Ecography

ECOG-01021
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., van den Hoff, J., Wienecke, B., Woehler, E. J., Wotherspoon, S. and Hindell, M. A. 2014. Important marine habitat off East Antarctica revealed by two decades of multi-species predator tracking. – Ecography doi: 10.1111/ecog.01021

Supplementary material

Appendix 1

Methods

The preferred habitat characteristics of each species were identified using a habitat selectivity approach, by modelling the environmental characteristics of the locations where the animals were recorded to be present (utilized habitat) relative to the environmental characteristics of the areas that they could potentially have used (available habitat). This is analogous to the presence-background approach to species distribution modelling (e.g. Phillips et al. 2009). The tracking data provided presence locations (i.e. areas utilized by the animals), but not absences. Areas without track locations either represent areas that the animals do not use, or areas that would have been utilized by the animals if a different sample had been acquired (e.g. different animals from the same colony, or animals from a different colony).

Habitat availability was estimated by simulating tracks that were statistically similar to the observed tracks. By simulating such tracks from the known deployment locations, we obtained an indication of where the animals could potentially have travelled if they did not have any preferences in terms of environmental conditions (while still respecting the constraints on their trip duration, travel speed, and departure locations). Similar approaches have previously been used (Wakefield et al. 2010, Žydelis et al. 2011). For each observed trip by each individual animal, 20 simulated trips were computed using a first-order vector autoregressive model:

$$\mathbf{z}(t) = \mathbf{A}\mathbf{z}(t-1) + \boldsymbol{\varepsilon}(t)$$

where $\mathbf{z}(t)$ is the zonal and meridional step length of the track at time *t*: $\mathbf{z}(t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}$, $\mathbf{A}$ is a matrix of coefficients, and $\boldsymbol{\varepsilon}(t) \sim N(\mathbf{0}, \boldsymbol{\Omega})$ (i.e. bivariate normal with zero mean and covariance $\boldsymbol{\Omega}$). Trips were simulated from the fitted model parameters by iteratively stepping from the deployment location, using bivariate normal samples from $N(\mathbf{0}, \boldsymbol{\Omega})$. Each simulated trip was rotated by a random angle so that the surrounding habitat in all directions was potentially visited, with a land mask to constrain tracks to oceanic locations.

20 simulated trips was found by trial and error to give a reasonable compromise between data set size and stability of results (i.e. with too few simulated tracks, the model predictions tended to be variable from one run to the next).

Environmental conditions $\boldsymbol{w}$ were collated at each point on the observed and simulated tracks. We define the response variable q to indicate habitat use (i.e. q=1 indicates that a certain habitat was utilized, and q=0 indicates that it was not). Ideally, we would like to be able to estimate p(q=1|$\boldsymbol{w}$), the probability that the species utilizes a location given its habitat $\boldsymbol{w}$. However, as noted above, the observed tracks only provide information about q=1 (i.e. areas that were utilized), and not q=0 (areas not utilized). Instead, we fit a binomial model with response variable s, where s takes the value 1 if the point is from an observed track, or 0 for a simulated track. This model is then used to predict the probability p(s=1|$\boldsymbol{w}$) that a point came from an observed track given the environmental conditions $\boldsymbol{w}$ at the location. Simulated tracks provide information about the available habitat, and the observed tracks about utilized habitat. The probabilities p(s=1|$\boldsymbol{w}$) can be therefore interpreted as a description of habitat use relative to availability (2008, 2010). However, it is important to note that these probabilities are not interpretable as direct estimates of the probabilities p(q=1|$\boldsymbol{w}$). Under mild assumptions, p(s=1|$\boldsymbol{w}$) can be shown to be nonlinearly but monotonically related to p(q=1|$\boldsymbol{w}$), with the relationship being dependent on the prevalence of the species (i.e. P(q=1), the probability that an individual utilizes a randomly sampled site) (Aarts et al. 2008, Phillips et al. 2009). Direct comparison of p(s=1|$\boldsymbol{w}$) between species with differing prevalences is therefore not meaningful.

However, consider a discrete set of grid cells, obtained by applying a threshold to the p(s=1|$\boldsymbol{w}$). That is, we identify grid cells with p(s=1|$\boldsymbol{w}$) values greater than some threshold. By appropriate selection of the threshold value, the area covered by these grid cells can be chosen to represent a certain fraction of the total area (say, 10% — thus identifying the most important 10% of geographic space associated with a particular model output). The monotonic relationship between p(s=1|$\boldsymbol{w}$) and p(q=1|$\boldsymbol{w}$) means that exactly the same grid cells would be obtained by applying a threshold to the p(q=1|$\boldsymbol{w}$), albeit with a different threshold value. Thus, given the predicted habitat preferences for a species, we can partition the study region into areas of decreasingly important habitat by thresholding the p(s=1|$\boldsymbol{w}$) estimates at decreasing levels. This

yields a transformed prediction map, wherein each value is a habitat importance percentile (by area). These percentile values can be compared across species, allowing us to quantify the degree of overlap between the different species.

