A novel method for quantifying habitat selection and predicting habitat use

Carla Freitas[1,2]*, Kit M. Kovacs[1], Christian Lydersen[1] and Rolf A. Ims[2]

[1]*Norwegian Polar Institute, Polar Environmental Centre, N-9296 Tromsø, Norway;* [2]*Institute of Biology, University of Tromsø, N-9037 Tromsø, Norway*

Summary

1. The study of habitat selection and habitat use are crucial for understanding the biological requirements of animals and the strategies they use to fulfil their needs. A variety of statistical techniques are available to quantify habitat selection, most of them based on the comparison of habitat attributes in sites used by the animals and in unused, but available, sites. Because of the difficulties in defining what is available from an animal's perspective for many species, statistical approaches such as first-passage time (FPT) analysis have been developed, which explore habitat-use intensities through the areas used by animals.

2. In this study, we expand on that approach by using FPT-derived data within the framework of mixed-effects Cox proportional hazards models (CPH models) to quantify habitat selection. By modelling FPTs, CPH models evaluate which habitat attributes best explain the 'risk' or 'hazard' of the animal leaving an area of a given radius. For quantitative interpretations, the coefficients of the CPH model can be used to calculate hazard ratios that estimate relative habitat preferences: the lower the hazard ratio (i.e. the lower the risk of leaving), the higher the preference. Accounting for individual variability in habitat use by means of a random-effect term added to CPH models provides inferences that more appropriately reflect the typical hierarchical structure of telemetric data on animal space use.

3. *Synthesis and applications.* In addition to providing inferences about habitat selection based on the estimated parameters, survival functions of fixed-effects CPH models can be used to construct spatial predictions, for instance maps of population-level space use. Such predictions can be particularly useful for applied purposes, for example as a basis for species conservation plans and reserve selection. The method is illustrated using two data sets from marine mammal species: ringed seals *Phoca hispida* and white whales *Delphinapterus leucas*, but it is broadly applicable to habitat selection and prediction studies of other highly mobile animals in marine or terrestrial systems.

Key-words: Cox regression, *Delphinapterus leucas*, first-passage times, habitat use, mixed-effects Cox proportional hazards models, *Phoca hispida*, random effects models, satellite telemetry, survival analysis

Introduction

Habitat use and selection studies are essential for understanding the biological requirements of animals and the strategies they use to fulfil their needs (Manly *et al.* 2002). In addition, information on habitat use patterns is crucial for conservation and management purposes (Scott *et al.* 2002; Guisan & Thuiller 2005).

A variety of statistical approaches are available to quantify habitat selection (Manly *et al.* 2002; Strickland & McDonald 2006). Commonly, these methods use Resource Selection Functions (RSFs) that compare resource attributes in areas where animals are observed and in areas that are considered available. A major difficulty when comparing used vs. available areas is the definition of 'available habitat', which critically alter the results (e.g. Johnson 1980; Hjermann 2000). Defining areas as available habitat implies that animals know about this availability and make their movement decisions based on this information (Fauchald & Tveraa 2003). Alternatively, habitat selection can be inferred from the intensity of use of habitat conditions only in sites experienced by the animals. One way to quantify intensity of habitat use along an

*Correspondence author. E-mail: carla@npolar.no

animals' movement trajectory is to measure first-passage time (FPT; Fauchald & Tveraa 2003). FPT is defined as the time required for an animal to cross a circle of a given radius *r*. Using FPT as a response variable in statistical habitat models does offer some challenges because it violates the assumptions underlying traditional parametric models such as Gaussian Generalized Linear Models (GLMs). However, since FPTs are continuous event–time measurements, methods which model time until an event occurs, can be used. The most popular statistical model used for this purpose is the Cox proportional hazards model (CPH model, Cox 1972). It has been used extensively in medical research to model survival times and in failure–time analysis of industrial products. CPH models have also been applied in resource selection studies to model the time until a resource (food or habitat) is selected (see examples in Manly *et al.* 2002, Chapter 6) and to investigate patch-leaving decisions in foraging studies (e.g. Wajnberg *et al.* 2003). A recent development of CPH models that incorporates random-effect terms (Pankratz, de Andrade & Therneau 2005) makes these models suitable to perform analyses on data with more than one stratum of variation. This makes them more suitable for use with location data collected in telemetry studies. In such analyses, it is essential to distinguish between random variation between individuals versus within-individual variation by including random-effect terms in the analyses.

This study proposes the use of mixed-effects CPH models to infer habitat selection from FPTs. It shows how the models can: (i) identify environmental variables that best explain the time spent in different areas, taking into account the individual variability; (ii) measure how animals respond to environmental conditions, by calculating relative habitat preferences; and (iii) predict habitat use from survival functions. This novel combination of methods is illustrated using two data sets from marine mammal species – ringed seals *Phoca hispida* and white whales *Delphinapterus leucas*.

Materials and methods

MODELLING INTENSITY OF HABITAT USE

The first step in the present method consists of calculating FPTs for the track of each individual animal at regular space intervals (see Fauchald & Tveraa 2003 and below). Once this information is available, together with measurements of habitat variables at the same locations where FPTs were calculated, CPH models can be used to investigate how FPTs (and hence animal movements) are influenced by the measured variables. The CPH model is written as:

$$h(t) = \exp(\beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \beta_p X_p)h_0(t), \qquad \text{eqn 1}$$

where $h(t)$ is the hazard function, i.e. the risk that an animal leaves an area at time *t*. $X_1, X_2, X_3, \ldots, X_p$ are the explanatory variables in the model, and $\beta_1, \beta_2, \beta_3, \ldots, \beta_p$ are the coefficients that describe the contribution of these variables. $h_0(t)$ is the baseline hazard function at time *t* (i.e. the risk of leaving an area where all explanatory variables are equal to zero or to a defined base value). In order to take individual variability into account, a random-effect term (*b*) can be added to the general CPH model, as described in Pankratz, de Andrade & Therneau (2005). In this case, the CPH model becomes:

$$h(t) = \exp(\beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \ldots + \beta_p X_p + b)h_0(t), \qquad \text{eqn 2}$$

where *b* is the per-subject random effects, which is assumed to be normally distributed with a mean of zero. Similar to the case for linear mixed-effects models (see Pinheiro & Bates 2000), the addition of this term enables estimation of the response of an average individual using the population mean (fixed-effects) and in addition provides a measure of the variability in the response across individuals (random-effects). The standard CPH model (equation 1) which assesses only fixed-effects is hereafter referred to as the fixed-effects CPH model, while models including the individual random effects are termed mixed-effects CPH models.

