

Predicting seabirds at sea in the Southern Indian Ocean

Ben Raymond*, Eric J. Woehler

Australian Antarctic Division, Channel Highway, Kingston, Tasmania 7050, Australia

ABSTRACT: Predictive models were constructed based on at-sea observations of 6 species of seabirds within the Prydz Bay region of East Antarctica. A spatial logistic model was used, and predictions were based on physical environmental data. The hypothesis that these models could predict observations of the same species in the disjunct Vincennes Bay region of East Antarctica was tested. The best results were obtained for the 2 endemic Antarctic breeding species examined (Antarctic and snow petrels). The accuracies of the predictions for Wilson's storm petrels and Cape petrels (which breed both in the Antarctic and on subantarctic islands) were low due to their broad distributions at sea. The predictions for 2 species that breed in subantarctic and temperate localities (white-chinned petrels and prion spp.) were poorest, reflecting the limited seasonal use of Antarctic waters for foraging by these species. The approach is applicable for estimating the probability of observing seabird species in unsurveyed areas of the Southern Ocean, but is likely to be limited to those species breeding regionally, and not for species foraging in the area and breeding elsewhere.

KEY WORDS: Seabirds at sea · Southern Ocean · Spatial model · Predictive · Environmental determinants

Resale or republication not permitted without written consent of the publisher

INTRODUCTION

Seabirds are major top-order predators in the Southern Ocean ecosystem, with annual estimates of prey biomass consumption exceeding 6.6×10^8 t (Croxdall 1984). This prey consumption is patchily distributed throughout the Southern Ocean. The factors that determine the distribution of seabirds at sea, and hence their prey consumption, have been investigated for the last 3 decades. These studies have been more successful in relating the distributions and abundances of seabirds at sea to their physical environment, such as sea surface temperatures, water mass characteristics, sea ice extent, and wind regime, rather than the distribution of their prey species (see Woehler et al. 2003 for a brief review). Of these, the presence/absence of sea ice and its extent and concentration are known to have strong influences on seabird distributions in the Southern Ocean (e.g. Ainley et al. 1993, 1994). These Southern Ocean studies have paralleled similar investigations elsewhere (for example Croll 1990, Leopold 1993, Mehlum et al. 1996).

While correlations between seabirds and aspects of their environments have been demonstrated, there are few published studies that build on these correlations to produce predictive models. Predictive methods have been popular in a range of ecological disciplines (e.g. Guisan & Zimmermann 2000, Austin 2002, De'ath 2002). The primary focus of such methods is the construction of statistical models, which are capable of predicting species distributions based on environmental data. These are in contrast to explanatory methods such as canonical correlation analysis (Gittins 1985) and canonical correspondence analysis (ter Braak & Prentice 1988), which attempt to identify the environmental variables that best explain the observed patterns in species' distributions. The predictive approach has been used for predicting changes in species' distributions in response to climate change (e.g. Scott et al. 1997, Sykes 1997), conservation planning (e.g. Franklin 1995, Cawsey et al. 2002), and improving faunistic and floristic atlases (e.g. Högmänder & Møller 1995, Leathwick 2001). Although the emphasis is on predic-

*Email: ben.raymond@aad.gov.au

tion rather than description, the relationships between species and their environment can be inferred from an understanding of the model's internal species-environment representations.

The predictive approach is particularly appealing for Antarctic seabird-at-sea studies. Species diversity amongst the breeding Antarctic seabirds is relatively low (between 9 and 25 breeding species at regional scales; Woehler et al. 2003), and observational studies have revealed that seabird assemblages at sea are remarkably consistent around the Antarctic continent (Ribic & Ainley 1988, 1989, Ainley et al. 1993, Woehler et al. 2003). Ice-associated assemblages have been identified in the Ross Sea (Ainley et al. 1984), South Atlantic, Amundsen and Bellingshausen Seas (Ainley et al. 1998) and Indian (Woehler et al. 2003) sectors of the Southern Ocean, with only minor differences in species composition due to differences in regional breeding species (Woehler et al. 2003). Predictive models developed in one region might, therefore, be applicable to other regions of the Southern Ocean.

As a consequence of the difficulties involved in undertaking synoptic seabird surveys in the Southern Ocean, the collection of at-sea data has been limited to 3 foci around the Antarctic: the Ross Sea (Ainley et al. 1984), the South Atlantic sector (Ainley et al. 1994), and Prydz Bay (Woehler 1995, 1997, Woehler et al. 2003). The remaining majority of the Antarctic coastline has received little attention. Predictive models applied to remotely sensed measurements of the physical environment could be used to provide an interim indication of likely seabird communities in these areas until regional surveys are undertaken. Predictive models could also be used to explore the implications of climate change, such as the reduction in sea ice extent documented for the Antarctic Peninsula region (Smith et al. 1999).

This study develops models for the prediction of at-sea distributions of 6 species of Southern Ocean seabirds in the Vincennes Bay region of East Antarctica, based on observations of these species within the Prydz Bay region, approximately 1000 km to the west, for the period 1980/81 to 2001/02.

MATERIALS AND METHODS

Seabird and environmental observations. Details of the observation methodologies have been described previously (Woehler 1995, 1997). In these previous studies, the Prydz Bay region was defined as that area of the Southern Ocean between 60 and 90°E and south of 60°S to the Antarctic continent. For this study, the Prydz Bay region (hereafter 'Prydz Bay') has been extended 10° farther north to include Heard Island

(53°05'S, 73°30'E), a breeding site for several seabird species investigated here. The Vincennes Bay region (hereafter 'Vincennes Bay') is defined for this study as being between 100 and 120°E, and south of 50°S to the Antarctic continent. Seabird-at-sea data from 1980/81 to 2001/02 were used in this study, except for observations from 1982/83, when a different observational methodology was used (Woehler 1997). Data for prions *Pachyptila* spp. have been pooled in analyses, as these are difficult to separate at sea (Woehler 1997).