Models were first fitted using boosted regression trees (De'ath 2007, Elith et al. 2008) as an exploratory step, to assist in the identification of relevant predictor variables. Final models were fitted as binomial generalized additive models (GAMs) with logit link. Variable selection was guided by expert knowledge of the species in question, previously published research, model accuracy, examination of the fitted component smooth terms, and the spatial pattern of those component terms.

The tracking data display serial correlation (successive locations come from the same individual animal). This has two important consequences. The first is that the smoothness estimation of model terms in GAM is likely to under-smooth. Appropriate smoothness of terms was enforced here by specifying the degrees of freedom of individual model terms. The second issue is that standard methods for estimating parameter uncertainty assume independence in the model residuals, and so in the presence of serial correlation will give uncertainty estimates that are too optimistic. Consequently, standard methods for model selection (e.g. using Akaike's Information Criterion, or likelihood-ratio tests) are unreliable. Instead, a cross-validation procedure was used for model selection and assessment. Individual animals were randomly assigned to one of ten data folds. Each model was trained on nine folds and tested on the remaining one, withholding each fold in turn. Predictive performance was then aggregated across the ten sets of results. The area under the receiver operating curve (AUC) was used as the index of predictive performance. AUC values were calculated for each individual animal in the testing data fold. The overall AUC performance for a given model is reported as the mean and standard deviation of the AUC results across all individual animals. Cross-validating by individual animal is important because the reported accuracies reflect inter-individual variability (i.e. the ability of the fitted model to predict the habitat preference of a previously-unseen individual animal).

The fitted model for each species was used to predict the habitat preference for that species across the entire region of interest. Predictions were made for each month November–February, and then averaged to give a single composite summer prediction.

Estimates of uncertainty in spatial predictions were calculated by a similar cross-validation method to the one described above. For a given set of predictor variables, cross-validating by individual gives an estimate of the uncertainty in the model predictions arising from individual variability. However, the predictor variables available for large-scale habitat modelling are by necessity either remote-sensed or modelled, in order to provide regional coverage at appropriate spatial and temporal resolution. Typically, these predictors are proxies for the actual processes and conditions that influence the behaviour of the animals. For example, there is no direct estimate of food availability for any of the species. Even parameters that might be considered directly relevant (e.g. satellite-derived sea ice concentrations as an indicator of ice-mediated accessibility of open water) are generally not so, because the spatial and temporal scales of the predictor data are much coarser than the actual environmental conditions experienced by the animals. As well as being proxies for more direct (but unmeasurable) information, the predictor variables in the Southern Ocean are typically highly correlated, because of the strong latitudinal and seasonal gradient that affects oceanic and atmospheric conditions. Because of these factors, it is rarely obvious which particular predictor variable is the most appropriate proxy to use in a given model. Predictive performance offers some guidance, but should not be relied upon exclusively, particularly with small sample size. Therefore, in order to obtain reasonable assessments of the uncertainties in the spatial predictions, we applied the cross-validation procedure across predictor variables as well as individual animals. That is, for a given model, each individual predictor variable was in turn swapped for one or more alternative predictors that were deemed to be plausibly able to act as a similar proxy. The model was re-fitted and the spatial predictions re-calculated each time. The range of predictions obtained (i.e. across all combinations of data folds and predictor variable sets) was used as an indication of the uncertainty in the spatial predictions for that model.

Individual species predictions were combined to quantify overlap. For each grid cell in the study region, the top four habitat importance values were averaged. Uncertainty in the overlap was estimated by resampling from the range of predictions for each species (i.e. combinations of data folds and predictor variable sets) and repeating the overlap calculations.

Predictor variables

The environmental predictor variables available for the habitat selectivity modelling are listed in Table S3. All predictions were made on a 0.1-degree grid from 30–150 °E, 71–55 °S. Interpolation of predictor data to the prediction grid or track points was by bilinear interpolation, except where noted. All distances were calculated as great-circle distances assuming a spherical earth of radius 6378.1 km. Data and further details are available from http://webdav.data.aad.gov.au/data/environmental/derived/.

The transport cost variables were intended to provide an estimate of the accessibility of an area to animals from a certain colony, accounting for the prevailing winds or ocean currents. These were calculated following a similar method to that used by Raymond et al. (Raymond et al. 2010). Briefly, the preferred speed of the animal through the air (for light-mantled albatrosses) or water (penguins and seals) was estimated from published studies. For light-mantled albatrosses, this was taken as the best glide speed (i.e. the air speed at which the forward speed is highest relative to the descent rate; Pennycuick 2008), which was estimated at 12.3 m/s. For penguins and seals, swimming speeds were taken from published estimates of "preferred" or "normal" swimming speed. Given this speed and the prevailing wind or current field, the time required to travel between two given locations can be calculated by simple vector arithmetic. A spatial grid was overlaid on the study region, and the time required to transition between each cell and its 8 adjoining neighbours calculated. Grid cells were randomly placed with average spacing of 0.2° in longitude and 0.1° in latitude. The minimum time required to travel from a given colony location to a particular destination (and return) was calculated using these transition costs and the Bellman-Ford shortest paths algorithm. The prevailing near-surface wind was estimated using a long-term average (2000–2010) of data from NCEP/DOE Reanalysis 2 (Kanamitsu et al. 2002). The surface ocean current field was estimated from a circumpolar Antarctic implementation of the Regional Ocean Modelling System (ROMS) data (Galton-Fenzi et al. 2012).