In both equations above, the baseline hazard function $h_0(t)$ is estimated from the β coefficients, while the β and *b* coefficients, which are the unknown parameters in the model, are estimated using the method of maximum likelihood (see Collett 2003; Pankratz, de Andrade & Therneau 2005). Hazard ratios (HR) can be calculated from the exponential of β (e^β) and used to get rigorous interpretations of covariate effects (Collett 2003; Murray 2006). For a continuous *X* variable, an HR equal to 1·8, for example, indicates that the risk of leaving increases 1·8 times (80%) per unit of *X*. For a categorical *X* variable with two or more levels, where the first level, for example, is considered the base level, an HR equal to 0·6 indicates that the risk of leaving when *X* assumes that level is 0·6 times (40% lower) the risk of leaving when *X* is equal to the base level. An HR greater than one represents a higher risk of leaving, while an HR lower than one is interpreted in the opposite way. Using this approach, the hazard ratios can be used to infer habitat selection, under the assumption that lower risk of leaving correspond to increased preference (affinity) for a given habitat condition.

PREDICTING HABITAT USE

Once β coefficients are estimated for equation 1, survival functions, which describe the probability of being in an area longer than a time *t*, can be obtained. The estimated survival function for a given *i*th individual is given by:

$$\hat{S}_i(t) = (\hat{S}_0(t))^{\exp(\hat{\beta}' x_i)}$$

where x_i is the vector of values of the explanatory variables for the *i*th individual, $\hat{\beta}$ is the vector of estimated coefficients (from equation 1) and $\hat{S}_0(t)$ is the baseline survivor function. (cf. for instance Collett 2003). Since $\hat{S}_i(t)$ predicts the probability of using an area longer than a time *t* under a given set of variables x_i, it can be used as a predictive measure of the intensity of use of that area. Survival functions $\hat{S}_i(t)$ can be predicted for example for an individual in different locations with specific values of explanatory variables. The estimated probability of staying in those locations longer than a given time *t* (e.g. longer than 24 h) can be extracted from these functions. If desired, the obtained values can be plotted to obtain a cartographic prediction of expected habitat-use intensity.

APPLICATION TO RINGED SEAL AND WHITE WHALE SATELLITE-TRACKING DATA SETS

The ringed seal is an arctic species that usually spends winter and spring in areas of annually formed sea ice (fast ice) inside fjords and bays (McLaren 1958; Lydersen & Gjertz 1986). Females give birth inside snow lairs in March–early April; mating takes place about 1

month later (Lydersen 1998). Following mating, ringed seals use the remaining annual sea ice as a platform for moulting (Smith 1987; Lydersen 1998). In summer, when the moulting season is over and the annual sea ice has melted, ringed seals leave coastal breeding and moulting areas to spend the rest of the year elsewhere. Ringed seals, similar to most ice-breeding seals, are quite nomadic for much of the year, unlike coastal seals (e.g. grey *Halichoerus grypus* or common seals *Phoca vitulina*) that are colonial and return to familiar, fixed locations on a regular schedule outside the breeding season. The present study uses a data set consisting of satellite-tracking locations from 18 ringed seals that were tagged in Svalbard, Norway, after the moulting seasons of 2002 and 2003 (Supplementary Material Table S1). It is a subset of an original data set consisting of 22 seals, for which two types of movements were observed (inshore and offshore trips; Freitas *et al.* 2008a). The present study only presents coastal movements collected between August and December of 2002 and 2003. Tagging methods and other details regarding the satellite-relay data loggers (SRDLs) used on the ringed seals can be found in Lydersen *et al.* (2004). The SRDLs were programmed to send data whenever possible, without duty cycling.

The white whale is also an arctic species. This species can be found in coastal and estuarine areas and in some seasons also in offshore areas, sometimes along ice-edges; large-scale movement patterns seem to vary markedly among populations from different regions (Smith & Martin 1994; Richard, Heide-Jørgensen & St Aubin 1998; Richard, Martin & Orr 2001; Lydersen *et al.* 2001). The data set used here consists of satellite-tracking locations for 12 adult animals tagged in Svalbard from 1997 to 2001 (Table S1). Field methods and tag specifications can be found in Lydersen *et al.* (2001, 2002). As for the ringed seals, SRDLs deployed on the white whales were set to send data at any time, with no duty cycling.

Both ringed seal and white whale satellite locations were filtered using the algorithm described in Freitas *et al.* (2008b) and available at the R software package 'argosfilter' (http://cran.r-project.org/; function sdafilter with all default settings). Filtering aimed to remove locations requiring unrealistic swimming speeds (> 2 m s^{-1}) and unlikely turning angles (> 165 degrees or > 155 degrees, depending on the length of such path deviations). The post-filter paths had a mean ($\pm$ SE) distance between locations of $3 \cdot 5 \pm 0 \cdot 07$ km. White whale paths had a mean distance between locations of $3 \cdot 9 \pm 0 \cdot 04$ km. Mean time between locations was $1 \cdot 6 \pm 0 \cdot 28$ h for ringed seals and $1 \cdot 2 \pm 0 \cdot 06$ h for white whales.