Physical environmental data collected contemporaneously with seabird observations included sea surface temperature (°C), sea state (Beaufort), cloud cover (oktas), precipitation, wind force (knots), and air pressure (hPa). Sea ice concentrations (% ice cover) were estimated from passive microwave satellite images (Cavalieri et al. 1999, 2002). The edge of the ice pack (defined as the northernmost latitude at which the sea ice cover dropped below 15%) was computed. The standard deviation of sea surface temperature within a spatial bin was calculated to detect frontal regions. Digital coastline data (Soluri & Woodson 1990), bathymetric data (National Geophysical Data Center 1988), the continental shelf break, and the southern boundary of the Antarctic Circumpolar Current (ACC; Orsi et al. 1995) were included in analyses. The shelf break is usually considered to be ca. 600 m depth. However, we did not have the complete 600 m isobath across the study area and so the 1000 m isobath (GEBCO Digital Atlas, British Oceanographic Data Centre 1994) was used. The average distance between these 2 isobaths was 4.3 km in the longitude interval 60 to 90°E: a small difference when compared to the spatial bin sizes used (see below). The pack-ice edge, coastline, shelf, and ACC data were included in binary form, where a '1' indicated that the feature intersected the spatial bin in question.

The 2 most commonly observed species in both Prydz Bay and Vincennes Bay in each of the 3 breeding categories outlined below were chosen for analyses. The selected species were snow petrels *Pagodroma nivea* and Antarctic petrels *Thalassoica antarctica*, which are Antarctic endemic breeding species with breeding populations in Prydz Bay and Vincennes Bay; Cape petrels *Daption capense* and Wilson's storm petrels *Oceanites oceanicus*, which breed along the Antarctic coast and at subantarctic localities; and white-chinned petrels *Procellaria aequinoctialis* and prions *Pachyptila* spp., which breed on islands in temperate latitudes and forage in Antarctic waters during the austral summer.

Statistical analyses. A discrete-space approach was used, in which the geographical sample space was divided into a grid of rectangular spatial bins. We used a bin size of 2° longitude by 2° latitude (approximately 95 × 220 km or 21 000 km² per bin at latitude 65°S).

An assessment of the sensitivity of the results to the choice of spatial bin size was undertaken.

The at-sea observation data were transformed to presence/absence format. The encounter rate for a particular species in a particular bin was calculated as the fraction of surveys of that bin in which the species was observed. The species' encounter rate was averaged across all bins with a non-zero encounter rate. The species was then marked as present in any bin for which the encounter rate was greater than 10% of this average. This procedure was used in order to avoid including bins with only infrequent observations of a particular species in that species' range. The presence or absence of observations of each species as a function of physical environment was modelled using logistic regression (McCullagh & Nelder 1989).

At-sea observations of seabirds display spatial correlation: observations taken at locations that are spatially proximate will be more closely related than observations taken from widely separated locations (Schneider 1990). The standard logistic model ignores this spatial correlation and assumes that the probabilities of observing a species in neighbouring spatial bins are independent. The predictive accuracy of a model can be improved in some applications if a mechanism for spatial dependence is included in the model (Cressie 1993, Augustin et al. 1996, Carroll & Pearson 2000). Accordingly, we also investigated the spatial logistic model (also referred to as the autologistic model; Besag 1974). This extends the logistic model by modelling presence as a function of environment and of presence in surrounding spatial bins. The estimation of the spatial logistic model parameters is problematic because the exact maximum likelihood solution is computationally intractable except where the number of bins is small. Approximate methods, such as maximum pseudo-likelihood (Besag 1975, Augustin et al. 1998), are commonly used. We used the Markov Chain Monte Carlo approximation to the maximum likelihood solution (Geyer & Thompson 1992, Huffer & Wu 1998) as implemented by LeSage (2000). For further details of the spatial logistic model, see Augustin et al. (1996, 1998) for an example.

The seabird observations were collected across more than 2 decades, and at different times during each year; hence, there is temporal and spatial variation within the data. Our interest was in the long-term relationships between observations and environment and so data from all years were pooled. The time of year was measured on a seasonal basis (beginning on July 1 and ending on June 30 the following year). We assumed that the relationships between seabird observation and environment might not be constant within this seasonal time frame, as the birds' behaviours will be driven by differing processes throughout the sea-

son. For each species we therefore fitted a sequence of models, each spanning a 30 d time period, and with consecutive models overlapping by 15 d. Thus, the first model in the sequence covered the date range July 1 to July 30 (data from all years combined), the second covered July 15 to August 14, and so on. The 30 d time period represents a compromise between the need for sufficient data for parameter estimation and sufficient temporal resolution to allow the within-season variation in bird observations to be captured. All environmental observations associated with seabird observations within the bin were averaged. Bins which were surveyed on less than 5 separate occasions were not included in the analyses.

Preliminary analyses (not detailed here) using forward selection of environmental variables by cross-validation prediction error revealed that wind force, cloud cover, precipitation, air pressure, the standard deviation of sea surface temperature, the edge of the ice pack, and the southern boundary of the ACC were weak predictors of observations. Thus, the final environmental variables used for analyses in this study were sea surface temperature, sea ice concentration, sea state, coastline, water depth, and shelf break. Covariance amongst predictor variables can make model interpretation difficult, and so the following variable selection procedure was used to reduce the number of predictor variables in each model. For each time period, the observation data from Prydz Bay were divided into 3 subsets of approximately equal size. This division was done by voyage; that is, the available voyages were randomly divided into 3 subsets. Variables were added to the model, using data from the first subset of voyages to estimate the parameters. The variable that gave the best improvement in predictive accuracy (calculated on data from the second voyage subset) was retained in the model. The improvement in predictive accuracy was noted and used to assess the importance of the variable as a predictor. Variables were added until no improvement in predictive accuracy was obtained (or all variables had been added). The predictive accuracy of the final model was estimated using data from the third subset of voyages. This procedure is known as cross-validation, and is a widely used method of assessing prediction error with limited data (Stone 1974, Hastie et al. 2001). Each model was also applied to data from Vincennes Bay to test the generality of the model when applied to data from a different geographic region. The model building, cross-validation, and testing procedures were repeated 10 times, using different randomly selected voyage subsets each time.

Predictive errors are presented as mean squared prediction error (MSE). The null MSE is given for comparison in each case. This is the error under a 'null'

model, which assumes a constant probability of observation across all spatial bins. The null model MSE reflects the baseline probability of observation of a given species. A model which is no more accurate than the null MSE is therefore of little value for providing predictive estimates. However, as discussed below, the internal representations formed by such models can still be of interest. The model MSE for each species was compared to the null MSE using a Wilcoxon paired-sample test at a significance level of $p = 0.05$.