The distance from deployment and transport cost variables require knowledge of the deployment location of an individual animal. For training data (i.e. observed and simulated tracks) this is known. For gridded predictions, known colony locations of the species of interest were used to compute the distance from deployment (i.e. colony) and transport cost for all colonies. The gridded predictor value at a given grid cell was taken

as the minimum distance or cost across all colonies. Similarly, the sea ice monthly, distance to sea ice monthly, and sea ice days since melt variables were matched to the actual times of the animal locations. For gridded predictions, the long-term mean values of these fields were used.

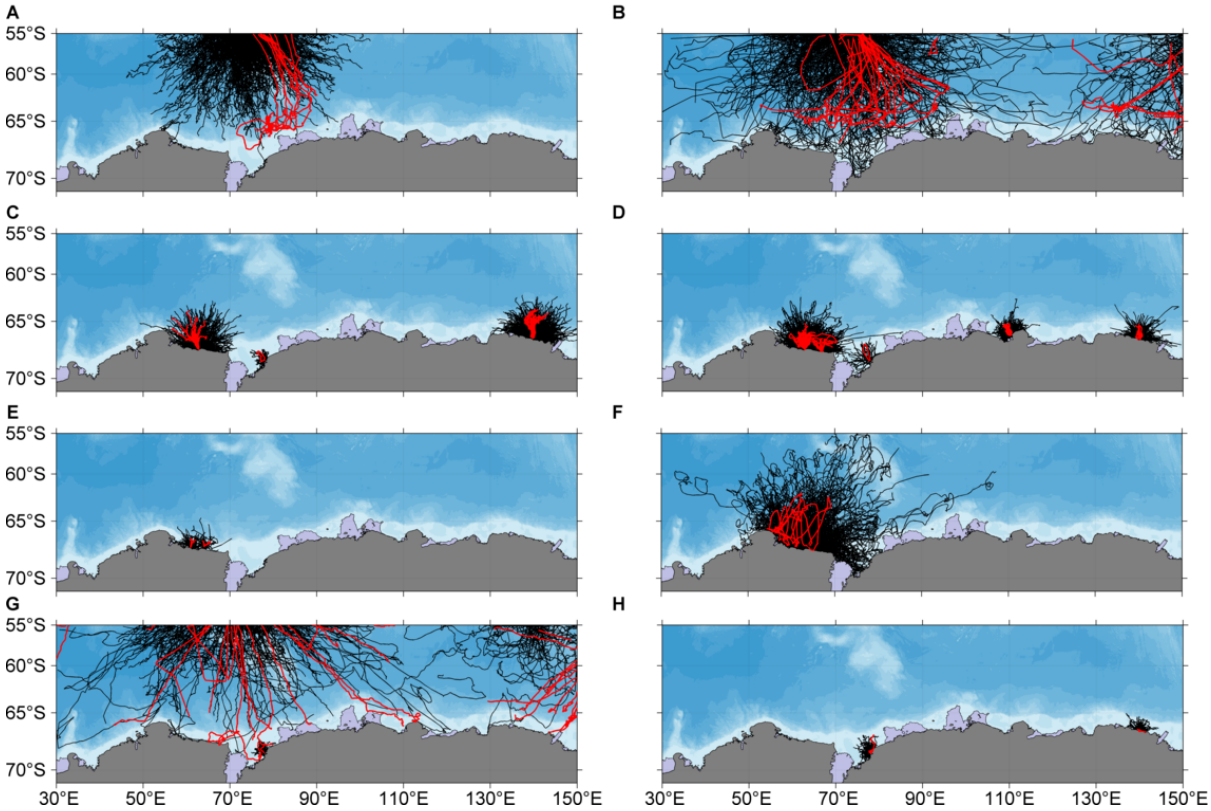

Figure A1. Observed (red) and simulated (black) tracks. (a) male Antarctic fur seals, (b) light-mantled albatrosses, Adélie penguins during the (c) incubation and (d) chick-rearing periods, emperor penguins during the (e) chick-rearing and (f) pre-moult periods, (g) southern elephant seals, and (h) female Weddell seals.

