

# Habitat use in a marine ecosystem: beluga whales *Delphinapterus leucas* in Cook Inlet, Alaska

Kimberly T. Goetz<sup>1,2,\*</sup>, David J. Rugh<sup>2</sup>, Andrew J. Read<sup>1</sup>, Roderick C. Hobbs<sup>2</sup>

<sup>1</sup>Duke University Marine Laboratory, 135 Pivers Island Road, Beaufort, North Carolina 28516-9721, USA

<sup>2</sup>National Marine Mammal Laboratory, Alaska Fisheries Science Center, NMFS, NOAA, 7600 Sand Point Way NE, Seattle, Washington 98115-6349, USA

**ABSTRACT:** Marine ecosystems have fluid habitat features that can be modeled for application in management decisions. With less than 360 beluga whales *Delphinapterus leucas* remaining in Alaska's Cook Inlet population and increasing development in the area, it is important to identify habitat requirements of this depleted population. We used a classification and regression tree (CART) model and a resource selection function (RSF) model to determine the importance of selected environmental variables in structuring the habitat use of beluga whales in Cook Inlet. Bathymetry, proximity to mudflats, and distance from rivers classified by flow accumulation values were evaluated with respect to the presence or absence of belugas. Although bathymetry, as applied to the model, was not a significant variable, mudflats were a significant predictor of beluga distribution during early summer months. The importance of flow accumulation varied slightly between the 2 models but, in general, belugas preferred higher flow accumulation areas. The CART model correctly classified 88 % of the sightings, with the majority of beluga sightings found to be within 2.7 km of mudflats and 11.5 km of medium flow accumulation inlets. Using a 0.09 probability threshold for habitat, the RSF model correctly discriminated between beluga sightings and non-sightings 92 % of the time. Similar regions of habitat were predicted by CART (habitat area,  $A = 2985 \text{ km}^2$ ) and RSF ( $A = 2373 \text{ km}^2$ ) models. The habitat models developed in this study will help managers meet current conservation goals and make future decisions to promote the recovery of Cook Inlet beluga whales.

**KEY WORDS:** Resource selection function · RSF · Classification and regression tree · CART · Geographic Information Systems · GIS · Receiver operating characteristic curve · ROC

—Resale or republication not permitted without written consent of the publisher—

## INTRODUCTION

Accurately describing and understanding the distribution of organisms is a fundamental problem in ecology with important conservation and management implications. Advances in statistical analysis now make it possible to predict the geographic distribution of species from environmental features (Guisan & Zimmermann 2000, Pearce & Ferrier 2000, Thuiller et al. 2003). In addition, technological innovations in geographic information systems (GIS) provide accurate, high resolution quantifications of many environmental variables. This combination of GIS and modern statisti-

cal techniques allows the creation of spatial habitat models to elucidate which environmental variables correlate with species' occurrence.

Determining habitat use is particularly challenging for marine organisms such as marine mammals due not only to the dynamicity of habitat requirements over space and time, but also the mobile nature of many marine species. In addition, difficulties can arise when applying habitat models to rare or endangered species, especially in cases when there are very few sightings or data lack locational accuracy (Palma et al. 1999, Engler et al. 2004). Despite these challenges, several studies modeling marine mammal habitat have em-

\*Email: kim.goetz@noaa.gov

ployed environmental variables such as sea surface temperature, tidal cycles, proximity to shore, distance from continental shelf, bathymetric slope, and temperature fronts (Davis et al. 1998, Hooker et al. 1999, Gregr & Trites 2001, Waring et al. 2001, Hamazaki 2002, Mendes et al. 2002, Baumgartner et al. 2003, Torres et al. 2003). However, the variables included in any particular model are context-specific; many of these variables would not be appropriate for other marine mammal populations.