FPTs were then calculated from the filtered tracks, at 5-km intervals. For ringed seals, FPTs were calculated for radii ranging from 5 to 100 km, at 5-km increments from 5 to 20 km and 10-km increments thereafter. FPTs for white whales were calculated for radii ranging from 5 to 80 km, at 5-km increments. The variance of FPT (log-transformed) was then plotted for each individual as a function of radii, in order to find the radius of maximum variance. This radius corresponds to the spatial scale at which animals concentrate their search effort (Fauchald & Tveraa 2003), and it is also the scale that best differentiates between low FPT areas (transitory areas) and high FPT areas (intensive search areas). Although the scale may vary among different individuals, finding a common spatial scale *SR* (Fauchald & Tveraa 2006) is recommended. This is because all estimates derived from the hazard and survival functions must be scaled to a specific unit area. The radius corresponding to the mean maximum variance is a natural choice for a common *SR*. This was found to be 10 km for ringed seals and 15 km for white whales; both values corresponded to distinct peaks in the mean variances (Fig. 1). In cases of flat or multi-modal variance profiles, there may be a need to stratify data and analyse individuals separately
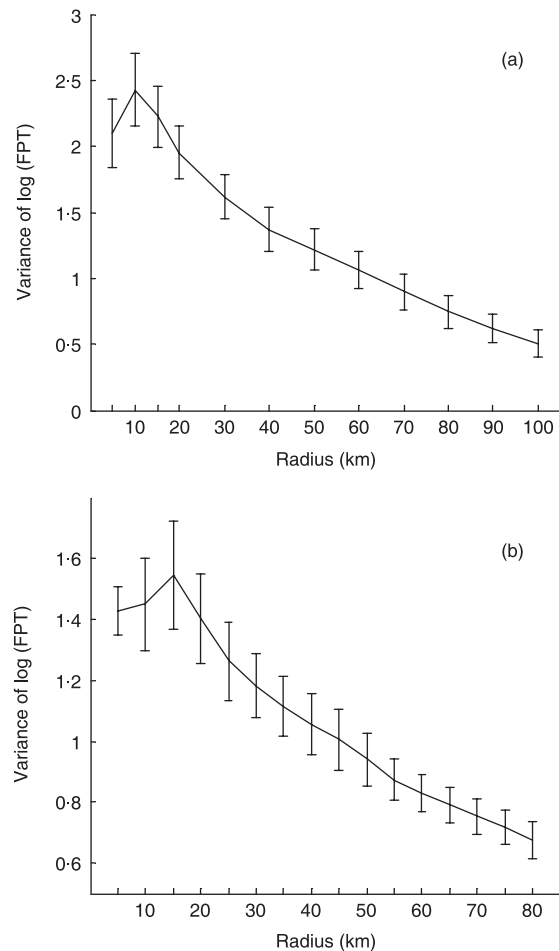

Fig. 1. Mean variance ($\pm$SE) in log-transformed first-passage time (FPT) as a function of the radius for 18 ringed seals satellite-tracked in Svalbard in 2002 and 2003 (a), and for 12 white whales tracked in Svalbard between 1997 and 2001 (b). The peak in variance occurred at the radius of 10 km for ringed seals and 15 km for white whales.

with different scales and associated ecological characteristics (e.g. ringed seals with different large-scale movement tactic; see Freitas *et al.* 2008a).

A set of explanatory variables that could potentially influence FPTs was obtained for the same locations where FPTs were calculated. These included: sea bottom depth; distance to the coast; and distance to the nearest glacier front. Additionally, sea surface temperature was recorded for ringed seals and presence or absence of sea ice for white whales. No sea ice information was available for the coastal areas where the ringed seal tracks were obtained. Seabed depths were extracted from $2 \cdot 5$ km resolution grid data from IBCAO (International Bathymetric Chart of the Arctic Ocean, Version $1 \cdot 0$, 2001). Distances to the coast and to the nearest glacier front were calculated from digital maps produced by the Norwegian Polar Institute (updated using aerial photographs of glacier fronts and coastlines taken from 1993 to 1998). Sea surface temperatures were obtained from the temperature data collected by the ringed seals themselves (see Lydersen *et al.* 2004). Temperatures, collected at irregular points along the seals tracks, were linearly interpolated to the positions of interest (5-km intervals along the tracks of the seals). The presence or absence of sea ice along the white whale

tracks was obtained from daily sea ice maps produced by the Norwegian Meteorological Institute. These charts are constructed using DMSP/SSM/I satellite data having a resolution of 25 km, NOAA/AVHRR satellite data with a 1·5-km resolution and observations from ships and aircraft.

Using FPTs at the *SR* radii as survival times, and the environmental variables listed above as explanatory variables, mixed-effects CPH models were fitted. Standard fixed-effects CPH models were also fitted for comparison. Although FPTs and environmental variables were calculated at intervals of 5 km along the individual tracks, only a subset of those locations, at intervals equal to *SR* was used. We recommend not using intervals larger than *SR* because this could result in high FPTs being missed (unless the highest FPTs are previously selected, as performed by Suryan *et al.* 2006). On the other hand, the use of smaller intervals, especially smaller than the resolution of the environmental data (2·5 km in this case) would cause serial autocorrelation in the data. This latter problem appeared not to be prevalent when using intervals equal to the *SR* radii.

All possible combinations of variables were used during model selection, since the number of explanatory variables was not high. The following interaction terms were also considered for white whales: depth × presence of ice; distance to coast × presence of ice; and distance to glacier × presence of ice. These interactions were considered in order to investigate whether FPTs at different depths, distances to the coast and distances to glaciers were different depending on whether sea ice was present or not.

