted responses (Figs. S1 – S3).

Similarly, spatial predictions of models fitted using the GAM, RF and BRT algorithms were comparable in the extent and location of suitable habitats identified, and in the scaling of the habitat suitability index (HSI) in these regions (Fig. 5). MaxEnt models, however, generated more spatially restricted predictions with overall lower HSI predicted throughout the accessible area. For these reasons, we did not include MaxEnt in the final EENMs per year. The location and extent of suitable habitats identifed and the scaling of HSI in EENM predictions integrated the predictions of the GAM, RF and BRT algorithms, smoothing over variation between model frameworks (Fig. 3). Spatial predictions of all single-algorithm models were similar in extent, location and HSI scaling among years (Fig. 5). EENM predictions showed a strong spatial concordance in the location and extent of suitable habitats identified in each year (Fig. 3; HSI, Mantel r=0.9599).

(C) Model Performance

EENMs were highlighted by AUC and Boyce Index as the best performing models in comparison with all single-algorithm models for both years. However, the True Skill Statistic (TSS) selected Random Forest (RF) as the best performing in both years (Table 2).

Evaluation metrics indicated similar performance of single-algorithm models across model sets, (variance, AUC=0.0002; TSS=0.001; Boyce Index=0.002; Table 2), and for each of these single-algorithm models among years (correlation, AUC r=0.999; TSS=0.935; Boyce Index=0.884; Table 2). There was little concordance between the rankings of model performance for single-algorithm models among the three model performance metrics used (AUC, TSS, Boyce Index), although AUC and TSS ranked single-algorithm models in a similar order in both years (e.g. AUC = RF, BRT, GAM, MaxEnt; Table 2).

The exclusion of MaxEnt models from the final EENMs per year had little effect on model performance metrics, although a slight improvement was evident in AUC, TSS and Boyce Index in both years (Table 2). The weighted mean EENM including predictions of GAM, RF and BRT models was retained as the final model for each year.

(B) EENM Extrapolative Performance

EENMs extrapolated across years to predict suitable foraging habitats over differing mesoscale oceanographic conditions performed well according to AUC and Boyce Index scores of projected models. All model performance metrics (AUC, TSS, Boyce Index) reveal the extrapolative performance of the 2011/12 EENM to be superior to that of the 2009/10 EENM. However, the TSS scores of both models dropped below the 0.7 threshold used to select best performing models for EENM creation.

Spatial predictions of EENMs extrapolated across years were broadly comparable to the predictions of each year-specific EENM, highlighting the suitable foraging habitats located to the north and west of the colony. Extrapolation of the 2011/12 EENM to the 2009/10 combined environmental data surface exhibited strong similarity with the 2009/10 EENM (HSI, Mantel r=0.9437), but extrapolation of the 2009/10 EENM on to 2011/12 conditions predicted more spatially restricted regions of high habitat suitability than those predicted by the year-specific model (HSI, Mantel r=0.8740; Fig. 3). The proportion of the area accessible to the population during this breeding phase in which suitable foraging habitats were predicted to occur was also comparable among years (Fig. 6).

(A) Discussion

Predictable foraging habitats for the grey-headed albatross population breeding at Bird Island, South Georgia appear to coincide with a set of persistent biophysical conditions characterised by particular thermal ranges and elevated primary productivity. Over the spatial and temporal

scales investigated by our models, EENM performed better than single-algorithm models in predicting the locations of suitable foraging habitats. These insights highlight the potential of EENM as a tool for use with movement data for identifying at-sea habitats of wide-ranging marine predators, with clear implications for conservation and management.

(B) Predictable foraging habitats

Our ensemble ecological niche models (EENMs) highlight sea surface temperature (SST) and median surface chlorophyll-*a* (chl-*a*) concentration (monthly synoptic fields) as important determinants of habitat suitability for foraging grey-headed albatrosses during the brood-guard phase. SST has been found to be a useful predictor of habitat preference for other albatross species at South Georgia, and elsewhere (Wakefield *et al.,* 2011; Deppe *et al.,* 2014, Kappes *et al.,* 2010; Awkerman *et al.,* 2005). GHA also appeared to respond to the frequency of mesoscale thermal front manifestation (*Tfreq*), which characterised the APFZ, and to water depth, although these predictors had less influence in models.

SST is a proxy for the spatial structuring of biophysical conditions over the vast ranges utilised by these ocean-wandering seabirds, and so often proves useful in identifying predictable habitats. Different foraging guilds of pelagic predators exploit prey types that associate with particular temperature regimes (Commins *et al.,* 2014). GHA are known to seize prey from the ocean surface (<2-3m depth; Huin & Prince 1997), and to feed predominantly on ommastrephid squid, including *Martialia hyadesi*, crustaceans, including Antarctic krill *Euphausia superba*, and, less commonly, lamprey *Geotria australis*, mesopelagic fish and gelatinous zooplankton (Rodhouse *et al.,* 1990, Reid *et al.,* 1996, Xavier *et al.,* 2003, Catry *et al.,* 2004). Although the diet of the tracked birds was not determined in the current study, their distribution was broadly comparable with previous years when all these prey types were recorded (Catry *et al.,* 2004, Xavier *et al.,* 2003). This suggests that the environmental conditions identified through this modelling procedure reflect the key habitats and main prey that are targeted by grey-headed albatrosses at South Georgia, which represent c. 50% of the global breeding population (ACAP 2009).

Chl-*a* was also identified as a predictor of the spatial distribution of foraging events. Overall, foraging activity was more likely in productive regions. Chl-*a* concentrations (monthly median) were highest on-shelf, with peak values recorded to the south-west of the colony. The APFZ was not characterised by elevated productivity over the spatial and temporal scales investigated in this model. Birds foraging in productive shelf waters around South Georgia are likely to be targeting Antarctic krill and icefish *Champsocephalus gunnari*, which are

more closely tied to bottom-up forcing mechanisms than the squid and mesopelagic fish found in the APFZ (Wakefield, Phillips & Belchier 2012).

High *Tfreq* values and narrow SST contours characterise the APFZ, which was identified by the EENM as a region of high habitat suitability for GHA. Plunge-diving GHA have been observed in association with large aggregations of *M. hyadesi* at the ocean surface within the APFZ (Rodhouse & Boyle 2010). Although few foraging events were observed in the APFZ during the tracking period, it is likely that those birds foraging in the APFZ region were targeting ommastrephid squid. The APFZ lies at the northernmost extreme of the observed foraging range during brood-guard, which might suggest that reproductive constraints influenced the strength of the association with this region. Regardless, the high spatial overlap between the APFZ and the distribution of GHA during other breeding stages and in the non-breeding period (Phillips et al. 2004, Croxall et al. 2005) suggest it is a key foraging area for this species, year-round.

In previous studies in the region, the spatial extent of the APFZ has been estimated using historical or averaged data, which did not match the temporal resolution of animal movement data. For example, Xavier *et al.* (2003) used the position of the Polar Front (PF) derived from survey data in 1997 to investigate habitat preference of birds tracked in 2000. However, the APFZ is a highly dynamic feature, characterised by intense mesoscale variability, and the PF can vary in position by as much as 100km in 10 days (Trathan *et al.,* 1997). Detecting fronts in a temporally-averaged SST composite can also mask the dynamic nature of these features. The *Tfreq* index, used here for the first time in the Southern Ocean, is an objective, synoptic product that enables incorporation of mesoscale oceanographic dynamics in broad-scale ecological niche models (Scales *et al.*, 2014).

In addition to the selection of environmental data layers, analytical scale is a key aspect of the construction of ecological niche models. Matching the spatial resolution of remotely-sensed datasets with the scales over which animals locate key foraging areas remains a major challenge in habitat modelling (Storch 2002, Luoto *et al.,* 2007), particularly in the marine realm (Araújo & Guisan 2006, Hirzel *et al.,* 2006). In our study, environmental data layers were interpolated to a standard 0.5 degree grid resolution, which was deemed appropriate given the extent of the area over which tracked birds roamed. In order to ensure scale match of the research question, response and environmental datasets, we also restricted temporal averaging of environmental data layers to one month, matching the duration of the brood-guard phase for the focal population.

(B) EENM vs single-algorithm models

(C) Model Predictions

Single-algorithm ecological niche models fitted on the same dataset can perform differently and generate contrasting predictions (Guisan & Zimmerman 2000, Thibaud *et al.,* 2014). Choosing a set of algorithms to fit an EENM is, therefore, central to its predictive capability. Here, several algorithms that are used widely in habitat models for wide-ranging marine vertebrates were combined in an ensemble.