The availability of survey data and accessibility of several environmental variables create ideal conditions for identifying habitat preferences for the depleted stock of beluga whales *Delphinapterus leucas* inhabiting the waters of Cook Inlet, Alaska. Understanding environmental variables important to this population of belugas are timely due to impacts associated with the continual development of Anchorage, Alaska's largest city. Identifying habitat features may help provide important management strategies for the future recovery of the stock.

Although the underlying mechanism of beluga distribution is not well understood, several studies report their occupancy in coastal mudflat areas and in proximity to river mouths during the summer (Calkins 1989, Smith & Martin 1994, Moore et al. 2000). Belugas reportedly seek prey aggregations at these locations, so physical and oceanographic features may indirectly affect their assemblages by influencing prey distributions (Huntington 2000, Moore et al. 2000). Moore et al. (2000) reviewed a variety of variables that may influence the distribution of Cook Inlet beluga whales, but, to date, no studies have quantified the influence of physical, ecological, or anthropogenic factors on the distribution of this population.

This study examines the ecological relationships between Cook Inlet belugas and 3 environmental variables. We used data collected during 12 yr of aerial surveys (1993 to 2004) plus classification and regression tree (CART) and resource selection function (RSF) models to predict beluga habitat. Variables used in the models were based on the availability of environmental data and previous research on beluga distribution. Since data on prey availability have not been collected in a manner suitable for beluga research, flow accumulation (a GIS term referring to the quantity of cells upstream of a given cell) was used as a mechanism to distinguish among tributaries entering Cook Inlet. The specific goal of this study was to quantify the relationship between the observed distribution of Cook Inlet beluga whales and (1) bathymetry, (2) mudflats, and (3) flow accumulation values. Results from this analysis will aid stakeholders, especially management agencies, in recognizing specific geographic regions important to belugas.

## MATERIALS AND METHODS

**Beluga whales.** Belugas are circumpolar in the northern hemisphere (Smith et al. 1990). In Alaskan waters, there are 5 known summer populations of beluga whales (Hazard 1988, Frost & Lowry 1990, O'Corry-Crowe et al. 1997). Variation in mitochondrial DNA indicates that, of these 5, the Cook Inlet population is the most genetically distinct (O'Corry-Crowe et al. 1997). In addition, this population is geographically isolated from the other 4 populations by the Alaska Peninsula (Laidre et al. 2000).

The tendency of belugas to consistently use shallow, estuarine locations during summer months has contributed to their overexploitation by hunting. The Cook Inlet beluga population declined from the 1970s to the mid-1990s (Rugh et al. 2000) and decreased approximately 50% from 1994 to 1998 (Hobbs et al. 2000). The annual mean harvest during this period was 21% of the estimated 1998 abundance (Hobbs et al. 2000). As a result, the National Marine Fisheries Service (NMFS) began regulating subsistence hunting of Cook Inlet belugas in 1998 and enacted an official moratorium on 21 May 1999 (Pub. L. No. 106-31, Section 3022, 113 Stat. 57, 100). Exceptions to the moratorium were permitted only through a co-management agreement between NMFS and an Alaskan Native Organization (ANO). With abundance estimates less than 450 (Hobbs et al. 2000), the Cook Inlet stock was listed as 'depleted' under the Marine Mammal Protection Act (MMPA) (65 FR 34590-34597) on 31 May 2000. To date, the population has failed to show a significant increase (Rugh et al. 2004) and is now being examined for evidence that would warrant an endangered listing under the 'Endangered Species Act' (ESA).

**Study area.** Located in south-central Alaska, Cook Inlet is a semi-enclosed tidal estuary stretching approximately 370 km from the southern limits of Kamishak Bay to the northeastern reaches of Turnagain and Knik Arm (Fig. 1). The inlet covers approximately 20 000 km<sup>2</sup> with 1350 km of shoreline (Rugh et al. 2005). Large tidal ranges and broad tidal flats result in currents up to 6.2 m s<sup>-1</sup> and significant changes in shoreline (Moore et al. 2000). At 9 m, tidal height variations at Anchorage are among the most extreme tidal cycles in the world (Mulherin et al. 2001).