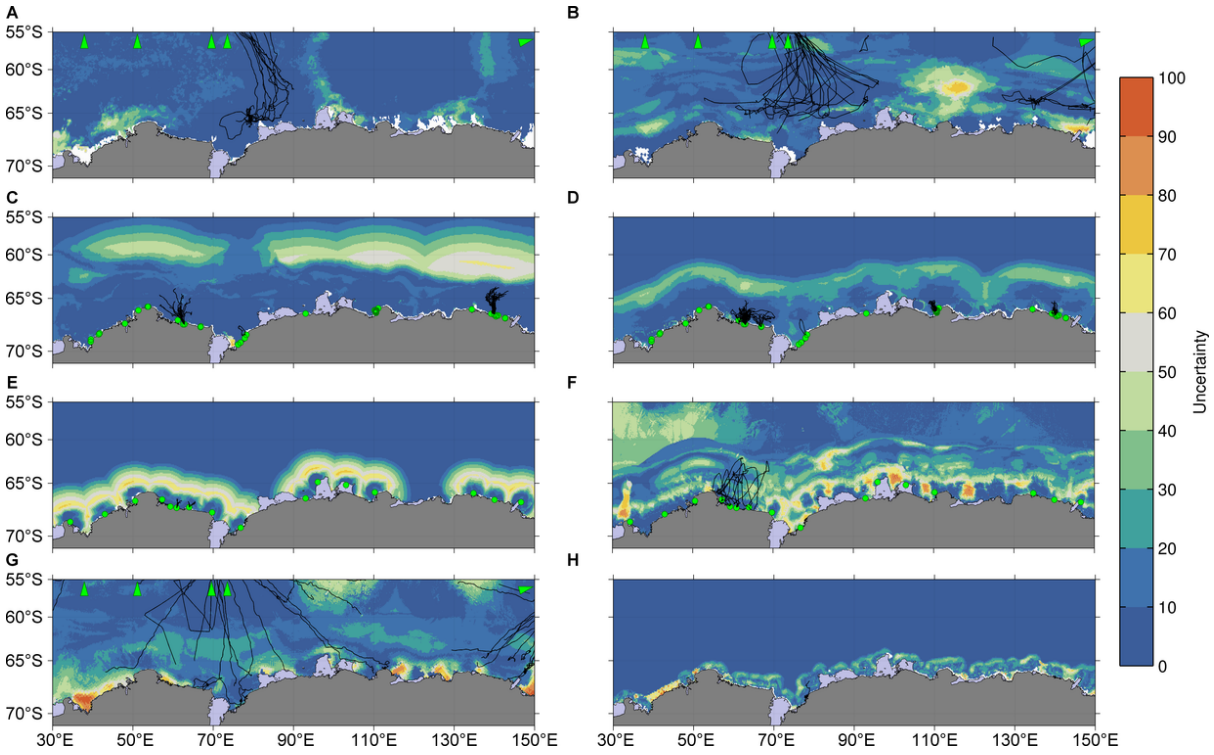

Figure A2. Uncertainty associated with the predicted habitat importance values shown in Figure 1. (a) male Antarctic fur seals, (b) light-mantled albatrosses, Adélie penguins during the (c) incubation and (d) chick-rearing periods, emperor penguins during the (e) chick-rearing and (f) pre-moult periods, (g) southern elephant seals, and (h) female Weddell seals. Uncertainty was calculated as the interquartile range of predicted values, using a cross-validation procedure (see text). Green points show colony locations for the Antarctic-breeding Adélie and emperor penguins; green arrows show the direction of (from west to east) subantarctic Marion and Prince Edward, Crozet, Kerguelen, Heard, and Macquarie islands, which host breeding colonies of Antarctic fur seals, southern elephant seals, and light-mantled albatrosses.

Table A1. Further details for each of the single-species models and data. Model specification is given in the format of a formula passed to the gam function in the R mgcv package. The predictor swaps were used in the estimation of the uncertainty in the predictions for each model (see Supplementary material Appendix 1).

Species	**Data notes and references**	**Data date range**	**Model specification**	**Predictor swaps**	
				Predictor variable or model term	**Alternatives**
Antarctic fur seal (post-breeding sub-adult/adult males)	All data were obtained from trackers deployed on post-breeding sub-adult and adult male individuals. References: (Gales et al. 2004, updated 2010)	1-Jan – 28-Feb. Data came from the 2003/04 austral summer season	p ~ s(log10(chl_summer_climatology), k = 5) + s(transport_cost, k = 5) + s(distance_to_polynya, k = 5)	Distance to polynya	Distance to sea ice monthly Sea ice cover
				log10(summer mean chlorophyll-*a*)	Primary production (February mean)
				Transport cost	Distance from deployment
Adélie penguin (incubation)	The GPS tracking studies of Adélie penguins at Dumont d'Urville were supported logistically by the French Polar Institute (IPEV) and the Terres Australes et Antarctiques Francaises (TAAF), as well as	24-Nov – 29-Dec. Data came from eight individual austral summer seasons from	p ~ s(transport_cost, k = 7) + te(seaice_cover, distance_upper_slope, k = c(3, 3))	te(sea ice cover, distance to upper slope)	te(sea ice monthly,distance to upper slope) te(sea ice cover, log10(bathymetry))
				Transport cost	Distance from deployment