Single-algorithm models used here ranked the relative importance of environmental variables differently in both years, yet overall concordance was observed in estimated variable importance between algorithms. Relative variable importance in final EENMs for each year broadly echo the consensus in variable ranking among GAM, RF and BRT model sets. Year-specific EENMs conflicted in the ranking of environmental variable importance. SST, *TFreq* and Depth were ascribed greater importance in the 2011/12 ensemble, whereas the importance of chl-*a* dropped from 2009/10 to 2011/12. This could be attributable to non-stationary processes that govern the foraging responses of grey-headed albatrosses to oceanographic conditions over the scales at which our analysis was focused (Jenouvrier *et al.,* 2005), or indicative of the need for additional environmental data to enhance the capacity of our models to sufficiently capture the foraging seascape experienced by this population.

Concordance in model response curves per environmental variable from single-algorithm models increases confidence in the capacity of these models to detect true responses to environmental conditions. We observed strong concordance between model response curves resulting from GAM, RF and BRT across all environmental variables in both years, and so included these model sets in final EENMs. EENM predictions integrating outputs of several single-algorithm models predicting broadly similar responses could be regarded as preferable to any single-model output in terms of confidence in predictions. Similarly, broadly matching spatial predictions, such as those predicted by GAM, RF and BRT in our analysis, increase confidence in these single-algorithm model outputs, and in the spatial predictions of the final EENMs. This is a key aspect of the utility of the EENM process in enabling the construction of more reliable predictive habitat-based models.

(C) Model Performance

Differences in model performance rankings using alternative metrics (i.e. AUC, TSS, Boyce Index) highlight the potential effect of choice of performance metric on model selection for EENM construction. There is, to our knowledge, no current consensus on which performance

445 metric would be preferable in this context, although the reliability of AUC has been heavily
446 criticised (Boyce *et al.*, 2022; Lobo *et al.,* 2008). The TSS is robust and independent of
447 sample size (prevalence), unlike the commonly used *kappa* statistic (Allouche *et al.,* 2006).
448 As TSS is implemented in the R Package 'biomod2' framework, we chose this metric over
449 AUC for model selection for EENM. We also implemented the Boyce Index as a comparative
450 measure of model performance (Boyce *et al.,* 2002; Hirzel *et al.,* 2006). As with all
451 movement datasets, our response variable is strictly presence-only, and so a presence-only
452 model evaluation metric is likely more appropriate than a presence-absence metric such as
453 AUC or TSS. However, we note that the use of multiple performance metrics in EENM
454 construction and evaluation, and comparison between these metrics, is clearly preferable to
455 any single metric (Allouche *et al.,* 2006, Jiménez-Valverde 2012, Thibaud *et al.,* 2014).
456 EENMs were selected as the best performing models in both years using the Boyce Index and
457 AUC methods, indicating that averaging the outputs of several single-algorithm models into
458 an ensemble has improved predictive capacity in our test case.
459
460 Our exploration of the utility of EENM in this context highlights the capacity of the technique
461 for comparing among the predictions of single-algorithm models and selecting the best
462 performing models for a particular dataset or application. A final model can be selected from
463 among the candidate EENMs and single-algorithm outputs. For example, taking a
464 conservative approach, we excluded MaxEnt from final EENMs, improving performance and
465 increasing confidence in predictions. EENM is useful for excluding strong bias and
466 smoothing over weaker biases in different model predictions. Our results exemplify the
467 potential of EENM for use with movement data in identifying predictable foraging habitats
468 for wide-ranging marine vertebrates over broad scales.
469
470 (B) EENM Extrapolative Performance
471
472 Ecological niche models constructed and validated over the same spatial and temporal extent
473 can show limited transferability in space and time (Randin *et al.,* 2006, Torres *et al.,* 2015).
474 While we did not have sufficient movement data to investigate transferability through space,
475 the extrapolative performance of our EENMs across the two years of this study was generally
476 good, although the 2011/12 ensemble performed better than that built for 2009/10 (2009/10,
477 AUC=0.9107, TSS=0.5194, Boyce Index=0.8536; 2011/12, AUC=0.9281, TSS=0.6630,
478 Boyce Index=0.9348). Changes in the performance of ensembles extrapolated across years are
479 indicative of poor transferability through time, because of variation in animal-environment

interactions or, more probably, the failure of models to fully capture the drivers of these interactions.

Further tests of EENM extrapolative performance through space and time, for example to other populations of the same species (e.g. Torres *et al.,* 2015), or through multiple years in the same region, are necessary to ascertain true extrapolative capabilities. Moreover, the multi-scale periodicity of oceanographic variability in the region (e.g. decadal-scale Southern Ocean Oscillation Index) is likely to influence extrapolative capabilities (e.g. Jenouvrier *et al*., 2005). Some key questions remain: for example, after how many years is the extrapolative performance of a year-specific model likely to fade? How do predictable habitats over decadal timescales align with predictable habitats on inter-annual timescales? Future work should investigate the degree of variability within and between years in prevailing oceanographic conditions and preferred foraging areas if these techniques are to prove valuable for predicting population-level responses to climate-driven ecosystem change.

Nevertheless, ensemble ecological niche models (EENMs) can incorporate differing predictions from species-habitat models fitted using alternative algorithms, where they are implemented with awareness of technical limitations (Marmion *et al.,* 2009, Oppel *et al.,* 2012). By better incorporating uncertainty, the output of EENMs provide a robust basis for recommendations relating to the conservation and management of marine vertebrate populations, particularly those of conservation concern

(A) Acknowledgements

We are grateful to the fieldworkers who assisted with the tracking studies at Bird Island. The authors thank Drs. Stephen C. Votier and Beth Scott for useful discussions. This study represents a contribution to the Ecosystems component of the British Antarctic Survey Polar Science for Planet Earth Programme, funded by the Natural Environment Research Council.

(A) References

Awkerman, J.A., Fukuda, A., Higuchi, H., & Anderson, D.J. (2005). Foraging activity and submesoscale habitat use of waved albatrosses *Phoebastria irrorata* during chick-brooding period. *Marine Ecology Progress Series*, **291**, 289-300.

Agreement on the Conservation of Albatrosses and Petrels (ACAP. (2009). ACAP Species assessment: Grey-headed Albatross *Thalassarche chrysostoma*. Downloaded from http://www.acap.aq on 17 November 2014

Allouche O., Tsoar A., Kadmon R. (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology* **43**, 1223-1232

Araújo M.B., Guisan A. (2006) Five (or so) challenges for species distribution modelling. Journal of *Biogeography* **33**, 677-1688

Araújo M.B., New M. (2007) Ensemble forecasting of species distributions. *Trends in Ecology & Evolution* **22**, 42-47

Ballard G., Jongsomjit D., Veloz S.D., Ainley D.G. (2012) Coexistence of mesopredators in an intact polar ocean ecosystem: The basis for defining a Ross Sea marine protected area. *Biological Conservation* **156**, 72-82

Barbet-Massin, M., et al. (2012). Selecting pseudo-absences for species distribution models: how, where and how many? *Methods in Ecology and Evolution* **3**, 327-338

Belkin I.M., Cornillon P.C., Sherman K. (2009) Fronts in large marine ecosystems. *Progress in Oceanography* **81,** 223-236

Bost C.A., Cotté C., Bailleul F., Cherel Y., Charrassin J.B., Guinet C., Ainley D.G., Weimerskirch H. (2009) The importance of oceanographic fronts to marine birds and mammals of the southern oceans. *Journal of Marine Systems* **78**, 363-376

Broenniman O., Petitpierre B., Randin C., Engler R., Breiner F., D'Amen M., Pellissier L., Pottier J., Pio D., Garcia Mateo R., Di Colla V., Hordijk W., Dubuis A., Scheerer D., Salamin N., Guisan A. (2014) ecospat: Spatial ecology miscellaneous methods. R package version 1.0.

Catry P., Phillips R.A., Phalan B., Silk J.R.D., Croxall J.P. (2004) Foraging strategies of grey-headed albatrosses *Thalassarche chrysostoma*: integration of movements, activity and feeding events. *Marine Ecology Progress Series* **280**, 261-273

Cayula J.-F., Cornillon P.C. (1992) Edge detection algorithm for SST images. *Journal of Atmospheric and Oceanic Technology* **9**, 67-80

Commins M.L., Ansorge I., Ryan P.G. (2014) Multi-scale factors influencing seabird assemblages in the African sector of the Southern Ocean. *Antarctic Science* **26**, 38-48

Croxall, J.P., Silk, J.R.D., Phillips, R.A., Afanasyev, V. and Briggs, D.R. (2005) Global circumnavigations: tracking year-round ranges of non-breeding albatrosses. *Science* **307**, 249-250.