The Cook Inlet watershed drains over 101 000 km<sup>2</sup> of land and supports 400 000 residents within its boundaries (Cook Inlet Keeper 2006, see [www.inletkeeper.org/about/watershed.htm](http://www.inletkeeper.org/about/watershed.htm)). Melting snow and ice from the Alaska Range, the Talkeetna Range, the Chugach Mountains, and the Aleutian Range drain into rivers which feed the waters of Cook Inlet (Brabets et al. 1999).

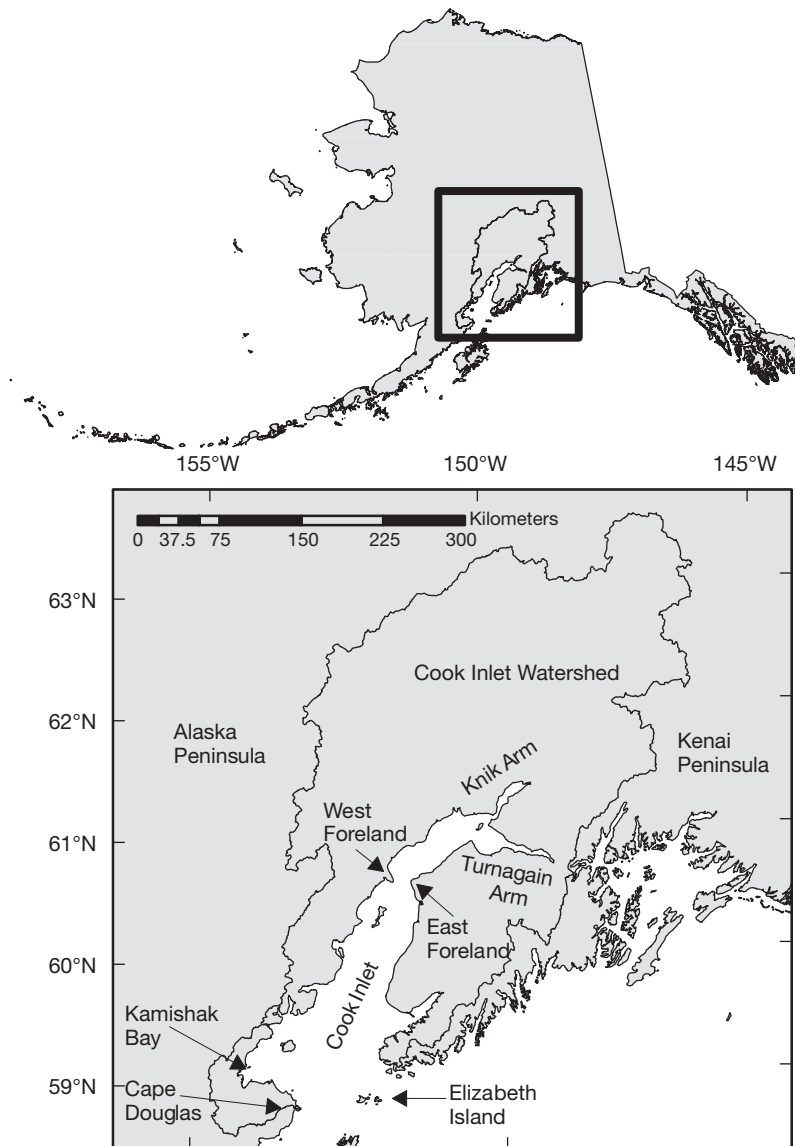


Fig. 1. Cook Inlet study area, south-central Alaska

**Survey protocol.** Using methods described by Rugh et al. (2004), we conducted aerial surveys each summer from 1993 to 2004. Survey duration ranged between 4 and 9 d with as many as 16 total flights per year. All surveys were conducted between 2 June and 29 July. The aircraft used for most surveys was a 5-passenger Aero Commander 680 FL with twin engines, high wings, and 10 h flying capability. Two observers were on the left and 1 on the right, each searching through large bubble windows. An intercom system provided communication between the observers, the data recorder, and the pilot (Rugh et al. 2004). The laptop computer, interfaced with a data entry system and a global positioning system (GPS), recorded the aircraft's location as well as sighting information.