Selection between the candidate models was done using the AIC corrected to the effective sample size (AIC$_c$)

$$\mathrm{AIC}_c = -2\log(L) + 2k + \frac{2k(k+1)}{n-k-1},$$

where *n* is the sample size (see Burnham & Anderson 2002), in this case the total number of FPT locations. AIC$_c$ values in the mixed-effects CPH models were calculated using penalized log likelihoods (as log(*L*)), which are the log likelihoods from the standard CPH model, treating β and *b* as ordinary covariates minus a quadratic penalty on *b* (see Pankratz, de Andrade & Therneau 2005). In addition, penalized degrees of freedom were used as the number of parameters in the model (*k*). Coefficients of determination (R^2) were also

calculated using the penalized log likelihoods. The assumption of proportional hazards, required by CPH models (see Therneau & Grambsch 2000; Collett 2003), was verified from fixed-effects CPH models, since it is not yet implemented for mixed-effects CPH models. This verification was done by checking the scaled Schoenfeld Residuals, both visually and by testing if their slope was zero (see Collett 2003). All variables used in the models satisfied this assumption.

Survival functions [$\hat{S}(t)$], predicting the probabilities of using an area longer than a time *t*, were also calculated from fixed-effects *h*(*t*) as the calculation of these functions from mixed-effects CPH models is not yet implemented. In order to do that, a grid with cell size of 2·5 km was created for the area of interest (Svalbard Archipelago) and all environmental variables used in *h*(*t*) were obtained for the central point of each grid cell *i*. A $\hat{S}_i(t)$ function was then estimated for each location *i*. The probability of being in the surrounding *SR* kilometres for longer than 24 h [$\hat{S}_i(24)$] was extracted from each survival function $\hat{S}_i(t)$ and plotted on a map, in order to generate a cartographic prediction of habitat-use intensities for those areas. Survival functions and CPH models were fitted using R software (packages kinship and survival).

Results

RINGED SEAL HABITAT USE

The ranking of alternative models (mixed-effects CPH models) for the ringed seal data set showed that habitat use was significantly affected by distance to glaciers, depth, distance to the coast, as well as time of year (Supplementary Material Table S2). The best model also included sea surface temperature although its effect was not significant when taking other variables and individual variability into account (see confidence intervals in Table 1). Distance to the nearest glacier front was the variable with greatest influence on habitat use; this variable explained 34% of the variability in the data (Table S2). The best model according to the AIC was the following:

Table 1. Estimated coefficients (β), hazard ratios (e^β) and 95% confidence intervals [CI (β)] of the CPH models (with and without random effects) for the covariates being selected according to AIC$_c$ for ringed seals (Table S2). Number of locations in each category is also given (*N*)

Variable	*N*	Random-effects included			No random-effects included		
		β	e^β	CI (β)	β	e^β	CI (β)
glac (0–5 km)	2945	–	–	–	–	–	–
glac (> 5–10 km)	880	0·983	2·671	0·88 1·09	1·060	2·885	0·97 1·15
glac (> 10–20 km)	998	0·922	2·515	0·81 1·04	0·932	2·538	1·00 1·25
glac (> 20 km)	543	1·072	2·921	0·91 1·23	1·069	2·912	0·94 1·20
coast	5366	0·014	1·014	0·00 0·02	0·019	1·019	0·01 0·03
depth	5366	0·008	1·008	0·01 0·01	0·006	1·006	0·00 0·01
month (August)	1304	–	–	–	–	–	–
month (September)	1535	0·301	1·352	0·22 0·39	0·107	1·113	0·03 0·18
month (October)	1555	0·321	1·379	0·21 0·43	0·045	1·046	−0·05 0·14
month (November)	601	−0·207	0·813	−0·35 −0·06	−0·378	0·685	0·50 −0·25
month (December)	371	−0·700	0·496	−0·87 −0·53	−0·732	0·481	−0·88 −0·58
temp	5366	−0·011	0·989	−0·04 0·01	−0·036	0·964	−0·05 −0·02

Model abbreviations: glac is distance to nearest glacier front; coast is distance to the coast (km); depth is sea bottom depth (m); month is the calendar month and temp is sea surface temperature (°C).

$$h(t) = \exp(\beta_i\, glac_i + 0.014\ coast + 0.008\ depth$$
$$+ \beta_j\, month_j - 0.011\ temp + b)\, h_0(t),$$

where $glac_i$ is the distance to the nearest glacier in the i category, $coast$ is the distance from the coast (km), $depth$ is the water depth (m), $month_j$ is the time of year in the j category (month), $temp$ is sea surface temperature (°C) and b is the per-individual random effect. All β coefficients for models with and without the random effects are presented in Table 1. The same best model structure was found in both cases, and in general, there was good consistency between β coefficients (Table 1). The biggest difference between the model types concerned the effect of sea surface temperature, which appeared to be much weaker and uncertain in the mixed-effects than in the fixed-effect model (Table 1). Fitting separate fixed-effects models to individual seals (not shown here) revealed that this discrepancy was due to contradictory responses among the individual seals to sea surface temperature. Apparently, this discrepancy was, to some extent, extracted by the random-effect term of the mixed model.

Based on the estimated coefficients (Table 1) several quantitative inferences could be made regarding ringed seal habitat selection. The risk of leaving an area was ~2·5–2·9 times higher when seals were located at more than 5 km from a glacier front than when this distance was less than 5 km (Table 1). For constant distances from a glacier, the risk of leaving an area increased with distance from the coast (at a rate of 14% 10 km^{-1}) and with water depth (at a rate of 8% 10 m^{-1}). Lower risks of leaving an area were observed in November and December relative to August (18·7% lower in November and 50·4% lower in December), while higher risks for leaving were found in September (35·2% higher) and October (37·9% higher) in relation to August.

The most parsimonious mixed-effects CPH model for ringed seals explained 40·9% of the variability in the data from all animals (Table S2). The variance component attributed to individual variability (b) was 0·271. The standard deviation of the per-individual random effects ($\sqrt{0.271} \approx 0.521$), indicates that the average spread of relative risk of leaving among individuals is $e^{0.521} \approx 1.68$, meaning that the per-individual variability in the risk of leaving is on average 68% higher or lower than the overall risk.