Deppe L., McGregor K.F., Tomasetto F., Briskie J.V., Scofield R.P. (2014) Distribution and predictability of foraging areas in breeding Chatham albatrosses *Thalassarche eremita* in relation to environmental characteristics. *Marine Ecology Progress Series* **498**:287-301

Dias M.P., Granadeiro J.P., Phillips R.A., Alonso H., Catry P. (2010) Breaking the routine: individual Cory's shearwaters shift winter destinations between hemispheres and across ocean basins. *Proceedings of the Royal Society B: Biological Sciences*, rspb20102114

Dray, S & Dufour, A.B. (2007) The ade4 package: implementing the duality diagram for ecologists. *Journal of Statistical Software* **22**, 1-20.

Fauchald P., Erikstad K.E., Skarsfjord H. (2000) Scale-dependent predator-prey interactions: the hierarchical spatial distribution of seabirds and prey. *Ecology* **81**, 773-783

Guisan A., Zimmerman N.E. (2000) Predictive habitat distribution models in ecology. *Ecological Modelling* **135,** 147-186

Hazen, E.L., Suryan, R.M., Santora, J.A., Bograd, S.J., Watanuki, Y., & Wilson, R.P. (2013). Scales and mechanisms of marine hotspot formation. *Marine Ecology Progress Series*, **487**, 177-183.

Hijmans R.J., van Etten J. (2012) raster: Geographic analysis and modelling with raster data. R package version 20-08 http://CRANR-projectorg/package=raster

Hirzel A.H., Le Lay G., Helfer V., Randin C., Guisan A. (2006) Evaluating the ability of habitat suitability models to predict species presences. *Ecological Modelling* **199,** 142-152

Howell, E. A., Hoover, A., Benson, S. R., Bailey, H., Polovina, J. J., Seminoff, J. A., & Dutton, P. H. (2015). Enhancing the TurtleWatch product for leatherback sea turtles, a dynamic habitat model for ecosystem‐based management. *Fisheries Oceanography*.

Huin N., Prince P. (1997) Diving behaviour of the grey-headed albatross. *Antarctic Science* **9,** 243-249

Jenouvrier S., Weimerskirch H., Barbraud C., Park Y. H., Cazelles B. (2005) Evidence of a shift in the cyclicity of Antarctic seabird dynamics linked to climate. *Proceedings of the Royal Society of London B: Biological Sciences*, **272**(1566), 887-895.

Jiménez‐Valverde A. (2012) Insights into the area under the receiver operating characteristic curve (AUC) as a discrimination measure in species distribution modelling. *Global Ecology and Biogeography* **21,** 498-507

Joiris C.R., Dochy O. (2013) A major autumn feeding ground for fin whales, southern fulmars and grey-headed albatrosses around the South Shetland Islands, Antarctica. *Polar Biology* **36**, 1649-1658

Kappes, M.A., Shaffer, S.A., Tremblay, Y., Foley, D.G., Palacios, D.M., Robinson, P.W., Bograd, S.J. & Costa, D.P. (2010). Hawaiian albatrosses track interannual variability of marine habitats in the North Pacific. *Progress in Oceanography*, **86**(1), 246-260.

Lascelles, B.G., Langham, G.M., Ronconi, R.A. & Reid, J.B. (2012) From hotspots to site protection: Identifying Marine Protected Areas for seabirds around the globe. *Biological Conservation,* **156,** 5-14.

Lobo J.M., Jiménez-Valverde A., Real R. (2008) AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography* **17**, 145-151

Louzao M., Pinaud D., Peron C., Delord K., Wiegand T., Weimerskirch H. (2011) Conserving pelagic habitats: seascape modelling of an oceanic top predator. *Journal of Applied Ecology* **48,** 121-132

Luoto M., Virkkala R., Heikkinen R.K. (2007) The role of land cover in bioclimatic models depends on spatial resolution. *Global Ecology and Biogeography* **16,** 34-42

Marmion M., Parviainen M., Luoto M., Heikkinen R.K., Thuiller W. (2009) Evaluation of consensus methods in predictive species distribution modelling. *Diversity and Distributions* **15**, 59-69

Miller P. (2009) Composite front maps for improved visibility of dynamic sea-surface features on cloudy SeaWiFS and AVHRR data. *Journal of Marine Systems* **78,** 327-336

Nel D.C., Lutjeharms J.R.E., Pakhomov E.A., Ansorge I.J., Ryan P.G., Klages N.T.W. (2001) Exploitation of mesoscale oceanographic features by grey-headed albatross *Thalassarche chrysostoma* in the southern Indian Ocean. *Marine Ecology Progress Series* **217,** 15-26

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

Patrick S.C., Weimerskirch H. (2014) Consistency pays: sex differences and fitness consequences of behavioural specialization in a wide-ranging seabird. *Biology Letters* **10,** 20140630

Phillips R.A., Silk J.R.D., Phalan B., Catry P., Croxall J.P. (2004) Seasonal sexual segregation in two Thalassarche albatross species: competitive exclusion, reproductive role specialization or foraging niche divergence? *Proceedings of the Royal Society of London Series B: Biological Sciences* **271,** 1283-1291

Phillips R.A., Silk J.R., Croxall J.P., Afanasyev V., Bennett V.J. (2005) Summer distribution and migration of nonbreeding albatrosses: individual consistencies and implications for conservation. *Ecology* **86,** 2386-2396

Phillips R.A., Silk J.R., Croxall J.P., Afanasyev V. (2006) Year-round distribution of white-chinned petrels from South Georgia: relationships with oceanography and fisheries. *Biological Conservation* **129,** 336-347

Piatt J.F., Wetzel J., Bell K., DeGange A.R., Balogh G.R., Drew G.S., Geernaert T., Ladd C., Byrd G.V. (2006) Predictable hotspots and foraging habitat of the endangered short-tailed albatross (*Phoebastria albatrus*) in the North Pacific: Implications for conservation. *Deep Sea Research Part II: Topical Studies in Oceanography* **53,** 387-398

Pikesley S.K., Maxwell S.M., Pendoley K., Costa D.P., Coyne M.S., Formia A., Godley B.J., Klein W., Makanga-Bahouna J., Maruca S., Ngouessono S., Parnell R.J., Pemo-Makaya E., Witt M.J. (2013) On the front line: integrated habitat mapping for olive ridley sea turtles in the southeast Atlantic. *Diversity and Distributions* **19**, 1518-1530

Randin C.F., Dirnböck T., Dullinger S., Zimmermann N.E., Zappa M., Guisan A. (2006) Are niche-based species distribution models transferable in space? *Journal of Biogeography* **33,** 1689-1703

Raymond, B., Lea, M.A., Patterson, T., Andrews-Goff, V., Sharples, R., Charrassin, J.B., Cottin, M.., Emmerson, L., Gales, N., Gales, R., Goldsworthy, S.D., Harcourt, R., Kato, A., Kirkwood, R., Lawton, K., Ropert-Coudert, Y., Southwell, C., ven den Hoff, J., Wienecke, B., Woehler, E.J., Wotherspoon, S. & Hindell, M. A. (2015). Important marine habitat off east Antarctica revealed by two decades of multi-species predator tracking. *Ecography* **38**, 121-129

Reid K., Croxall J., Prince P. (1996) The fish diet of black-browed albatross *Diomedea melanophris* and grey-headed albatross *D. Chrysostoma* at South Georgia. *Polar Biology* **16,** 469-477

Rodhouse P., Prince P., Clarke M., Murray A. (1990) Cephalopod prey of the grey-headed albatross *Diomedea chrysostoma*. *Marine Biology* **104,** 353-362

Rodhouse P.G., Boyle P.R. (2010) Large aggregations of pelagic squid near the ocean surface at the Antarctic Polar Front, and their capture by grey-headed albatrosses. *ICES Journal of Marine Science* **67,** 1432-1435

Rodhouse P.G., White M.G. (1995) Cephalopods occupy the ecological niche of epipelagic fish in the Antarctic Polar Frontal Zone. *The Biological Bulletin* **189,** 77-80

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. *Journal of The Royal Society Interface*, **11**(100), 20140679

Scheffer, A., Bost, C.A., & Trathan, P.N. (2012). Frontal zones, temperature gradient and depth characterize the foraging habitat of king penguins at South Georgia. *Marine Ecology Progress Series*, **465**, 281-297.

Storch I. (2002) On spatial resolution in habitat models: can small-scale forest structure explain capercaillie numbers? *Conservation Ecology* **6,** 6

Thibaud E., Petitpierre B., Broennimann O., Davison A.C., Guisan A. (2014) Measuring the relative effect of factors affecting species distribution model predictions. *Methods in Ecology and Evolution* **5,** 947-955

Thuiller, W., Lafourcade, B., Engler, R., & Araújo, M. B. (2009). BIOMOD–a platform for ensemble forecasting of species distributions. *Ecography*, **32**(3), 369-373.