	financially by the WWF and the Zone Atelier	1991/92–2011/12			
Adélie penguin (chick-rearing)	Antarctique et subantarctique. References: (Clarke et al. 1998, Clarke et al. 2006, Cottin et al. 2012, Emmerson et al. 1999, updated 2013, Kerry et al. 1997, Nicol et al. 2008, Wienecke et al. 2000)	22-Dec – 16-Feb. Data came from 13 individual austral summer seasons from 1991/92–2011/12	p ~ s(log10(bathymetry), k = 7) + te(seaice_summer_variability, seaice_cover, k = c(4, 4)) + s(transport_cost, k = 5) + s(distance_upper_slope, k = 7)	Sea ice cover Transport cost	Sea ice monthly Distance to sea ice monthly Distance from deployment
Emperor penguin (chick-rearing)	References: (Wienecke et al. 2004)	5-Dec – 13-Dec. Data came from the 2000/01 austral summer season	p ~ s(transport_cost, k = 5) + s(fast_ice, k = 4)	Transport cost Fast ice cover	Distance from deployment Distance to fast ice
Emperor penguin (pre-moult)	References: (Wienecke et al. 2004)	15-Dec – 26-Feb. Data came from the 2000/01 austral	p ~ s(log10(bathymetry), by = on_shelf, k = 5) + s(distance_to_fast_ice, k = 5)	log10(bathymetry), by=on_shelf Distance to fast ice	Distance to Antarctica Distance to upper slope Distance to sea ice monthly

		summer season			
Light-mantled albatross (chick-rearing)	All data were obtained from trackers deployed on adult individuals during the chick-rearing phase of the breeding season. References: (Lawton et al. 2008, Weimerskirch and Robertson 1994)	14-Nov – 28-Feb. Data came from three individual austral summer seasons from 1992/93–2003/04	p ~ s(surface_zonal_wind_summer, k = 7) + s(seaice_days_since_melt, by = seaice_zone_flag, k = 7) + s(transport_cost, k = 7)	Summer mean zonal wind Sea ice days since melt Transport cost	Distance to upper slope Distance to sea ice monthly Sea ice cover Distance from deployment
Southern elephant seal (post-breeding/post-moult males and females)	Some elephant seal data was sourced from the Integrated Marine Observing System (IMOS), which is supported by the Australian Government through the National Collaborative Research Infrastructure Strategy and the Super Science Initiative. Data were also	11-Nov to 28-Feb. Data came from eight individual austral summer seasons from 2003/04–2010/11	p ~ te(distance_from_deployment, log10(bathymetry), k = c(4,4)) + s(seaice_monthly, k = 5)	te(distance from deployment, log10(bathymetry)) Sea ice monthly	te(distance from deployment, distance to upper slope) te(distance_antarctica,log10(bathymetry)) Distance to sea ice monthly Sea ice cover

	obtained from the MEOP (Marine Mammals Exploring the Oceans Pole to Pole) International Polar Year programme and the SEaOS (Southern Elephant Seals as Oceanographic Samplers) project. The French elephant seal data collected as part of SEaOS and MEOP were provided by the SO-MEMO project (PI C. Guinet). References: (Bestley et al. 2012, Biuw et al. 2007)				
Weddell seal	Some Weddell seal data was sourced from the Integrated Marine Observing System (IMOS), which is supported by the	1-Nov – 28-Feb. Data came from four individual austral	p ~ s(log10(bathymetry), k = 5) + s(distance_to_fast_ice, k = 4) + s(distance_to_polynya, k = 4)	log10(bathymetry) Distance to fast ice Distance to polynya	Distance to Antarctica Distance to sea ice monthly Sea ice cover

	Australian Government through the National Collaborative Research Infrastructure Strategy and the Super Science Initiative. References: (Andrews-Goff et al. 2010, Heerah et al. 2012, Lake et al. 2006)	summer seasons from 1999/2000–2007/08

Table A2. Summary of the important environmental dependencies in the individual species models. “Post-polynyas” refers to locations corresponding to winter polynyas, but in the spring/summer when they are technically no longer polynyas. Note that interpretation of this table is complicated by the indirect nature of many of the predictor variables, and also by the strong, common latitudinal and seasonal structuring of many environmental processes in the East Antarctic region, which makes it difficult to ascribe particular importance to individual environmental variables.