RESULTS

The density of surveys between 1980/81 and 2001/02 is shown in Fig. 1. We have provided maps of observations of the 6 species included in these analyses online (available at <http://purl.oclc.org/net/wov/>). Fig. 2 shows the prediction error of the logistic models for each of the 6 species listed in Table 1; errors are given for both Prydz Bay and Vincennes Bay. Fig. 3 shows the relative importance of each environmental variable as a predictor of each of these 6 species. The correlations between environmental variables are given in Table 2. We present the results of the spatial logistic model for each species in turn.

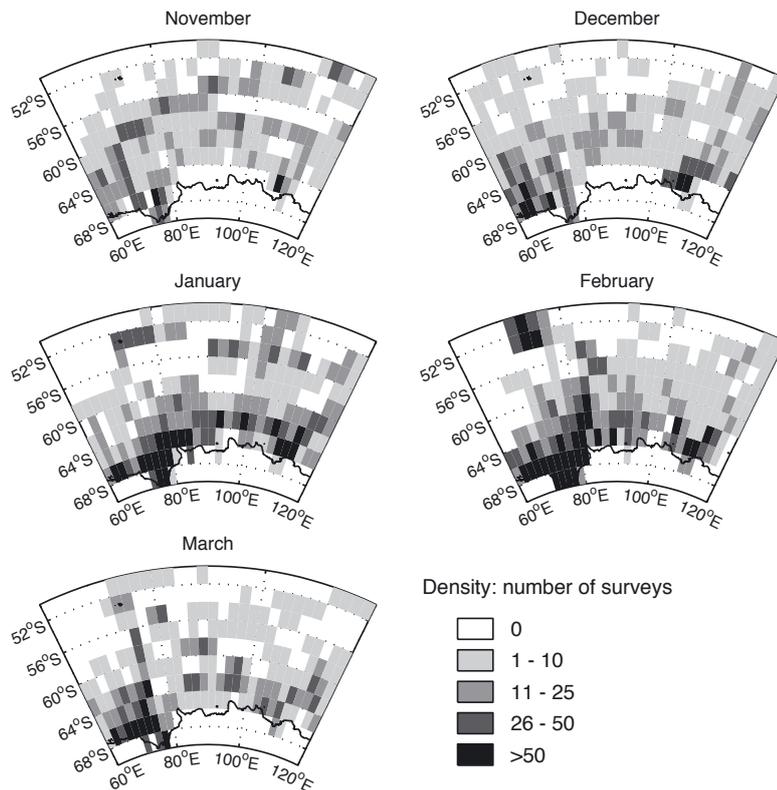


Fig. 1. Density of at-sea surveys of seabirds in the Prydz Bay and Vincennes Bay regions, East Antarctica, between 1980/81 and 2001/02

Snow petrel

The distributions of at-sea sightings of snow petrels in Prydz Bay indicate that snow petrels occupy the southern areas of Prydz Bay, extending as far north as 56°S during October and November, and contracting southward to the coastal regions of Prydz Bay (south of 64°S) for the period from January to March. There were insufficient data after March, and so the northward dispersion of snow petrels from their breeding colonies around Prydz Bay after this time was not recorded. The distribution of snow petrel sightings in Vincennes Bay was similar to that in Prydz Bay, with a southward contraction to the coastal areas throughout the season. The spatial logistic models of snow petrel sightings in Prydz Bay were significantly better than the null model in each time period except late October (Fig. 2). The same models applied to the Vincennes Bay data also produced better performance than the null in each time period except late October, indicating good generalisation. Sea ice concentration (positive model coefficient, indicating that snow petrels were associated with sea ice) and sea state (negative coefficient; indicating an association with small values of sea state) were the most important predictors from November until early January (Fig. 3). The shelf break

was an important feature around January (positive coefficient, indicating an association with the shelf break). Sea surface temperature (negative coefficient) was the most important predictor during February and March.

Antarctic petrel

The distributions of sightings of Antarctic petrels in Prydz Bay and Vincennes Bay were similar to those of snow petrels. The sightings of Antarctic petrels during January and February extended slightly farther from the Prydz Bay coast than did those of snow petrels; additionally, Antarctic petrels were not seen in southern most Prydz Bay (approximately 70°S, 75°E) from January to March. The spatial logistic models of Antarctic petrel sightings in Prydz Bay were more accurate than the null, except during the early (November) and late (March) parts of the season (Fig. 2). When applied to the Vincennes Bay data, the model errors were also significantly lower than the null, except during the early and later parts of the