Thuiller W., Georges D., Engler R. (2014) biomod2: Ensemble platform for species distribution modeling. R package version 3.1-64. http://CRAN.R-project.org/package=biomod2

Torres L.G., Sutton P., Thompson D.R., Delord K., Weimerskirch H., Sagar P.M., Sommer E., Dilley B.J., Ryan P.G., Phillips R.A. (2015) Poor transferability of species distribution models for a pelagic predator, the Grey Petrel, indicates contrasting habitat preferences across ocean basins. *PLoS ONE* **10**(3): e0120014. doi: 10.1371/journal.pone.0120014

Trathan P., Brandon M., Murphy E. (1997) Characterization of the Antarctic Polar Frontal Zone to the north of South Georgia in summer 1994. *Journal of Geophysical Research: Oceans* (1978–2012) **102**, 10483-10497

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

Wakefield, E.D., Phillips, R.A. & Belchier, M. (2012) Foraging black-browed albatrosses target waters overlaying moraine banks-a consequence of upward benthic-pelagic coupling? *Antarctic Science,* **24,** 269-280.

Weimerskirch H. (2007) Are seabirds foraging for unpredictable resources? *Deep Sea Research Part II: Topical Studies in Oceanography* **54,** 211-223

Weimerskirch H., Bartle J.A., Jouventin P., Stahl J.C. (1988) Foraging ranges and partitioning of feeding zones in three species of southern albatrosses. *Condor* **90***,* 214-219

Weimerskirch H., Mougey T., Hindermeyer X. (1997) Foraging and provisioning strategies of black-browed albatrosses in relation to the requirements of the chick: natural variation and experimental study. *Behavioral Ecology* **8,** 635

Weimerskirch H., Salamolard M., Sarrazin F., Jouventin P. (1993) Foraging strategy of wandering albatrosses through the breeding season: a study using satellite telemetry. *The Auk* **110**:325-342

Xavier J., Croxall J., Trathan P., Wood A. (2003) Feeding strategies and diets of breeding grey-headed and wandering albatrosses at South Georgia. *Marine Biology* **143,** 221-232

Xavier J., Louzao M., Thorpe S., Ward P., Hill C., Roberts D., Croxall J.P, Phillips R.A. (2013) Seasonal changes in the diet and feeding behaviour of a top predator indicate a flexible response to deteriorating oceanographic conditions. *Marine Biology* **160,** 1597-1606

Biosketch

This research was carried out by an inter-disciplinary team of authors from multiple institutions, each with expertise in linking animal movements and behaviours to oceanographic conditions in dynamic marine systems. Author contributions: K.S., P.M., R.P & S.I. conceived the ideas; R.P. collected the tracking data; K.S. and P.M. prepared the remotely-sensed data; K.S. analysed the data and led manuscript preparation; P.M.,E.H., S.B. and R.P. contributed to manuscript preparation and edits.

Figures & Tables

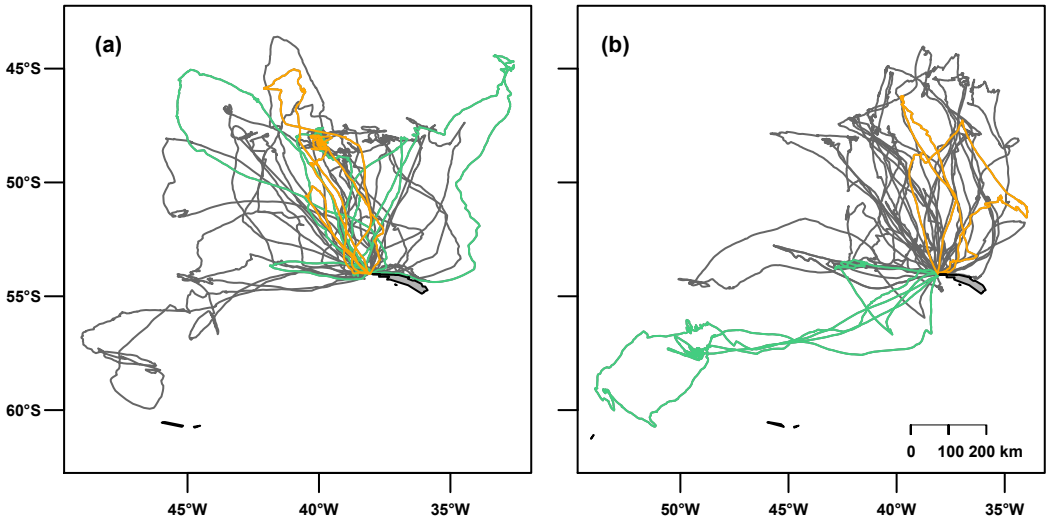

Figure 1 GPS tracking of grey-headed albatrosses (GHA) from Bird Island, South Georgia. Trips used to identify the spatial distribution of foraging events during the (a) 2009/10 (n=25) and (b) 2011/12 (n=30) breeding seasons (brood-guard phase). Birds for which sexes are known are highlighted in orange for female (n=3, 2009/10, n=2, 2011/12) and green for male (n=5, 2009/10; n=3, 2011/12).

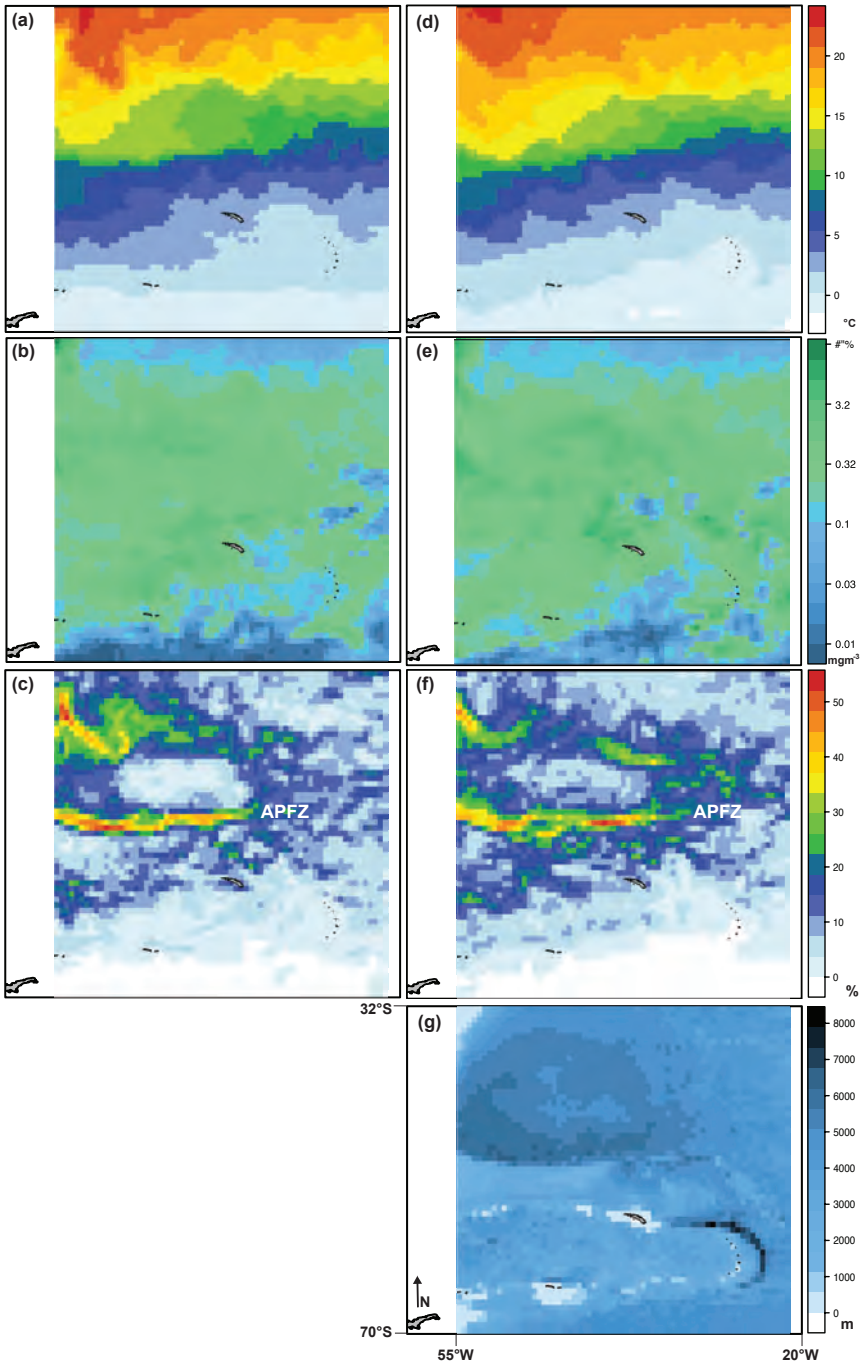

Figure 2 Environmental data layers for brood-guard period (end December – end January). Dynamic variables, (a) Sea Surface Temperature (SST, °C; monthly median composite) for 2009/10, (b) Chlorophyll-a (chl-a, mg m^{-3}; monthly median composite; log-transformed), for 2009/10 (c) Thermal front frequency (Tfreq, % time; 0.4°C front detection threshold; monthly synoptic composite) for 2009/10. (d)-(f) Dynamic variables for 2011/12. (g) GEBCO Depth (30 arc-second resolution).