Species	**Features**
Antarctic fur seal (post-breeding sub-adult/adult males)	Marginal ice zone. Areas of elevated productivity in the marginal ice zone and to the east of the Kerguelen Plateau. Moderately constrained to near-continental areas generally south of the breeding colony.
Adélie penguin (incubation)	Offshore polynyas or areas or reduced ice cover. Marginal ice zone, avoiding areas of heavy sea ice cover. Shelf slope. Constrained to within approx. 400 km of breeding colony.
Adélie penguin (chick-rearing)	General use of shelf post-polynya areas, but not exclusively. Areas of moderate variability in summer sea ice cover, avoid areas of heavy cover. Shelf slope. Constrained to within approx. 150–200 km of breeding colony.
Emperor penguin (chick-rearing)	Shelf post-polynya locations. Near fast ice, but avoiding areas typically fast-ice covered. Constrained to within approx. 100 km of breeding colony.
Emperor penguin (pre-moult)	Within approx. 100km of fast ice. Open-ocean areas north of the shelf slope. Deep areas of the shelf.
Light-mantled albatross (chick-rearing)	Marginal ice zone, particularly recently ice-covered areas. Open-ocean areas north of the shelf slope, in the easterly wind band south of about 66 °S. Moderately constrained to oceanic areas generally south of the breeding colony.
Southern elephant seal (post-breeding/post-moult males and females)	Shelf post-polynyas. Marginal ice zone, avoiding heavy ice cover. Shallow parts of shelf. Large dispersal distances.

Weddell seal	Within approx. 100km of winter polynya locations and approx. 100km of fast ice. Shallow parts of shelf. Highly territorial.

Table A3. Environmental predictor variables.

Variable	Description and source data	Source data	Description and processing steps
Bathymetry	Measured and estimated sea floor topography from satellite altimetry and ship depth soundings	Smith and Sandwell (Smith and Sandwell 1997) V15.1 at 1-minute resolution	
Bathymetry slope	Slope of sea floor	As above	Slope calculated on 0.1-degree gridded depth data (above)
Mean summer chlorophyll-*a*	Near-surface chlorophyll-*a* summer climatology	MODIS Aqua at 9km resolution (Feldman and McClain 2010)	Climatology spans the 2002/03 to 2012/13 austral summer seasons
Distance to Antarctica	Distance to the nearest part of the Antarctic continent	World map shapefile courtesy of ESRI	
Distance from deployment	Distance from deployment location		Distance for each individual calculated from the deployment location where known, otherwise from the first recorded position for that individual
Distance to maximum ice extent	Distance to the nearest point on the line of mean maximum winter sea ice extent	SMMR-SSM/I passive microwave estimates of sea ice concentration at 25km resolution (Cavalieri et al. 1996, updated yearly)	Mean maximum winter sea ice extent over the 1979/80 to 2008/09 austral summer seasons was derived from daily estimates of sea ice concentration as described in this metadata record
Distance to canyon	Distance to the axis of the nearest canyon	Seafloor geomorphic feature dataset (Post, unpublished data, expanded from O'Brien et al. 2009)	
Distance to fast ice	Distance to the nearest location where fast ice is typically present	20-day composite records of landfast sea-ice at 2km resolution, derived from MODIS imagery (Fraser et al.	Fast ice considered to be "typically present" at pixels that were associated with fast ice presence for more than half of the year on

		2012)	average
Distance to polynya	Distance to nearest winter polynya location	AMSR-E satellite estimates of daily sea ice concentration at 6.25km resolution (Spreen et al. 2008)	The sea ice coverage layer (below) was used. Pixels which were, on average, covered by sea ice for less than 35% of the year were identified as polynya pixels. The threshold of 35% was chosen to give a good empirical match to the polynya locations identified by Arrigo & van Dijken (Arrigo and van Dijken 2003), although the results were not particularly sensitive to the choice of threshold
Distance to sea ice monthly	Distance to nearest sea ice	SMMR-SSM/I passive microwave estimates of sea ice concentration at 25km resolution (Cavalieri et al. 1996, updated yearly)	Mean monthly sea ice data, matched to the time of each observation, were used. A threshold of 15% ice concentration was used as the cutoff between open and ice-covered water
Distance to upper slope	Distance to the "upper slope" geomorphic feature	Seafloor geomorphic feature dataset (Post, unpublished data, expanded from O'Brien et al. 2009)	
Fast ice coverage	The average proportion of the year for which landfast sea ice is present	20-day composite records of landfast sea-ice at 2km resolution, derived from MODIS imagery (Fraser et al. 2012)	The average proportion of the year for which each pixel was covered by landfast sea ice was calculated as an average across 2001–2008
Floor temperature	Seafloor water temperature	Floor temperatures derived from World Ocean Atlas 2005 data at 1-degree resolution (Clarke et al. 2009)	Isolated missing pixels (i.e. single pixels of missing data with no surrounding missing pixels) were filled using bilinear interpolation. Subsequent interpolation to grid or track points by nearest neighbour