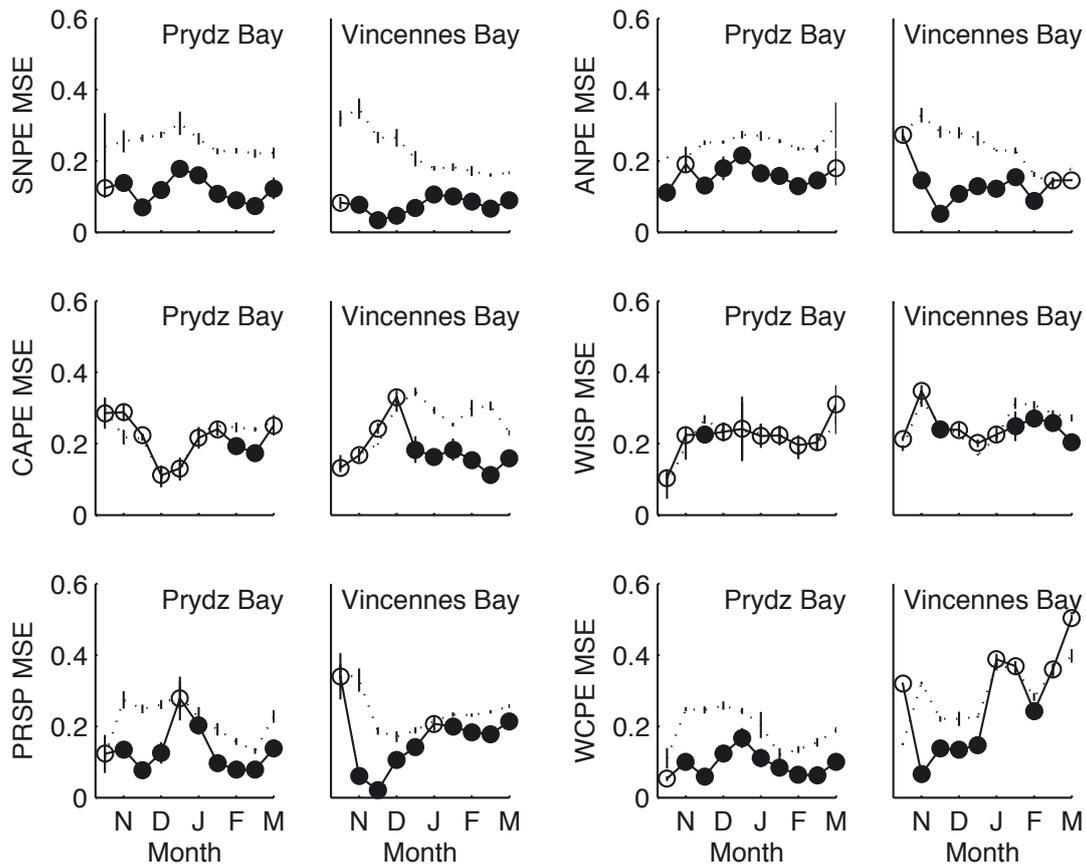


Fig. 2. Mean square prediction errors (MSE) of the logistic (solid line) and null (dotted line) models applied to at-sea observations of 6 species of seabirds. A filled circle indicates that the logistic MSE is significantly less than the corresponding null MSE ($p < 0.05$, Wilcoxon paired-sample test). Error bars show standard deviation. Data are taken from the Prydz Bay and Vincennes Bay regions. SNPE = snow petrels, ANPE = Antarctic petrels, CAPE = Cape petrels, WISP = Wilson's storm petrels, PRSP = prion species, WCPE = white-chinned petrels

season. The pattern of variable importance was similar to that of snow petrels: sea ice concentration (positive coefficient) was the dominant influence during November and December; thereafter the shelf break (positive coefficient) was the dominant influence (Fig. 3). Sea

surface temperature (negative coefficient) also contributed throughout the season. These results are consistent with previous observations that Antarctic petrels are observed near the shelf early in the summer and farther offshore later (Ainley et al. 1984, 1993).

Table 1. Number of surveys during which sightings of 6 species of seabird were made in the Prydz Bay and Vincennes Bay regions of East Antarctica, from 1980/81 to 2001/02

Species/taxa	No. of sightings in Prydz Bay (no. of surveys: 12 049)	No. of sightings in Vincennes Bay (no. of surveys: 5051)	Total no. of sightings (total surveys: 17 100)
Snow petrel	5344	1513	6857
Antarctic petrel	5021	1100	6121
Cape petrel	3816	1203	5019
Wilson's storm petrel	3200	1099	4299
Prion spp.	7291	1863	9154
White-chinned petrel	4032	1054	5086

Cape petrel

Cape petrel sightings in Prydz Bay were concentrated in the northern part of Prydz Bay (N of 60°S) during October, extending southwards throughout the entire Prydz Bay region during November and December, and then contracting to the coastal regions adjacent to breeding sites around Prydz Bay and Heard Island during January and February. Observations in March indicate that Cape petrels were beginning to leave

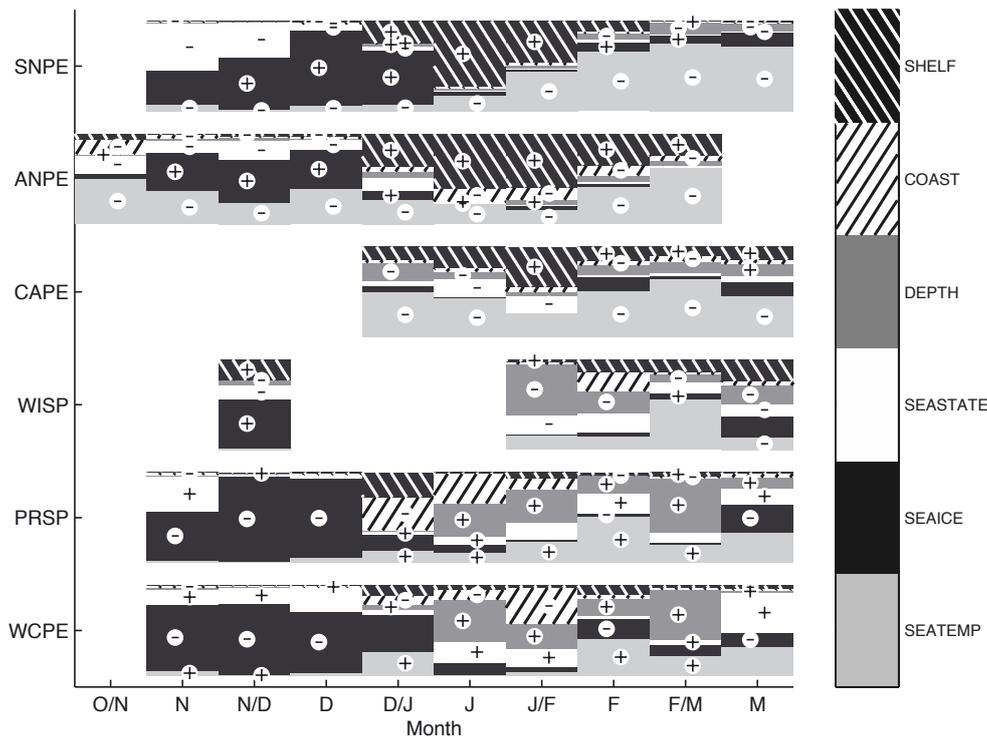


Fig. 3. Relative importance of environmental predictor variables for logistic models of at-sea observations of 6 species of seabird in the Prydz Bay region, East Antarctica. The height of a shaded segment shows the relative importance of that variable in predicting observations of a given species for a given 15 d period (see 'Materials and methods'). A positive sign indicates that the association was positive (i.e. the species in question was more likely to be seen with increasing values of the variable). A sign is shown only where the sign of the association was consistent across all cross-validation iterations. SNPE = snow petrels, ANPE = Antarctic petrels, CAPE = Cape petrels, WISP = Wilson's storm petrels, PRSP = prion species, WCPE = white-chinned petrels. SHELF = shelf break, COAST = coast, DEPTH = water depth, SEASTATE = sea state (height), SEAICE = % sea ice cover, SEATEMP = sea surface temperature