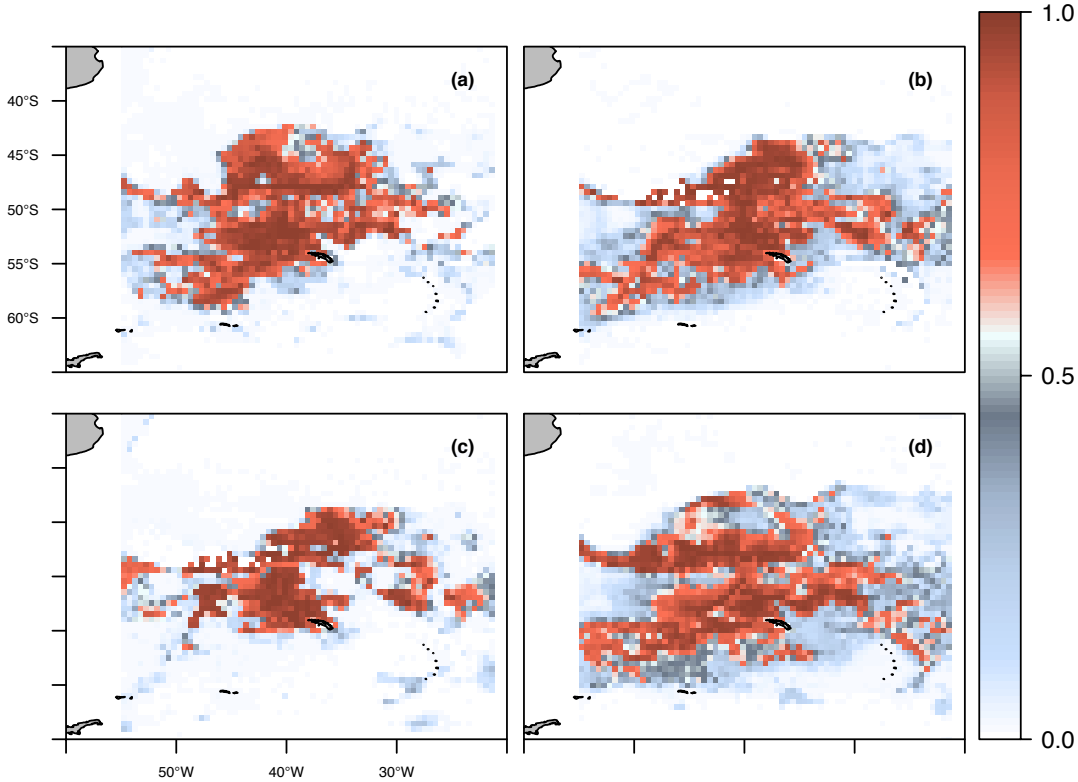

Figure 3 Spatial predictions of ensemble ecological niche models (EENMs), and cross-validation among years. Spatial predictions of final EENM (weighted mean, removal of MaxEnt predictions) for (a) 2009/10 and (b) 2011/12. Cross-validation of (c) 2009/10 EENM onto 2011/12 environmental conditions and (d) 2011/12 EENM onto 2009/10 environmental conditions. Spatial predictions displayed as Habitat Suitability Index (HSI) per grid cell, scaled from 0 to 1. Greater similarity between (a), (b) and (c),(d) indicates better EENM transferability among years.

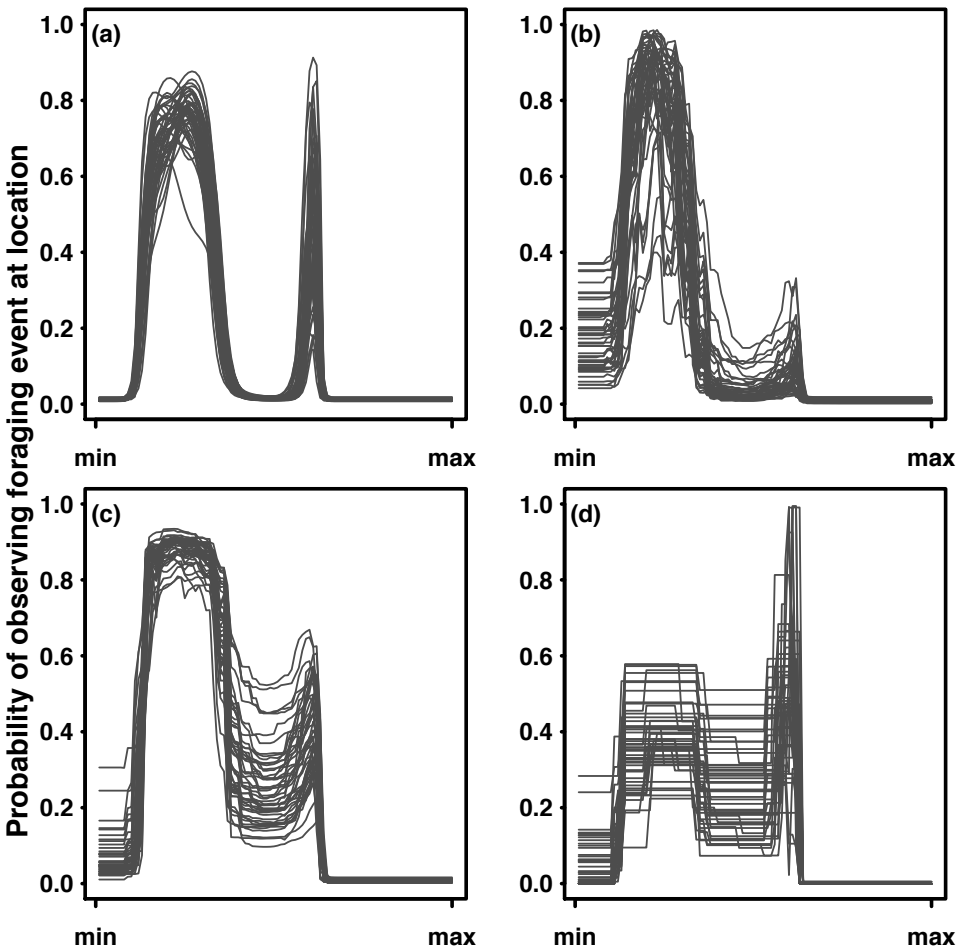

Figure 4 Model Response Curves for SST in 2011/12 model sets, per algorithm, (a) GAM, (b) RF, (c) BRT, (d) MaxEnt