WHITE WHALE HABITAT USE

The ranking of alternative models (Supplementary Material Table S3) and model coefficients (Table 2) indicate that white whale FPTs were affected significantly by distance to glacier fronts, depth and presence of sea ice. Two models had similar fits to the data (see AIC$_c$ values and Akaike weights on Table S3). An additional interaction term included in one of the models (distance to the coast × presence of ice) did not seem to affect the fit of the model. Distance to the coast itself did not have a significant effect upon the risk of leaving when other variables were taken into account (see confidence intervals for this variable in Table 2). Distance to glacier fronts was again the most important explanatory variable, accounting for 21% of the variability (Table S3). The best model explained 27·5% of the variability in the data. Similar to ringed seals, lower risk of leaving was found in areas located less than 5 km from a glacier front. Probabilities of leaving increased 2·1 times in areas 5–10 km away from a glacier, 2·9 times when 10–20 km from glaciers and 3·7 times when the whales were at distances greater than 20 km from a glacier (Table 2). Depth also significantly affected the risk of leaving, with maximum preferences being found for depths of less than 10 m (significant

Table 2. Estimated coefficients (β), hazard ratios (e^β) and 95% confidence intervals [CI (β)] of the CPH models (with and without random effects) for the covariates being included in the selected model for white whales (Table S3). Number of locations in each category is also given (*N*)

Variable	*N*	Random-effects included			No random-effects included		
		β	e^β	CI (β)	β	e^β	CI (β)
glac (0–5 km)	2116	–	–	–	–	–	–
glac (> 5–10 km)	669	0·76	2·13	0·64 0·87	0·77	2·15	0·65 0·88
glac (> 10–20 km)	625	1·07	2·91	0·94 1·19	1·02	2·76	0·89 1·14
glac (> 20 km)	410	1·31	3·71	1·15 1·48	1·30	3·69	1·15 1·46
depth (0–10 m)	2161	–	–	–	–	–	–
depth (> 10–20 m)	431	0·58	1·78	0·45 0·70	0·67	1·95	0·55 0·79
depth (> 20–50 m)	576	0·49	1·63	0·36 0·61	0·54	1·72	0·42 0·66
depth (> 50 m)	652	0·23	1·26	0·06 0·39	0·30	1·35	0·15 0·46
coast	3820	0·00	1·00	−0·01 0·02	0·00	1·00	−0·01 0·01
no_ice	3116	–	–	–	–	–	–
ice	704	0·42	1·52	0·23 0·61	0·82	2·26	0·64 0·99
glac (> 5–10 km) × ice	144	−0·81	0·45	−1·08 −0·53	−0·79	0·45	−1·06 −0·52
glac (> 10–20 km) × ice	183	−0·94	0·39	−1·25 −0·63	−0·83	0·44	−1·14 −0·53
glac (> 20 km) × ice	166	−1·10	0·33	−1·48 −0·71	−1·02	0·36	1·40 −0·65
depth (> 10–20 m) × ice	37	−0·39	0·68	−0·79 0·01	−0·58	0·56	−0·98 −0·18
depth (> 20–50 m) × ice	131	−0·60	0·55	−0·88 −0·32	−0·61	0·55	−0·88 −0·33
depth (> 50 m) × ice	392	−0·31	0·73	−0·65 0·02	−0·29	0·75	−0·61 0·03

Model abbreviations: glac is distance to nearest glacier front; depth is sea bottom depth; coast is distance to the coast (km); no_ice is absence of sea ice; and ice is presence of sea ice.

increases in the risk of leaving were found for increased depths; see Table 2). The presence of sea ice also affected habitat use in this species, with animals showing increased risk of leaving in ice-covered waters (Table 2). However, areas of deeper water, far from glaciers had lower probabilities of leaving when ice-covered compared to ice-free areas (Table 2). The β coefficients for mixed-effects and fixed-effects models (Table 2) show good consistency and the same model structure was selected according to AIC$_c$. The variance component attributed to individual variability (*b*) was 0·087 in this data set, meaning that the individual-specific relative risk of leaving was on average only 34% (1·34 times) higher or lower than the average risk ($e^{\sqrt{0.087}} \approx 1.34$). Animals tagged in the same year might belong to the same social group (see tagging dates in Table S1). In order to verify whether the low inter-individual variability documented was a result of group behaviour, the same CPH model was fitted using year as random effect. The variance component attributed to year was 0·095; year-specific relative risk of leaving was thus 1·36 times higher or lower than the average risk ($e^{\sqrt{0.095}} \approx 1.36$). Variability between years (~social groups) and between individuals was therefore approximately the same.

HABITAT-USE PREDICTIONS

Predictions of habitat-use intensity can be made based on the CPH models. Figure 2 depicts predicted habitat-use intensities
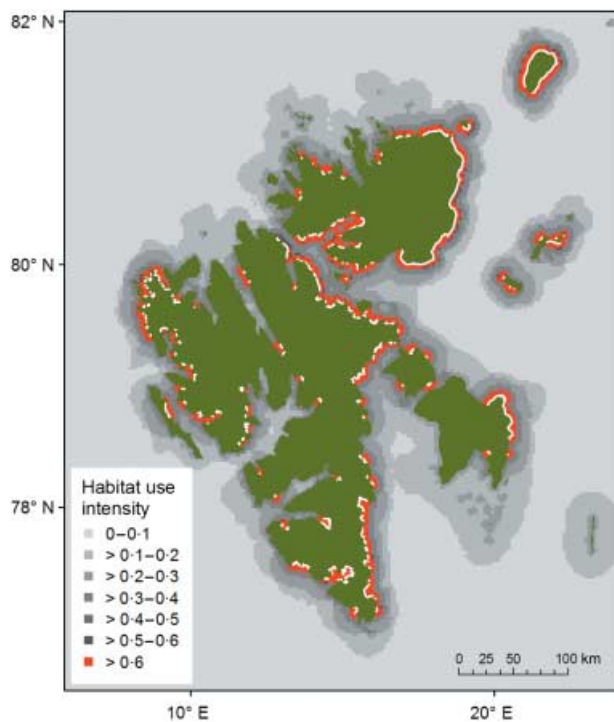
Fig. 2. Predicted habitat-use intensities map for white whales, displaying the estimated probabilities of an animal being in the surrounding 15 km for more than 24 h. Predictions were based on the depth, distance to the coast and distance to the nearest glacier front of the central point of each grid. Grid size is 2·5 km. Glacier fronts are shown in white.