the Prydz Bay coastal regions. The distribution of Cape petrel sightings in Vincennes Bay was similar, with the exception that all sightings from January to March were restricted to the region of the Vincennes Bay coast (S of approx 60°S), as there is no northerly breeding-locality equivalent to Heard Island in Vincennes Bay. The errors of the spatial logistic models of Cape petrel observations in Prydz Bay were lower than those of the null only during February and early March (Fig. 2). This finding reflects the distribution of breeding populations of this species in the Antarctic and subantarctic. For the majority of the season, indi-

viduals from colonies will forage widely within Prydz Bay, resulting in a broad at-sea distribution of observations. Around February, adults are feeding chicks and do not forage far from the colonies, leading to more predictable distributions of observations. When applied to the Vincennes Bay data, the model errors were significantly lower than the null error from late December until March (Fig. 2). Sea surface temperature was the dominant environmental predictor, with a negative coefficient (indicating an association with colder waters, Fig. 3).

Wilson's storm petrel

The distributions of observations of Wilson's storm petrels in Prydz Bay and Vincennes Bay were superficially similar to those of Cape petrels, but with a later initial influx (late November to early December, compared with October for Cape petrels), and with a less pronounced restriction to coastal areas during January and February. The spatial logistic models of observations of Wilson's storm petrels in Prydz Bay were no more accurate than the null (Fig. 2). Although the breeding distributions of Wilson's storm petrels are similar to those of Cape petrels, Wilson's storm

Table 2. Spearman correlations between environmental variables. For the purposes of this table, binary variables **coast** and **shelf** (see 'Materials and methods') have been replaced with **distance to nearest coast** and **distance to shelf break**. Asterisks denote significant correlations ($p < 0.01$, assessed using the method of Dutilleul 1993)

	Sea ice concentration	Sea state	Water depth	Distance to nearest coast	Distance to shelf break
Sea surface temperature	-0.69*	0.52*	0.34	0.42*	0.62*
Sea ice concentration		-0.64*	-0.37*	-0.50*	-0.64*
Sea state			0.31	0.44*	0.49*
Water depth				0.76*	-0.54*
Distance to nearest coast					0.77*

petrels tended to be observed further from the coast than Cape petrels during the later part of the season. This less pronounced restriction to coastal areas led to less accurate models in Prydz Bay. However, when applied to the Vincennes Bay data, the model errors were lower than the null from late January through to March. No clear pattern was evident in the predictor importances (Fig. 3).

Prion spp.

Sightings of prions in Prydz Bay and Vincennes Bay showed a southward advance throughout the season. Restricted to N of 60°S during October and November, prions were observed almost to the Prydz Bay and Vincennes Bay coasts by January. The Prydz Bay spatial logistic model errors were generally lower than the null model errors, but not so during late October or late December (Fig. 2). For Vincennes Bay, the model errors were lower than the null, except in late October and January. Sea ice concentration (negative coefficient) and sea state (positive coefficient) were important during November and December (Fig. 3), indicating an association with ice-free waters during this time. For the remainder of the season, depth, sea surface temperature, and sea state were important (each with a positive coefficient, indicating an association with the deeper, warmer waters of northern Prydz Bay).

White-chinned petrel

White-chinned petrel sightings in Prydz Bay were distributed in a similar manner to those of prions, showing a southward trend throughout the season. The distribution of white-chinned petrel sightings in Vincennes Bay also showed a southward advance, but with a bias to the western half of the region (W of 110°E) from January to March. Strikingly, there were few sightings of white-chinned petrels in Vincennes Bay during March, when sightings were common in Prydz Bay. The spatial logistic model errors for white-chinned petrels in Prydz Bay were less than the null, except during late October (Fig. 2). With the Vincennes Bay data, the model errors were lower than the null model from November to early January, but generally not so during the remainder of the season. The importances of the predictors were similar to those of prions, showing an association with ice-free waters during November and

December, and with the deeper, warmer waters in northern Prydz Bay from January to March (Fig. 3).

General results

The standard logistic model (results not shown) was in general only marginally less accurate than the spatial logistic model. The differences reached statistical significance (Wilcoxon paired-sample test at $p < 0.05$) on 5 data sets: snow petrel observations in Vincennes Bay, Antarctic petrels (both in Prydz Bay and Vincennes Bay), prions (Prydz Bay), and white-chinned petrels (Vincennes Bay). However, these differences were negligible: the largest median difference in predictive accuracy occurred for the models for white-chinned petrels in Vincennes Bay (0.9%). In one case (Cape petrels in Prydz Bay), the logistic model error was lower than that of the spatial logistic model, but the median difference was small (0.4%).

The size of the spatial bins used here (2° longitude by 2° latitude) was largely an arbitrary choice. To ensure that the results of these analyses were not an artefact of the selected spatial scale, we repeated the analyses at differing bin sizes. Fig. 4 shows the prediction error of the logistic model of sightings of prions in Prydz Bay (averaged over all time bins) for varying choices of spatial bin size. The model error was unchanged across bin sizes ranging between 0.5° longitude by 0.5° latitude and 3° longitude by 2° latitude. Results for other species were similarly unaffected by choice of spatial bin size.