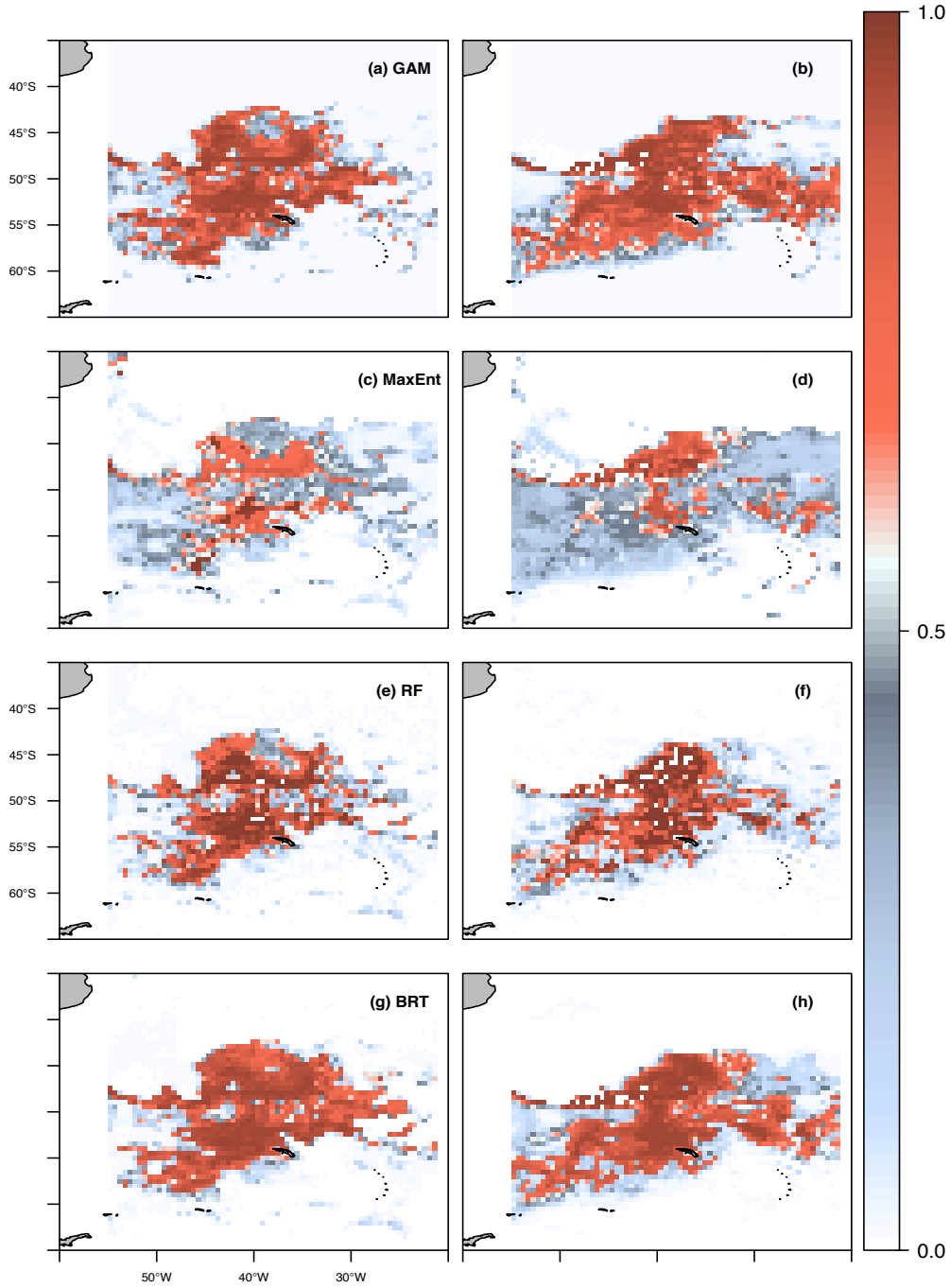

Figure 5 Spatial predictions of ecological niche models per algorithm, (a) Generalised Additive Models, GAM, 2009-10 (b) GAM, 2011/12; (c) Maximum Entropy, MaxEnt, 2009/10, (d) 2011/12; (e) Random Forest, 2009/10, (f) 2011/12; (g) Boosted Regression Trees, 2009/10, (h) 2011/12. Spatial predictions displayed as Habitat Suitability Index (HSI) per grid cell, scaled from 0 to 1 (mean over all model runs, n=50 per algorithm).

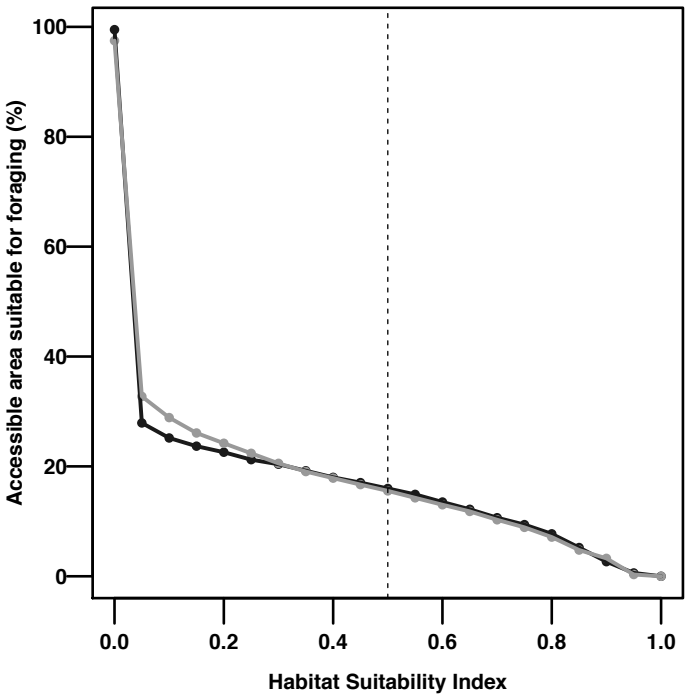

Figure 6 Percentage of area accessible during brood-guard phase (estimated using whole-dataset maximum displacement from colony) containing oceanographic conditions suitable for foraging against EENM-predicted Habitat Suitability Index (HSI). 2009/10 EENM (weighted mean) as black line; 2011/12 in grey.

Figure S1 Model Response Curves for Chl-*a* in 2011/12 model sets, per algorithm, (a) GAM, (b) RF, (c) BRT, (d) MaxEnt

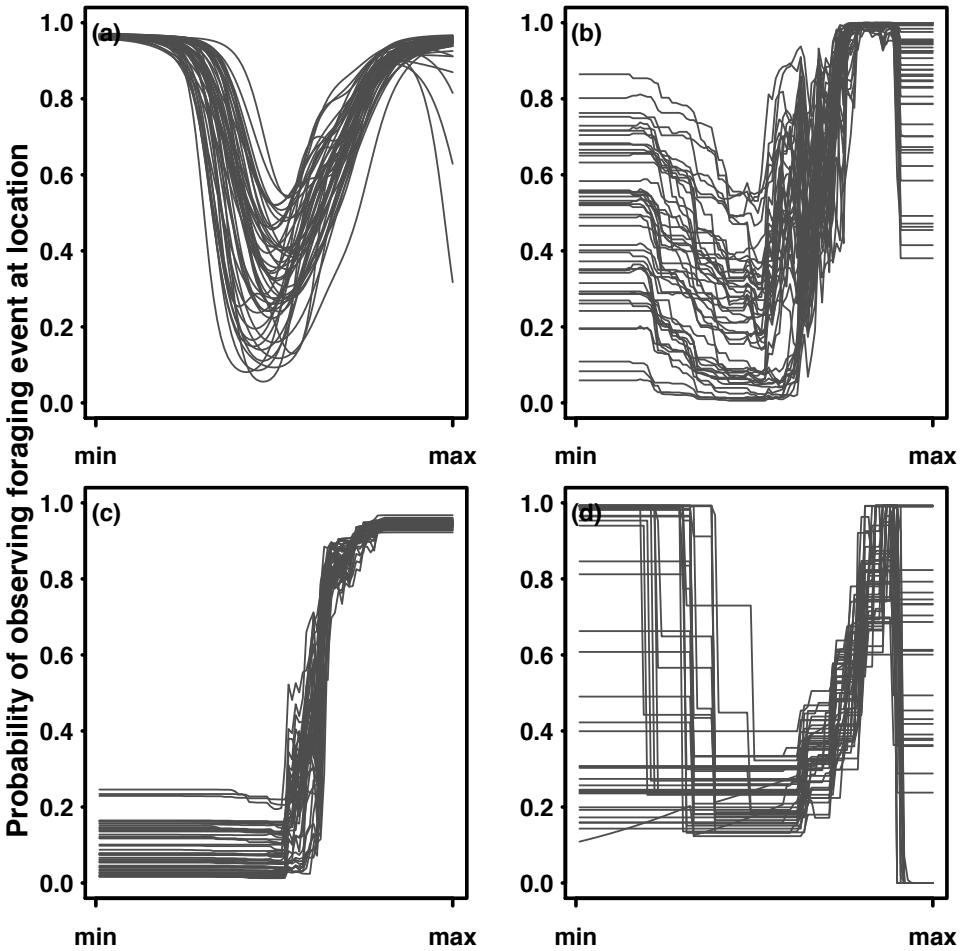

Figure S2 Model Response Curves for *TFreq* in 2011/12 model sets, per algorithm, (a) GAM, (b) RF, (c) BRT, (d) MaxEnt