for white whales within coastal waters of Svalbard, based on water depth, distance to the nearest glacier front and distance to the coast. Ice is assumed to be absent. Habitat-use intensities are given as probabilities of being in the surrounding 15 km for more than 24 h. As expected from the CPH model, higher probabilities are observed in shallow areas, close to glacier fronts. Figure 3 presents the complete survival curves (probabilities of being in the same 15 km radius area for more than 1 to 350 h), for white whales in two areas located at distinct distances from glacier fronts.

Discussion

The modelling approach presented in this study quantifies habitat use based on an animal's response to conditions it has actually experienced, thus avoiding the difficulties of defining availability. This analytical framework is especially useful for animals inhabiting heterogeneous, dynamic environments, where conditions available globally are unlikely to be known by animals. In such systems, short-term experience of the actual conditions seems to be the ultimate factor influencing small-scale movement patterns (Fauchald 1999; Fortin 2003; Hays *et al.* 2006). In other, more stable and predictable systems, long-term experience can lead animals to use areas repeatedly where food resources or mates have been abundant before (e.g. Born *et al.* 2005). In such cases, methods that compare used locations with unused, available ones may be more appropriate.

In this study, we used FPTs as a means to assess habitat use, based on the assumption that time spent in an area reflects the presence of resources (e.g. food, resting platforms, cover that provides predator protection, etc.) that makes that particular
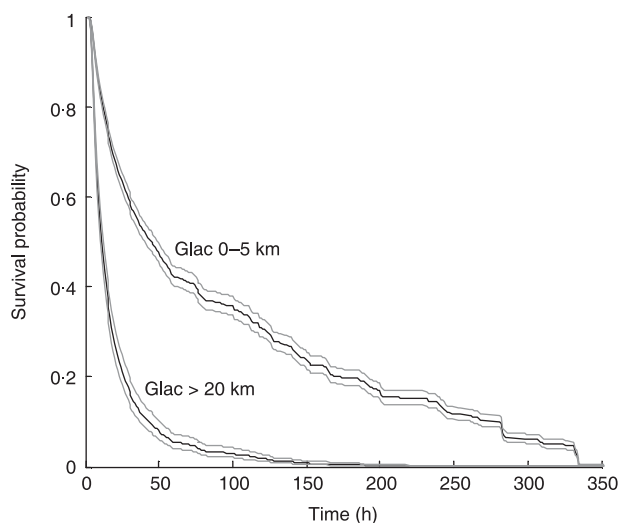
Fig. 3. Survival curves, with 95% confidence intervals, showing the estimated probabilities of white whales remaining in the surrounding 15 km for more than 1 to 350 h. The uppermost curve refers to areas located up to 5 km from the nearest glacier front and the other to areas located at distances greater than 20 km. Distance to the coast, depth and ice are, respectively, 500 m, 10 m and absent in both cases.

area preferred for some unit of time (see Fauchald & Tveraa 2003 and Buskirk & Millspaugh 2006). An interesting, alternative measure of habitat-use intensity is Utilization Distributions (UDs, Marzluff *et al.* 2004; Millspaugh *et al.* 2006). UDs are calculated through kernel or other point-density modelling methods, which use the number of points in an area and a smoothing function to estimate the probability of using each area. The use of kernel methods can be advantageous over FPTs for study areas where habitat conditions are very patchy, because telemetry errors can lead to erroneous classification of the habitat conditions actually used by animals. However, the problem of defining availability is still present when using UDs if all grid points within the home range (study area or other spatial area under consideration) are treated as available (Millspaugh *et al.* 2006). In addition, since FPTs are a direct measurement of habitat-use intensities and not an estimated probability, they are not dependent on the parameters used to produce the UDs, such as the smoothing factor in the kernels. Different smoothing factors used in the kernels can result in very different habitat-use probabilities and home range limits for the same data set (see Worton 1989; Getz & Wilmers 2004; Millspaugh *et al.* 2006). Additionally, kernel methods, at least those based on Euclidean distances, perform poorly in areas where the study species are located along edges (such as coast lines), corridors, or in areas encircling unusable habitats (such as islands for most marine animals and lakes for terrestrial species, see Getz & Wilmers 2004). In such conditions and in cases where long and fast movements are interrupted with Area Restricted Searches (ARS; Benhamou 1992), FPTs are expected to provide a better measure of habitat-use intensities.

FPTs have been used previously to model habitat use, using Compositional Analysis (Pinaud & Weimerskirch 2005) and GLMs (Suryan *et al.* 2006). The first technique is useful when investigating the effect of single, categorical variables on FPTs. The use of GLMs gives the possibility of integrating random effects (see Pinheiro & Bates 2000), similar to the present study. However, the nature of FPT data (and any other time–event variable) is such that they are likely to severely violate the requirements of linearity of the predictors and parametric distribution of the residuals. A possible approach might be to transform FPTs into a binary variable and model this new variable (that now has a binomial distribution) with a binomial GLM with an appropriate link function. In this case, models would compare areas of high FPT (intensive use) and low FPT (transitory areas), in a way that is similar to what is usually done with presence–absence data obtained from surveys or telemetry data (e.g. Cañadas *et al.* 2005; Gillies *et al.* 2006). The main disadvantage of binary classification of a continuous time variable is, of course, loss of information. There is also a potential problem with subjectivity in choice of classification criteria. Additionally, no proper estimates of displacement risks are obtained in these analyses.