			interpolation
Geomorphology	Geomorphic feature classes	Seafloor geomorphic feature dataset (Post, unpublished data, expanded from O'Brien et al. 2009). Mapping based on GEBCO contours, ETOPO2, seismic lines	
Summer mean mixed layer depth	Summer mixed layer depth climatology	ARGO float data at 2-degree resolution (de Boyer Montegut et al. 2004)	Interpolation to grid or track points by nearest neighbour interpolation
Sea ice coverage	The average proportion of the year for which sea ice is present	AMSR-E satellite estimates of daily sea ice concentration at 6.25km resolution (Spreen et al. 2008)	Concentration data from 1-Jan-2003 to 31-Dec-2010 was used. The fraction of time each pixel was covered by sea ice of at least 85% concentration was calculated
Sea ice days since melt	Number of days since the location was last covered by sea ice	SMMR-SSM/I passive microwave estimates of sea ice concentration at 25km resolution (Cavalieri et al. 1996, updated yearly)	A threshold of 15% ice concentration was used as the cutoff between open and ice-covered water
Sea ice monthly	Monthly mean sea ice cover, matched to the month of the observation	SMMR-SSM/I passive microwave estimates of sea ice concentration at 25km resolution (Cavalieri et al. 1996, updated yearly)	
Sea ice summer variability	Variability of sea ice cover during summer months	AMSR-E satellite estimates of daily sea ice concentration at 6.25km resolution (Spreen et al. 2008)	Daily estimates of sea ice concentration across December, January, and February of a given austral summer season were collated. For each pixel, the standard deviation of these values was calculated. The values given here are averaged over the 2002/03 to 2009/10 austral summer seasons

Sea surface height	Mean dynamic topography (sea surface height relative to geoid)	CNES-CLS09 Mean Dynamic Topography v1.1 at 0.25-degree resolution (Rio et al. 2011)	
Sea surface height spatial gradient	The spatial gradient (in mm/km) of the mean dynamic topography	CNES-CLS09 Mean Dynamic Topography v1.1 at 0.25-degree resolution (Rio et al. 2011)	Gradient calculated on the native 0.25-degree grid
Mean summer sea surface temperature	Sea surface temperature summer climatology	MODIS Aqua at 9km resolution (Feldman and McClain 2010)	Climatology spans the 2002/03 to 2009/10 austral summer seasons
Sea surface temperature spatial gradient	Spatial gradient of mean summer SST (above)	MODIS Aqua at 9km resolution (Feldman and McClain 2010)	Spatial gradient of the SST (degrees C per km) calculated on the original 9km resolution data
Summer surface wind speed	Mean wind speed at 10m height	NCEP/DOE Reanalysis 2 at 2.5-degree resolution (Kanamitsu et al. 2002)	Calculated from 2000–2010 monthly mean values
Summer zonal surface wind speed	Mean zonal wind speed at 10m height	NCEP/DOE Reanalysis 2 at 2.5-degree resolution (Kanamitsu et al. 2002)	Calculated from 2000–2010 monthly mean values
Summer meridional surface wind speed	Mean meridional wind speed at 10m height	NCEP/DOE Reanalysis 2 at 2.5-degree resolution (Kanamitsu et al. 2002)	Calculated from 2000–2010 monthly mean values
Transport cost	See description above		
Primary productivity	Mean February net primary productivity estimated from a vertically generalized production model (VGPM)	VGPM data at 1/6-degree resolution (Behrenfeld and Falkowski 1997) obtained from the Oregon State ocean productivity web site	Mean monthly February productivity, 2003–2010

References

Aarts, G. et al. 2008. Estimating space-use and habitat preference from wildlife telemetry data. – Ecography 31: 140–160.

Andrews-Goff, V. et al. 2010. Factors influencing the winter haulout behaviour of Weddell seals: consequences for satellite telemetry. – Endangered Species Research 10: 83–92.

Arrigo, K. R. and van Dijken, G. L. 2003. Phytoplankton dynamics within 37 Antarctic coastal polynya systems. – J. Geophys. Res. 108: 3271.

Behrenfeld, M. J. and Falkowski, P. G. 1997. Photosynthetic rates derived from satellite-based chlorophyll concentration. – Limnol. Oceanogr. 42: 1–20.

Bestley, S. et al. 2012. Integrative modelling of animal movement: incorporating in situ habitat and behavioural information for a migratory marine predator. – Proceedings of the Royal Society of London. Series B: Biological Sciences 280: 20122262.

Biuw, M. et al. 2007. Variations in behavior and condition of a Southern Ocean top predator in relation to in situ oceanographic conditions. – Proc. Natl. Acad. Sci. U. S. A. 104: 13705–13710.

Cavalieri, D. et al. 1996, updated yearly. Sea ice concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS passive microwave data. Boulder, Colorado USA: NASA DAAC at the National Snow and Ice Data Center.

Clarke, A. et al. 2009. Spatial variation in seabed temperatures in the Southern Ocean: implications for benthic ecology and biogeography. – J. Geophys. Res. 114: G03003.

Clarke, J. et al. 1998. Sex differences in Adélie penguin foraging strategies. – Polar Biol. 20: 248–258.

Clarke, J. et al. 2006. Environmental conditions and life history constraints determine foraging range in breeding Adélie penguins. – Mar. Ecol. Prog. Ser. 310: 247–261.

Cottin, M. et al. 2012. Foraging strategies of male Adélie penguins during their first incubation trip in relation to environmental conditions. – Marine Biology 159: 1843–1852.