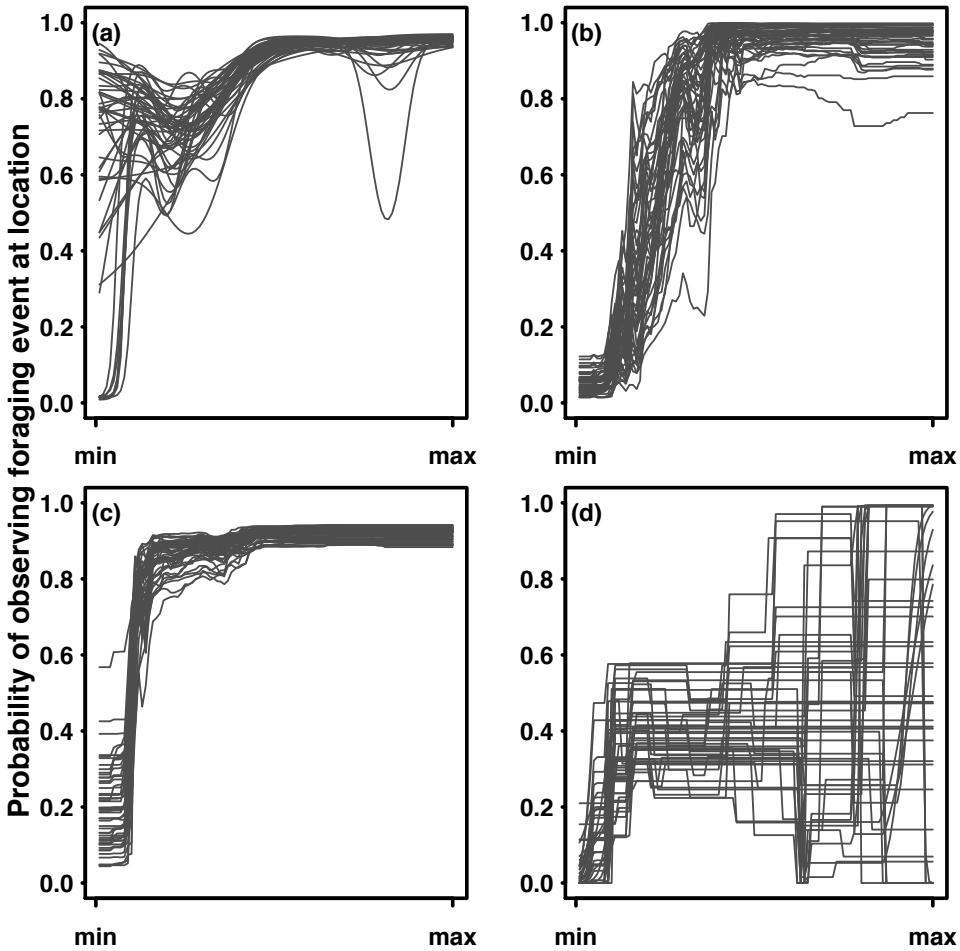

Figure S3 Model Response Curves for depth in 2011/12 model sets, per algorithm, (a) GAM, (b) RF, (c) BRT, (d) MaxEnt

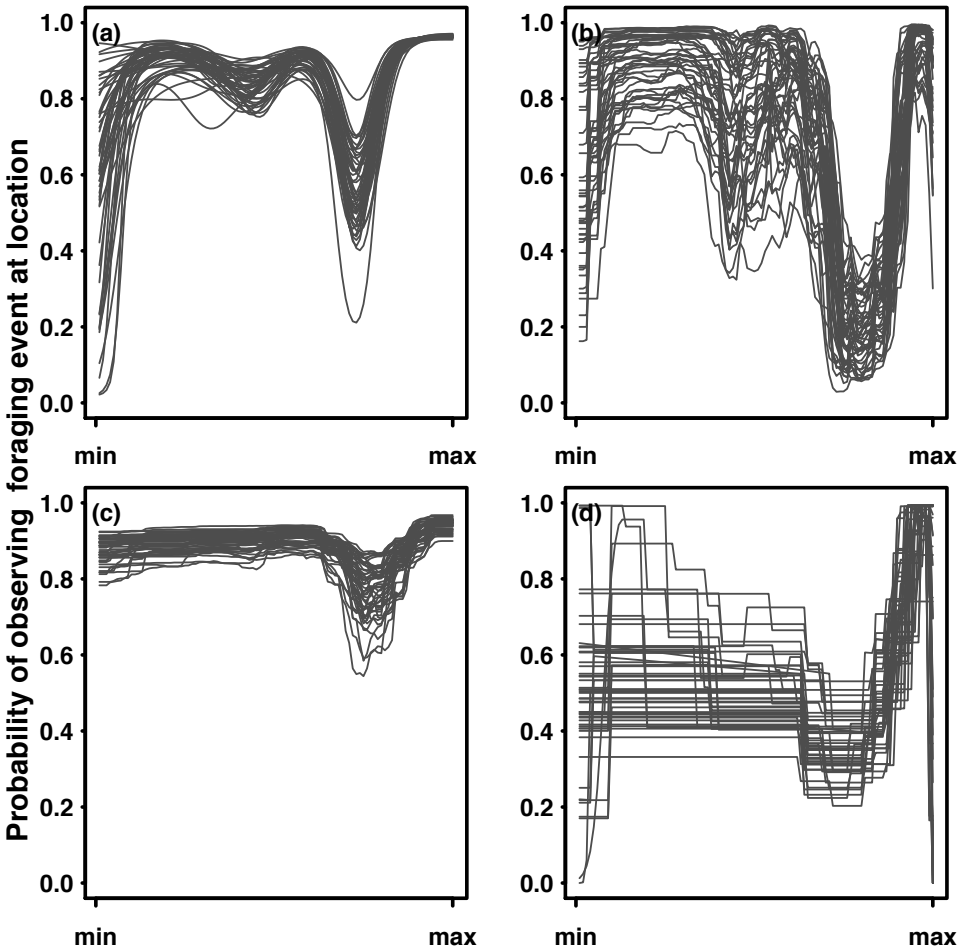

Table 1 Variable Importance (Mean over all model sets per algorithm), scaled as Relative Importance of Contribution to model Coefficients (RICC), from 0 to 1. Variable importance rankings in brackets

	Variable Importance, 2009/10				**Variable Importance, 2011/12**			
	SST	**Chl-*a***	**TFreq**	**Depth**	**SST**	**Chl-*a***	**TFreq**	**Depth**
GAM	0.61396 (1)	0.4570 (2)	0.06512 (4)	0.17284 (3)	0.92174 (1)	0.09860 (3)	0.07752 (4)	0.16574 (2)
MaxEnt	0.45498 (2)	0.48992 (1)	0.06060 (4)	0.12338 (3)	0.55658 (1)	0.21478 (3)	0.31830 (2)	0.18928 (4)
RF	0.46120 (2)	0.52012 (1)	0.08466 (4)	0.16598 (3)	0.51792 (1)	0.27812 (2)	0.24914 (3)	0.20358 (4)
BRT	0.5644 (1)	0.56014 (2)	0.01672 (4)	0.05316 (3)	0.59350 (1)	0.29776 (2)	0.22872 (3)	0.0805 (4)
EENM	**0.577 (2)**	**0.585 (1)**	**0.037 (4)**	**0.086 (3)**	**0.744 (1)**	**0.150 (3)**	**0.155 (2)**	**0.100 (4)**

Table 2 Model Performance Metrics (Mean over all model sets per algorithm). Area Under Receiver Operating Characteristic Curve (AUC) scaled 0 to 1; True Skill Statistic (TSS) scaled 0 to 1; Boyce Index scaled -1 to +1. Highest-scoring model for each performance metric highlighted in bold. EENM rows have metrics for final EENM, without MaxEnt (black) and EENM with MaxEnt (grey). Performance rankings per metric in brackets.

	Model Evaluation, 2009/10			**Model Evaluation, 2011/12**		
Model Set	**AUC**	**TSS**	**Boyce Index**	**AUC**	**TSS**	**Boyce Index**
GAM	0.9421 (3)	0.8237 (2)	0.9213 (2)	0.9372 (3)	0.7835 (3)	0.8943 (3)
MaxEnt	0.9276 (4)	0.7740 (4)	0.9300 (1)	0.9101 (4)	0.7184 (4)	0.9051 (1)
RF	0.9523 (1)	**0.8277** (1)	0.8329 (3)	0.9563 (1)	**0.8283** (1)	0.8998 (2)
BRT	0.9444 (2)	0.8176 (3)	0.7130 (4)	0.9418 (2)	0.7843 (2)	0.8615 (4)
EENM	**0.9547** 0.9479	0.7914 0.7514	**0.9512** 0.8990	**0.9610** 0.9591	0.7871 0.7791	**0.9656** 0.9626
EENM Extrapolation	0.9107 0.9038	0.5194 0.5188	0.8536 0.7138	0.9281 0.9267	0.6630 0.6208	0.9358 0.9540

Table S1 Variable importance per iteration of control locations, 2009/10. Mean importance of environmental variables (Sea Surface Temperature, SST; Chlorophyll-a, chl-a; thermal front frequency, Tfreq; depth) over model runs (10-fold cross-validation) per iteration of control locations, for each model algorithm (Generalised Additive Models, GAM; Maximum Entropy modelling, MaxEnt; Random Forest, RF; Boosted Regression Trees, BRT). Mean of Relative Importance to the model Coefficients (RICC) metric over successive iteration of control locations.