The inclusion of a random-effect term in CPH models (Pankratz, de Andrade & Therneau 2005) represents a powerful tool for dealing with data collected from individuals whose intrinsic behaviours differ. By quantifying this individual-level heterogeneity in terms of random effects, the model more appropriately reflects the structure of the data. For instance, the problem of unbalanced sample sizes, because of the typical uneven number of observations obtained per individual in telemetry studies, is overcome. Without taking the inherent individual-level variability into account, animals having more locations could bias the results. However, the development of mixed-effects CPH models is very recent and in many ways, it is still in its infancy. Future developments that would be very welcome include options for including random β coefficients (i.e. random slope models) and temporal covariance terms. A preliminary precautionary measure to deal with serially correlated data is to ensure that the temporal resolution of the data is sufficiently coarse so as to yield independent residuals. Moreover, estimations of survival functions from random-effects CPH models are not yet implemented in available software. However, in the present study, the high degree of consistency of coefficients of models with and without random effects, allowed us to use the fixed-effects model for approximate predictions of population space use.

The modelling of FPTs with CPH models in this study enabled us to measure the responses of ringed seals and white whales to a set of habitat variables. Depending on the species and context, such responses can reflect distinct biological preferences and needs. The strong preference for areas containing glacier fronts by ringed seals and white whales, for example, is likely to be related to the high productivities of these areas (see Lydersen *et al.* 2001) and the concomitant availability of food for these predators in these locations. The temporal changes observed for ringed seal habitat use, with decreased risks of leaving being observed as winter approached, reflect the need for ringed seals to construct and maintain breathing holes as the fast ice forms in the coastal waters of Svalbard. In the case of white whales, sea ice formation along the shore seems to displace them from the coast to areas where there is open water.

In addition to providing inferences about habitat selection based on the estimated parameters, CPH models can also be used for predicting population space use, given the sample of individuals included and the spatio-temporal sampling frame. Such predictions can be particularly useful for applied purposes, for example, as a basis for species conservation plans and reserve selection. Although this novel combination of methods was exemplified here for two marine mammals, it is expected to have wide applicability in studies of habitat use of other highly mobile animals both in marine and terrestrial systems.

Acknowledgements

References

Benhamou, S. (1992) Efficiency of area-concentrated searching behavior in a continuous patchy environment. *Journal of Theoretical Biology*, **159**, 67–81.

Born, E.W., Acquarone, M., Knutsen, L.Ø. & Toudal, L. (2005) Homing behaviour in an Atlantic walrus (*Odobenus rosmarus rosmarus*). *Aquatic Mammals*, **31**, 23–33.

Burnham, K.P. & Anderson, D.R. (2002) *Model Selection and Multimodel Inference: a Practical Information-theoretic Approach*. Springer-Verlag, New York.

Buskirk, S.W. & Millspaugh, J.J. (2006) Metrics for studies of resource selection. *Journal of Wildlife Management*, **70**, 358–366.

Cañadas, A., Sagarminaga, R., De Stephanis, R., Urquiola, E. & Hammond, P.S. (2005) Habitat preference modelling as a conservation tool: proposals for marine protected areas for cetaceans in southern Spanish waters. *Aquatic Conservation: Marine and Freshwater Ecosystems*, **15**, 495–521.

Collett, D. (2003) *Modelling Survival Data in Medical Research*. Chapman & Hall/CRC, Boca Raton, FL, USA.

Cox, D.R. (1972) Regression models and life tables (with discussion). *Journal of the Royal Statistical Society, Series B*, **74**, 187–220.

Fauchald, P. & Tveraa, T. (2003) Using first-passage time in the analysis of area-restricted search and habitat selection. *Ecology*, **84**, 282–288.

Fauchald, P. & Tveraa, T. (2006) Hierarchical patch dynamics and animal movement pattern. *Oecologia*, **149**, 383–395.

Fauchald, P. (1999) Foraging in a hierarchical patch system. *The American Naturalist*, **153**, 603–613.

Fortin, D. (2003) Searching behavior and use of sampling information by free-ranging bison (*Bos bison*). *Behavioral Ecology and Sociobiology*, **54**, 194–203.

Freitas, C., Kovacs, K.M., Ims, R.A., Fedak, M.A. & Lydersen, C. (2008a) Ringed seal post-moulting movement tactics and habitat selection. *Oecologia*, **155**, 193–204.

Freitas, C., Lydersen, C., Fedak, M.A. & Kovacs, K.M. (2008b) A simple new algorithm to filter marine mammal Argos locations. *Marine Mammal Science*, **24**, 315–325.

Getz, W.M. & Wilmers, C.C. (2004) A local nearest-neighbor convex-hull construction of home ranges and utilization distributions. *Ecography*, **27**, 489–505.

Gillies, C.S., Hebblewhite, M., Nielsen, S.E., Krawchuk, M.A., Aldridge, C.L., Frair, J.L., Saher, D.J., Stevens, C.E. & Jerde, C.L. (2006) Application of random effects to the study of resource selection by animals. *Journal of Animal Ecology*, **75**, 887–898.

Guisan, A. & Thuiller, W. (2005) Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, **8**, 993–1009.

Hays, G.C., Hobson, V.J., Metcalfe, J.D., Righton, D. & Sims, D.W. (2006) Flexible foraging movements of leatherback turtles across the North Atlantic Ocean. *Ecology*, **87**, 2647–2656.