DISCUSSION

Predictive ability of the models

Comparable model performance on the Prydz Bay and Vincennes Bay data sets indicates that the internal species-environment relationships developed by the

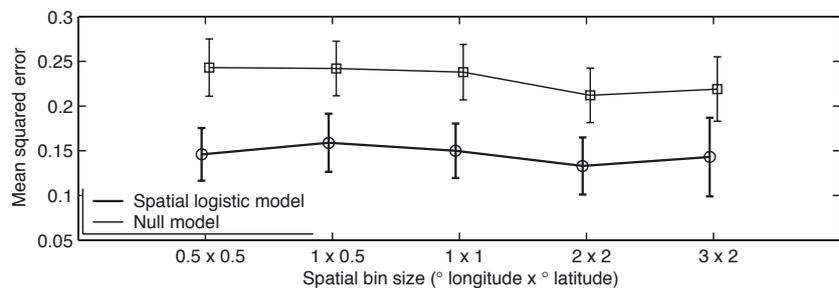


Fig. 4. Mean square prediction errors (\pm SD) of a spatial logistic model of at-sea observations of prions (*Pachyptila* spp.) in the Prydz Bay region with varying spatial bin size. Points have been offset slightly on the x-axis to avoid overlap

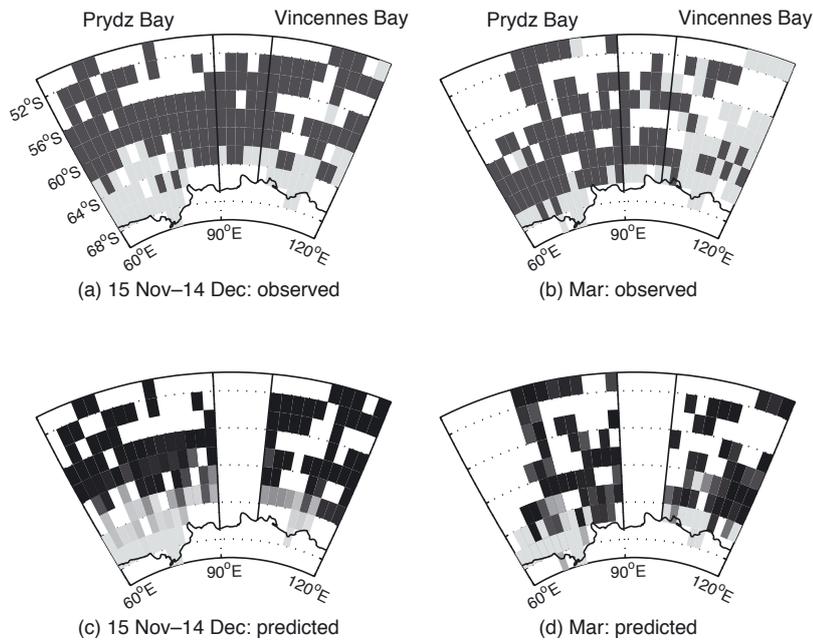


Fig. 5. Actual observations and predicted probability of observations of white-chinned petrels in the Prydz Bay and Vincennes Bay regions, from 1980 to 2002. Data in (a) and (c) are for mid-November to mid-December, and (b) and (d) are for March. The empty arc of bins shows the separation between the area used to develop the models (Prydz Bay) and that used for testing (Vincennes Bay). No predictions were made for the sector between 90 and 100° E. For actual observations, dark grey = observed, light grey = not observed, white = no data. Predicted probabilities range from 0 (light grey) to 1 (black). Data from all years have been pooled

model for Prydz Bay were also relevant for Vincennes Bay. This was found to be the case for snow and Antarctic petrels. For these 2 species it can be seen that the model error on the Prydz Bay data was approximately constant throughout the season, and similarly so with the Vincennes Bay data (Fig. 2). These results suggest a broader applicability of the models for these 2 species to other areas of the Southern Ocean adjacent to East Antarctica, and perhaps farther afield.

The remaining 4 species showed discrepancies between the Prydz Bay and Vincennes Bay model errors. For white-chinned petrels, the model errors on the Prydz Bay and Vincennes Bay data were similar from November until late December. From January through to March, however, the average Vincennes Bay model error (0.37 ± 0.09) was more than 4 times that in Prydz Bay (0.08 ± 0.02). Clearly, the processes governing the distribution of white-chinned petrels in Prydz Bay differ from those in Vincennes Bay during the latter half of the summer. Fig. 5b shows that observations of white-chinned petrels during March are widespread in Prydz Bay, but uncommon in Vincennes Bay. Earlier in the season, there are no obvious differences in the distributions of white-chinned petrel sightings in the 2 bays (Fig. 5a). Fig. 5c,d shows the associated model

predictions. White-chinned petrels overwinter in the subantarctic, and these data suggest an earlier departure from Vincennes Bay than from Prydz Bay. A similar but less pronounced pattern was seen in the observations of prions. The average model error in Prydz Bay from late January to March was 0.10 ± 0.03 ; in Vincennes Bay it was 0.19 ± 0.02 during the same period. As with white-chinned petrels, prions overwinter in the subantarctic and temperate regions of the Southern Hemisphere and these results suggest differences in departures from summer foraging areas around the Antarctic.

For Cape petrels and Wilson's storm petrels, the model performed better in Vincennes Bay than Prydz Bay during the latter part of the season (approximately January to March, Fig. 2). This is an unusual result. It is more common to find the performance on an independent test set (i.e. Vincennes Bay) to be worse than that obtained with the data used to build the model (Prydz Bay). Usually, this is a result of overfitting the model to the training data (Hastie et al. 2001). However, the situation has been reversed in this instance. This is a result

of topological differences between Prydz Bay and Vincennes Bay: specifically, the presence of Heard Island in the north of Prydz Bay, for which there is no equivalent in Vincennes Bay. We have already noted that Cape petrels and Wilson's storm petrels breed both on the Antarctic continent and at subantarctic locations, such as Heard Island. Observations of these birds suggest that from January to March, they forage in waters that are relatively proximate to their colonies (Woehler et al. 1991), and the models have captured this relationship (see Fig. 3). In Prydz Bay, they have widespread distributions, and the model accuracies are consequently little or no better than the null. In Vincennes Bay, where the colonies are only on the Antarctic coast, the models are more accurate. This finding illustrates that even though a model may not be sufficiently accurate for practical predictive use, the internal species-environment relationships formed by the model can still be instructive.

Spatial logistic model

The differences in predictive accuracies between the standard and spatial logistic models were small. In

order to obtain good performance with the standard logistic model, binary predictor variables relating to physical features (namely the coast and shelf break) were replaced by the distances to those features. Although this yielded good predictive accuracies, the species-environment relationships (predictor variable importances) were often less intuitive and ecologically plausible than those obtained from the spatial logistic model. For example, the standard logistic model indicated that the distance to the coast was the most important predictor of Antarctic petrel observations during November. Using the same data, the spatial logistic model indicated that the most important predictor was sea ice concentration. Examination of the data revealed that the sea ice during November extended to approximately 900 km from the coast, and it was this distance that the standard logistic model had identified as important. Clearly, it is sea ice rather than relative proximity to the coast that is the more likely driver of Antarctic petrel observations in this scenario. This finding is in agreement with other applications of spatial models in ecology (e.g. Leathwick 1998, Keitt et al. 2002), which indicate that the inclusion of a term that explicitly models spatial correlation can markedly change the conclusions drawn from species-environment models.