De'ath, G. 2007. Boosted trees for ecological modeling and prediction. – Ecology 88: 243-251.

de Boyer Montegut, C. et al. 2004. Mixed layer depth over the global ocean: an examination of profile data and a profile-based climatology. – J. Geophys. Res. 109: C12003.
Elith, J. et al. 2008. A working guide to boosted regression trees. – J. Anim. Ecol. 77: 802–813.
Emmerson, L. et al. 1999, updated 2013. Satellite Tracking of Adelie Penguins Around Mawson Station. Antarctica Australian Antarctic Data Centre, Catalogue of Australian Antarctic and Subantarctic Metadata. http://data.aad.gov.au/aadc/metadata/metadata_redirect.cfm?md=/AMD/AU/Tracking_BI.
Feldman, G. C. and McClain, C. R. 2010. Ocean Color Web, MODIS Aqua Reprocessing, NASA Goddard Space Flight Center. Eds. Kuring, N., Bailey, S.W. http://oceancolor.gsfc.nasa.gov/.
Fraser, A. D. et al. 2012. East Antarctic landfast sea ice distribution and variability, 2000–08. – Journal of Climate 25: 1137–1156.
Gales, N. et al. 2004, updated 2010. Animal Tracking at Heard Island 2003/2004 — ARGOS data. Australian Antarctic Data Centre, Catalogue of Australian Antarctic and Subantarctic Metadata. http://data.aad.gov.au/aadc/metadata/metadata_redirect.cfm?md=/AMD/AU/HI_animaltracks_ARGOS.
Galton-Fenzi, B. K. et al. 2012. Modeling the basal melting and marine ice accretion of the Amery Ice Shelf. – Journal of Geophysical Research: Oceans 117: C09031.
Heerah, K. et al. 2012. Ecology of Weddell seals during winter: Influence of environmental parameters on their foraging behaviour. – Deep Sea Research Part II: Topical Studies in Oceanography 88–89: 23–33.
Kanamitsu, M. et al. 2002. NCEP-DEO AMIP-II Reanalysis (R-2). – Bulletin of the American Meteorological Society 83: 1631–1643.
Kerry, K. R. et al. 1997. The foraging range of Adélie pengins — implications for CEMP and interactions with the krill fishery. – CCAMLR Science 4: 75–87.
Lake, S. et al. 2006. Movements of adult female Weddell seals during the winter months. – Polar Biol. 29: 270–279.

Lawton, K. et al. 2008. Preferred foraging areas of Heard Island albatrosses during chick raising and implications for the management of incidental mortality in fisheries. – Aquatic Conservation: Marine and Freshwater Ecosystems 18: 309–320.
Nicol, S. et al. 2008. Krill (*Euphausia superba*) abundance and Adélie penguin (*Pygoscelis adeliae*) breeding performance in the waters off the Béchervaise Island colony, East Antarctica in 2 years with contrasting ecological conditions. – Deep Sea Research Part II: Topical Studies in Oceanography 55: 540–557.
O'Brien, P. E. et al. 2009. Antarctic-wide geomorphology as an aid to habitat mapping and locating vulnerable marine ecosystems. – In: Commission for the Conservation of Antarctic Marine Living Resources Vulnerable Marine Ecosystems Workshop, Paper WS-VME-09/10. CCAMLR.
Pennycuick, C. J. 2008. Modelling the Flying Bird. – Elsevier.
Phillips, S. J. et al. 2009. Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. – Ecol. Appl. 19: 181–197.
Raymond, B. et al. 2010. Shearwater foraging in the Southern Ocean: the roles of prey availability and winds. – PLoS ONE 5: e10960.
Rio, M. H. et al. 2011. New CNES-CLS09 global mean dynamic topography computed from the combination of GRACE data, altimetry, and in situ measurements. – Journal of Geophysical Research: Oceans 116: C07018.
Smith, W. H. F. and Sandwell, D. T. 1997. Global seafloor topography from satellite altimetry and ship depth soundings. – Science 277: 1957–1962.
Spreen, G. et al. 2008. Sea ice remote sensing using AMSR-E 89 GHz channels. – J. Geophys. Res. 113: C02S03.
Wakefield, E. D. et al. 2010. Habitat preference, accessibility, and competition limit the global distribution of breeding Black-browed Albatrosses. – Ecol. Monogr. 81: 141–167.
Weimerskirch, H. and Robertson, G. 1994. Satellite tracking of light-mantled sooty albatrosses. – Polar Biol. 14: 123–126.
Wienecke, B. et al. 2004. Pre-moult foraging trips and moult locations of Emperor penguins at the Mawson Coast. – Polar Biol. 27: 83–91.
Wienecke, B. C. et al. 2000. Adélie penguin foraging behaviour and krill abundance along the Wilkes and Adélie land coasts, Antarctica. – Deep Sea Research Part II: Topical Studies in Oceanography 47: 2573–2587.

Žydelis, R. et al. 2011. Dynamic habitat models: using telemetry data to project fisheries bycatch. – Proceedings of the Royal Society of London. Series B: Biological Sciences 278: 3191–3200.