Hjermann, D.O. (2000) Analyzing habitat selection in animals without well-defined home ranges. *Ecology*, **81**, 1462–1468.

Johnson, D.H. (1980) The comparison of usage and availability measurements for evaluating resource preference. *Ecology*, **61**, 65–71.

Lydersen, C. & Gjertz, I. (1986) Studies of the ringed seal (*Phoca hispida* Schreber 1775) in its breeding habitat in Kongsfjorden, Svalbard. *Polar Research*, **4**, 57–63.

Lydersen, C. (1998) Status and biology of ringed seals (*Phoca hispida*) in Svalbard. *NAMMCO Scientific Publications*, **1**, 46–62.

Lydersen, C., Martin, A.R., Kovacs, K.M. & Gjertz, I. (2001) Summer and autumn movements of white whales *Delphinapterus leucas* in Svalbard, Norway. *Marine Ecology Progress Series*, **219**, 265–274.

Lydersen, C., Nøst, O.A., Kovacs, K.M. & Fedak, M.A. (2004) Temperature data from Norwegian and Russian waters of the northern Barents Sea collected by free-living ringed seals. *Journal of Marine Systems*, **46**, 99–108.

Lydersen, C., Nøst, O.A., Lovell, P., McConnell, B.J., Gammelsrød, T., Hunter, C., Fedak, M.A. & Kovacs, K.M. (2002) Salinity and temperature structure of a freezing Arctic fjord – monitored by white whales (*Delphinapterus leucas*). *Geophysical Research Letters*, **29**, No. 23, 2119.

Manly, B.F.J., McDonald, L.L., Thomas, D.L., McDonald, T.L. & Erikson, W.P. (2002) *Resource selection by animals: statistical design and analysis for field studies*, Kluwer Academic Publishers, Dordrecht, The Netherlands.

Marzluff, J.M., Millspaugh, J.J., Hurvitz, P. & Handcock, M.S. (2004) Relating resources to a probabilistic measure of space use: forest fragments and Steller's jays. *Ecology*, **85**, 1411–1427.

McLaren, I.A. (1958) The biology of the ringed seal (*Phoca hispida* Schreber) in the eastern Canadian Arctic. *Bulletin of the Fisheries Research Board of Canada*, **118**, 1–97.

Millspaugh, J.J., Nielson, R.M., McDonald, L., Marzluff, J.M., Gitzen, R.A., Rittenhouse, C.D., Hubbard, M.W. & Sheriff, S.L. (2006) Analysis of resource selection using utilization distributions. *Journal of Wildlife Management*, **70**, 384–395.

Murray, D.L. (2006) On improving telemetry-based survival estimation. *Journal of Wildlife Management*, **70**, 1530–1543.

Pankratz, V.S., de Andrade, M. & Therneau, T.M. (2005) Random-effects Cox proportional hazards model: general variance components methods for time-to-event data. *Genetic Epidemiology*, **28**, 97–109.

Pinaud, D. & Weimerskirch, H. (2005) Scale-dependent habitat use in a long-ranging central place predator. *Journal of Animal Ecology*, **74**, 852–863.

Pinheiro, J.C. & Bates, D.M. (2000) *Mixed Effects Models in S and S-PLUS*. Springer-Verlag, New York.

Richard, P.R., Heide-Jørgensen, M.P. & St Aubin, D. (1998) Fall movements of belugas (*Delphinapterus leucas*) with satellite-linked transmitters in Lancaster Sound, Jones Sound, and northern Baffin Bay. *Arctic*, **51**, 5–16.

Richard, P.R., Martin, A.R. & Orr, J.R. (2001) Summer and autumn movements of belugas of the Eastern Beaufort Sea stock. *Arctic*, **54**, 223–236.

Scott, J.M., Heglund, P.J., Haufler, J.B., Morrison, M., Raphael, M.G. & Wall, W.B. (2002) *Predicting Species Occurrences: Issues of Accuracy and Scale*, Island Press, Covelo, CA, USA.

Smith, T.G. & Martin, A.R. (1994) Distribution and movements of belugas, *Delphinapterus leucas*, in the Canadian high Arctic. *Canadian Journal of Fisheries and Aquatic Sciences*, **51**, 1653–1663.

Smith, T.G. (1987) The ringed seal, *Phoca hispida*, of the Canadian western arctic. *Bulletin of the Fisheries Research Board of Canada*, **216**, 1–81.

Strickland, M.D. & McDonald, L.L. (2006) Introduction to the special section on resource selection. *Journal of Wildlife Management*, **70**, 321–323.

Suryan, R.M., Sato, F., Balogh, G.R., Hyrenbach, K.D., Sievert, P.R. & Ozaki, K. (2006) Foraging destinations and marine habitat use of short-tailed albatrosses: A multi-scale approach using first-passage time analysis. *Deep Sea Research Part II: Topical Studies in Oceanography*, **53**, 370–386.

Therneau, T.M. & Grambsch, P.M. (2000) *Modeling Survival Data: Extending the Cox Model*. Springer-Verlag, New York.

Wajnberg, E., Gonsard, P.A., Tabone, E., Curty, C., Lezcano, N. & Colazza, S. (2003) A comparative analysis of patch-leaving decision rules in a parasitoid family. *Journal of Animal Ecology*, **72**, 618–626.

Worton, B.J. (1989) Kernel methods for estimating the utilization distribution in home-range studies. *Ecology*, **70**, 164–168.

Supplementary material

The following supplementary material is available for this article:

Table S1. Summary information for the tracking records for the 18 ringed seals *Phoca hispida* and 12 white whales *Delphinapterus leucas* used in the present study

Table S2. Model ranking for ringed seals

Table S3. Model ranking for white whales

This material is available as part of the online article from: http://www.blackwell-synergy.com/doi/abs/10.1111/j.1365-2664.2008.01505.x
(This link will take you to the article abstract).

Please note: Blackwell Publishing are not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.