Importance of variables

For the first part of the season (approximately November to early January) the sea ice concentration and its covariate sea state were the most important predictors of seabird observations (Fig. 3). Snow and Antarctic petrels were ice-associated (observed near to or within the pack ice), whereas the reverse was true for prions and white-chinned petrels. This finding is in agreement with earlier work (Fraser & Ainley 1986, Ainley et al. 1993). From late December to February, low sea surface temperatures and the shelf break were important variables for snow, Antarctic, and Cape petrels. This is probably indicative of inshore foraging relatively proximate to nesting sites, rather than a strict association with the shelf break. The observational data suggest that snow petrels in particular were associated with areas near to remnant sea ice, although this association was not identified by the models. The other physical environmental parameters that were collected with the at-sea observations (sea level air pressure, precipitation, wind force, sea state, and cloud cover) were found to have poor predictive power and were not used in the models. This finding does not necessarily imply a lack of interaction between these elements of the physical environment and seabird observations. Rather, it might indicate that there is a mismatch

between the scales (both temporal and spatial) of the analyses presented here, and the scales of these environmental processes. For example, a temporal scale of hours or days would probably be appropriate for assessing variations in precipitation. This is a much finer temporal scale than the 30 d time windows used in this study.

Limitations of the study

One limitation of this study was the opportunistic nature of data collection. Voyages were undertaken primarily for logistic purposes (station resupply and personnel movement) rather than dedicated surveys. Seabird observers thus had little control over ship tracks, and a designed sampling strategy was not possible. The most obvious consequence of this is the lack of data before October and after March; sea ice during the winter months makes ship travel to the Antarctic continent difficult. Other more subtle variations in survey density exist, such as the relative paucity of observations north of 56°S during most months and the frequent surveying of areas near to Australia's Antarctic stations (see Fig. 1). Our analysis methods were chosen to minimise the effects of these imbalances. Presence/absence data are likely to be more robust than abundance data where survey effort is variable. The use of presence/absence data also avoids possible inter-observer bias in abundance estimates. However, presence/absence data are susceptible to underestimation of ranges of species (failing to detect presence), particularly in bins with few surveys. Bins with a low number of surveys were therefore discarded during analyses (see 'Materials and methods'). More formal methods of incorporating survey effort into the analyses exist, for example, by estimating the probability of detecting species which are actually present in the survey area (e.g. Stauffer et al. 2001).

The model selection and validation procedures were carried out using voyage as a sampling unit. This avoided over-fitting models to unusual characteristics of individual voyages, and ensured that the cross-validation estimates of prediction error were not artefactually optimistic. The similarity of prediction errors on the Prydz Bay and Vincennes Bay data in most cases indicates that this strategy was successful. Voyages also tended to be aligned along the dominant environmental gradient (north to south), leading to a natural approximation of the gradsect sampling method (Gillison 1984).

Some of the predictor variables used in this study are surrogates for the relevant physical processes. Breeding birds (snow, Antarctic, Cape petrels and Wilson's storm petrels) appear to remain in relative proximity to

their nesting sites during the later part of the season (approximately late December through to February). Nesting site data might therefore be a useful addition to the models. However, any model using this information could not be applied to areas adjacent to unsurveyed nesting areas. Further, there are no independent data on prey distribution and abundance. Clearly, such data might be expected to contribute to an increased predictability in the distribution of foraging seabirds.

The models presented here are linear with respect to the predictor variables. Non-linear terms or interactions between variables might yield models which are ecologically more plausible (Austin 2002). Generalised additive models (Hastie & Tibshirani 1990) are commonly used for species-environment analyses for this reason. The aim of this study was to demonstrate the applicability of predictive models, and to compare spatial and non-spatial versions of a relatively simple model.

Acknowledgements. We thank all observers who have recorded seabird-at-sea observations over the past 22 yr under, at times, difficult conditions. David Watts and David Smith (Australian Antarctic Data Centre) provided database and GIS support. The input of 3 anonymous reviewers improved the manuscript considerably.