Control Location Iteration	**Model Algorithm**	**Variable Importance** (mean over 10 runs per pseudo-absence iteration)			
		SST	**Chl-*a***	**TFreq**	**Depth**
1	**GAM**	0.6160	0.4646	0.0721	0.1762
	MaxEnt	0.4840	0.5192	0.0784	0.1140
	RF	0.4746	0.5360	0.1122	0.1285
	BRT	0.5679	0.5618	0.0139	0.0396
2	**GAM**	0.6089	0.4589	0.0690	0.1503
	MaxEnt	0.4779	0.4031	0.1327	0.2149
	RF	0.4523	0.5474	0.0694	0.1651
	BRT	0.5808	0.5655	0.0146	0.0447
3	**GAM**	0.5992	0.4509	0.0430	0.1572
	MaxEnt	0.4449	0.4771	0.0345	0.1019
	RF	0.4645	0.5094	0.0891	0.1683
	BRT	0.5559	0.5690	0.0244	0.0417
4	**GAM**	0.6040	0.4803	0.0910	0.1544
	MaxEnt	0.3937	0.5321	0.0364	0.0852
	RF	0.4614	0.5267	0.0743	0.1499
	BRT	0.5470	0.5718	0.0131	0.0544
5	**GAM**	0.6417	0.4303	0.0505	0.2261
	MaxEnt	0.4744	0.5181	0.0210	0.1009
	RF	0.4532	0.4811	0.0783	0.2181
	BRT	0.5704	0.5326	0.0176	0.0854
mean of means	**GAM**	**0.61396**	**0.4570**	**0.06512**	**0.17284**
	MaxEnt	**0.45498**	**0.48992**	**0.06060**	**0.12338**
	RF	**0.46120**	**0.52012**	**0.08466**	**0.16598**
	BRT	**0.5644**	**0.56014**	**0.01672**	**0.05316**

Table S2 Variable importance per iteration of control locations, 2011/12. Mean importance of environmental variables (Sea Surface Temperature, SST; Chlorophyll-a, chl-a; thermal front frequency, Tfreq; depth) over model runs (10-fold cross-validation) per iteration of control locations, for each model algorithm (Generalised Additive Models, GAM; Maximum Entropy modelling, MaxEnt; Random Forest, RF; Boosted Regression Trees, BRT). Mean of Relative Importance to the model Coefficients (RICC) metric over successive iteration of control locations.

Control Location Iteration	Model Algorithm	Variable Importance (mean over 10 runs per pseudo-absence iteration)			
		SST	Chl-*a*	TFreq	Depth
1	GAM	0.9427	0.0941	0.0669	0.1390
	MaxEnt	0.5170	0.2031	0.4323	0.1567
	RF	0.4893	0.2765	0.2358	0.1887
	BRT	0.5819	0.2770	0.2378	0.0778
2	GAM	0.9277	0.0861	0.0580	0.1997
	MaxEnt	0.5942	0.2101	0.2982	0.1814
	RF	0.5094	0.2904	0.2906	0.1838
	BRT	0.5621	0.3188	0.2943	0.0681
3	GAM	0.9310	0.1234	0.0423	0.1522
	MaxEnt	0.4932	0.1673	0.1621	0.2250
	RF	0.5145	0.2910	0.2369	0.1892
	BRT	0.6279	0.3018	0.1764	0.0690
4	GAM	0.8950	0.0873	0.1362	0.1821
	MaxEnt	0.7395	0.3093	0.5689	0.1517
	RF	0.5737	0.2619	0.2485	0.2424
	BRT	0.6172	0.2780	0.2186	0.1113
5	GAM	0.9123	0.1021	0.0842	0.1557
	MaxEnt	0.4390	0.1841	0.1300	0.2316
	RF	0.5027	0.2708	0.2339	0.2138
	BRT	0.5784	0.3132	0.2165	0.0763
mean of means	**GAM**	**0.92174**	**0.09860**	**0.07752**	**0.16574**
	MaxEnt	**0.55658**	**0.21478**	**0.31830**	**0.18928**
	RF	**0.51792**	**0.27812**	**0.24914**	**0.20358**
	BRT	**0.59350**	**0.29776**	**0.22872**	**0.0805**

Table S3 Model performance metrics per iteration of control locations, 2009/10. Evaluation metrics (Area Under Receiver Operating Curve, AUC; True Skill Statistic, TSS). Mean over model runs (10-fold cross-validation) per iteration of control locations, for each model algorithm (Generalised Additive Models, GAM; Maximum Entropy modelling, MaxEnt; Random Forest, RF; Boosted Regression Trees, BRT).

Control Location Iteration	**Evaluation Metric**	**Model Algorithm** (mean over 10 runs per Pseudo-Absence iteration)			
		GAM	**MaxEnt**	**RF**	**BRT**
1	AUC	0.9362	0.9166	0.9511	0.9407
	TSS	0.8172	0.7599	0.8273	0.8094
	Boyce Index	0.9155	0.9391	0.8635	0.681
2	AUC	0.9520	0.9358	0.9641	0.9552
	TSS	0.8383	0.7967	0.8632	0.8452
	Boyce Index	0.9174	0.9343	0.8215	0.6572
3	AUC	0.9593	0.9287	0.9431	0.9374
	TSS	0.8209	0.7871	0.8164	0.8110
	Boyce Index	0.9154	0.9695	0.8195	0.6966
4	AUC	0.9494	0.9315	0.9604	0.9518
	TSS	0.8466	0.7749	0.8352	0.8256
	Boyce Index	0.9164	0.9624	0.8336	0.7599
5	AUC	0.9337	0.9253	0.9428	0.9369
	TSS	0.7956	0.7514	0.7963	0.7967
	Boyce Index	0.9419	0.8436	0.8263	0.7701
Mean of means	**AUC**	**0.9421**	**0.9276**	**0.9523**	**0.9444**
	TSS	**0.8237**	**0.7740**	**0.8277**	**0.8176**
	Boyce Index	**0.9213**	**0.9300**	**0.8329**	**0.7130**

Table S4 Model performance metrics per iteration of control locations, 2011/12. Evaluation metrics (Area Under Receiver Operating Curve, AUC; True Skill Statistic, TSS). Mean over model runs (10-fold cross-validation) per iteration of control locations, for each model algorithm (Generalised Additive Models, GAM; Maximum Entropy modelling, MaxEnt; Random Forest, RF; Boosted Regression Trees, BRT).

Control Location Iteration	**Evaluation Metric**	**Model Algorithm** (mean over 10 runs per Pseudo-Absence iteration)			
		GAM	**MaxEnt**	**RF**	**BRT**
1	AUC	0.9311	0.9058	0.9461	0.9334
	TSS	0.7824	0.7214	0.8111	0.7745
	Boyce Index	0.9125	0.9484	0.9040	0.8692
2	AUC	0.9344	0.9055	0.9551	0.9418
	TSS	0.7748	0.7019	0.8196	0.7810
	Boyce Index	0.8638	0.8955	0.9065	0.8397
3	AUC	0.9463	0.9126	0.9658	0.9496
	TSS	0.7892	0.7136	0.8345	0.7842
	Boyce Index	0.8778	0.8398	0.8697	0.8447
4	AUC	0.9365	0.9122	0.9581	0.9403
	TSS	0.7871	0.7399	0.8394	0.7908
	Boyce Index	0.8968	0.9237	0.8989	0.8564
5	AUC	0.9376	0.9143	0.9565	0.9437
	TSS	0.7842	0.7154	0.8369	0.7908
	Boyce Index	0.9206	0.9181	0.9197	0.8976
Mean of means	**AUC**	**0.9372**	**0.9101**	**0.9563**	**0.9418**
	TSS	**0.7835**	**0.7184**	**0.8283**	**0.7843**
	Boyce Index	**0.8943**	**0.9051**	**0.8998**	**0.8615**

Table S5 Model Parameterisation settings

GAM	package = ‘mgcv’, family = ‘binomial’ (link = ‘logit’), type = ‘s’ (spline-based smooth), model formula =
RF	number of trees = 500, node size = 5; Boosted Regression Trees
BRT	distribution = ‘bernoulli’, number of trees = 2500, shrinkage = 0.001, bag fraction = 0.5, train fraction = 1, cross-validation folds = 3
MaxEnt	maximum training iterations = 200, linear/quadratic/product/threshold/hinge features enabled, default prevalence = 0.5