LITERATURE CITED

- Ainley DG, O'Connor EF, Boekelheide RJ (1984) The marine ecology of birds in the Ross Sea. *Ornithol Monogr* 32
- Ainley, DG, Ribic CA, Spear LB (1993) Species-habitat relationships among Antarctic seabirds: a function of physical or biological factors? *Condor* 95:806–816
- Ainley DG, Ribic CA, Fraser WR (1994) Ecological structure amongst migrant and resident seabirds of the Scotia-Weddell Confluence region. *J Anim Ecol* 63:347–364
- Ainley DG, Jacobs SS, Ribic CA, Gaffney I (1998) Seabird distribution and oceanic features of the Amundsen and southern Bellingshausen Seas. *Antarct Sci* 10:111–123
- Augustin NH, Mugglestone MA, Buckland ST (1996) An autologistic model for the spatial distribution of wildlife. *J Appl Ecol* 33:339–347
- Augustin NH, Mugglestone MA, Buckland ST (1998) The role of simulation in modelling spatially correlated data. *Environmetrics* 9:175–196
- Austin MP (2002) Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecol Model* 157:101–118
- Besag J (1974) Spatial interaction and the statistical analysis of lattice systems (with discussion). *J R Stat Soc B* 36: 192–236
- Besag J (1975) Statistical analysis of non-lattice data. *Statistician* 24:179–195
- Carroll SS, Pearson DL (2000) Detecting and modeling spatial and temporal dependence in conservation biology. *Conserv Biol* 14:1893–1897
- Cavaliere D, Parkinson C, Gloerson P, Zwally HJ (1999, updated 2002) Sea ice concentrations from Nimbus-7 SMMR and DMSP SSM/I passive microwave data. CD-ROM. National Snow and Ice Data Center, Boulder, CO
- Cawsey EM, Austin MP, Baker BL (2002) Regional vegetation mapping in Australia: a case study in the practical use of statistical modelling. *Biodiversity Conserv* 11: 2239–2274
- Cressie NAC (1993) *Statistics for spatial data* revised edition. Wiley-Interscience, New York
- Croll DA (1990) Physical and biological determinants of the abundance, distribution, and diet of the common murre in Monterey Bay, California. *Stud Avian Biol* 14:139–148
- Croxall JP (1984) Seabirds. In: Laws RM (ed) *Antarctic ecology*. Vol. 2. Academic Press, London, p 533–619
- De'ath G (2002) Multivariate regression trees: a new technique for modeling species-environment relationships. *Ecology* 83:1105–1117
- Dutilleul P (1993) Modifying the *t* test for assessing the correlation between two spatial processes. *Biometrics* 49: 305–314
- Franklin J (1995) Predictive vegetation mapping: geographic modeling of biospatial patterns in relation to environmental gradients. *Progr Phys Geogr* 19:474–499
- Fraser WR, Ainley DG (1986) Ice edges and seabird occurrences in Antarctica. *BioScience* 36:258–263
- Geyer CJ, Thompson EA (1992) Constrained Monte Carlo maximum likelihood for dependent data (with discussion). *J R Stat Soc B* 54:657–699
- Gillison AN (1984) Gradient oriented sampling for resource surveys—the gradsect method. In: Myers K, Margules C, Musto I (eds) *Survey methods for nature conservation*. Vol. 2. CSIRO Division of Water and Land Resources, Canberra, p 349–374
- Gittins R (1985) *Canonical analysis—a review with applications in ecology*. Springer-Verlag, Berlin
- Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology. *Ecol Model* 135:147–186
- Hastie T, Tibshirani RJ (1990) *Generalised additive models*. Chapman & Hall, London
- Hastie T, Tibshirani RJ, Friedman J (2001) *The elements of statistical learning*. Springer-Verlag, New York
- Högmander H, Møller J (1995) Estimating distribution maps from atlas data using methods of statistical image analysis. *Biometrics* 51:393–404
- Huffer FW, Wu H (1998) Markov Chain Monte Carlo for autologistic regression models with application to the distribution of plant species. *Biometrics* 54:509–524
- Keitt TH, Bjørnstad O, Dixon P, Citron-Pousty S (2002) Accounting for spatial pattern when modeling organism-environment interactions. *Ecography* 25:616–625
- Leathwick JR (1998) Are New Zealand's *Nothofagus* species in equilibrium with their environment? *J Veg Sci* 9:719–732
- Leathwick JR (2001) New Zealand's potential forest pattern as predicted from current species-environment relationships. *NZ J Bot* 39:447–464
- Leopold MF (1993) Seabirds in the shelf edge waters bordering the Banc d'Arguin, Mauritania, in May. *Hydrobiologia* 258:197–210
- LeSage JP (2000) Bayesian estimation of limited dependent variable spatial autoregressive models. *Geogr Anal* 32: 19–35
- McCullagh P, Nelder JA (1989) *Generalized linear models*. 2nd edn. Chapman & Hall/CRC, Boca Raton, FL
- Mehlum F, Hunt GL Jr, Klusek Z, Decker MB, Nordlund N (1996) The importance of prey aggregations to the distribution of Brünnich's guillemots in Storfjorden, Svalbard. *Polar Biol* 16:537–547
- National Geophysical Data Center (1988) Data announce-

- ment 88-MGG-02: digital relief of the surface of the earth. NOAA, National Geophysical Data Center, Boulder, CO
- Orsi A, Whitworth T III, Nowlin WD Jr (1995) On the meridional extent and fronts of the Antarctic Circumpolar Current. *Deep-Sea Res* 42:641–673
- Ribic CA, Ainley DG (1988/89) Constancy of seabird species assemblages: an exploratory look. *Biol Oceanogr* 6: 175–202
- Schneider DC (1990) Spatial autocorrelation in marine birds. *Polar Res* 8:89–97
- Scott PA, Lavoie C, MacDonald GM, Sveinbjörnsson B, Wein RW (1997) Climate change and the future position of the Arctic treeline. In: Oechel WC, Callaghan T, Gilmanou T, Holten JI, Maxwell B, Molau U, Sveinbjörnsson B (eds) *Global change and arctic terrestrial ecosystems*. Springer-Verlag, New York, p 245–265
- Smith RC, Ainley D, Baker K, Domack E and 7 others (1999) Marine ecosystem sensitivity to climate change. *Bio-Science* 49:393–404
- Soluri EA, Woodson VA (1990) World vector shoreline. *Int Hydrogr Rev* LXVII(1):27–36
- Stauffer HB, Ralph CJ, Miller SL (2001) A component of accuracy assessment: incorporating uncertainty of detection into presence-absence survey design and analysis. In: Scott JM, Heglund PJ, Morrison M, Raphael M, Hauffer J, Wall B (eds) *Predicting species occurrences: issues of scale and accuracy*. Island Press, Covello, CA
- Stone M (1974) Cross-validatory choice and assessment of statistical predictions. *J R Stat Soc* 39:44–47
- Sykes MT (1997) The biogeographic consequences of forecast changes in the global environment: individual species' potential range changes. In: Huntley B, Cramer W, Morgan AV, Prentice HC, Allen JRM (eds) *Past and future rapid environmental changes: the spatial and evolutionary responses of terrestrial biota*. NATO Series 1: Global Environmental Change. Springer-Verlag, New York, p 427–440
- ter Braak CJF, Prentice IC (1988) A theory of gradient analysis. *Adv Ecol Res* 18:271–317
- Woehler EJ (1995) Variability in foraging ecology and food consumption by seabirds at high latitudes. PhD thesis, University of California, Irvine
- Woehler EJ (1997) Seabird abundance, biomass and prey consumption within Prydz Bay, Antarctica, 1980/81 to 1992/93. *Polar Biol* 17:371–383
- Woehler EJ, Hodges CL, Watts DJ (1991) An atlas of the pelagic distribution and abundance of seabirds in the southern Indian Ocean. *ANARE Res Notes* 77
- Woehler EJ, Raymond B, Watts DJ (2003) Decadal-scale seabird assemblages in Prydz Bay, East Antarctica. *Mar Ecol Prog Ser* 251:299–310

Editorial responsibility: Otto Kinne (Editor), Oldendorf/Luhe, Germany

*Submitted: May 19, 2003; Accepted: October 5, 2003
Proofs received from author(s): November 24, 2003*