The survey protocol involved both coastal surveys and off-shore transects. All surveys were conducted approximately 244 m above sea level with a sighting distance of 1.4 km on either side of the aircraft, although good viewing conditions permitted searches well beyond that distance. Documented whale locations were determined by the aircraft flying directly over whale groups. Although some whale groups were possibly missed during these surveys, consistency of inter-year distributional information minimizes concern about overlooked groups.

**GIS analysis. Tracklines and sightings data:** Time segments not solely dedicated to the searching of beluga groups or coincident with 'poor' or 'useless' visibility were considered 'off-effort' and excluded from the analyses. Using the animal movement extension in Hawth's analysis tools (Beyer 2004<sup>1</sup>), we created individual tracklines from the GPS point locations recorded during the aerial surveys. All tracklines were buffered 1.4 km on each side (considered to be the maximum sighting distance) and then merged to create a single polygon layer. Over the 12 yr of aerial surveys, search effort covered approximately 75% of the Cook Inlet study area. All years were considered equal (i.e. no trend was tested), and although coastal areas of the northern part of Cook Inlet were surveyed 3 to 6 times each year and some areas were surveyed only once a year, both survey effort and sightings were cumulative, countering the effect of differential sampling effort.

Off-effort sighting data and observations made outside the study area boundaries were not included in analyses. Each of the 226 sightings considered in this study represented the 'best' estimated location for each beluga group. An additional 226 random positions were generated within the buffered, on-effort polygon layer using Hawth's random point generator tool (Beyer 2004<sup>1</sup>). The 452 points (whale sightings plus random point locations) were used in statistical analyses to distinguish potential beluga habitat from non-habitat. Although there is a reasonable probability that aerial observers missed some beluga groups (as estimated in Hobbs et al. 2000), the consistency of sighting locations across the 12 yr of our analysis indicates that the sighting data are

<sup>1</sup> Available at [www.spatial ecology.com/htools](http://www.spatial ecology.com/htools)

fairly representative of the true distribution. Therefore, concerns of false negatives are minimal relative to the probability of a random point landing on a location where belugas could have been but were not recorded.

**Bathymetry:** Because a GIS bathymetry coverage was not available for the Cook Inlet area, we combined soundings data from 2 sources: (1) NOAA Electronic Navigation Charts (NOAA 2004, see [http://nauticalcharts.noaa.gov/csdl/ctp/encdirect\\_new.htm](http://nauticalcharts.noaa.gov/csdl/ctp/encdirect_new.htm)) and (2) the National Ocean Service Hydrographic Data Base (USGS 2005, see <http://nhd.usgs.gov/data.html>). Tidal data for all soundings were referenced in mean lower low water (MLLW). We used the 'inverse distance weighted' (IDW) method to interpolate bathymetry into a 100 × 100 m cell resolution raster. IDW interpolation determines cell values using a linearly weighted combination of a set of sample points as a function of inverse distance. To compensate for values greater than 0.0 (above water during MLLW), we added 9.0 m to the entire grid to achieve depths at maximum tidal height (4.2 to 223.8 m) Because of the extreme tidal range, specific depths at the time and location of a sighting were not taken into consideration.

**Mudflats:** A polygon shapefile depicting mudflats was used to determine whether belugas were selecting sites based on this substrate as suggested by previous studies (Rugh et al. 2000). We used the 'Spatial Analyst' extension in Arc/Info (Version 8.3) to create a Euclidean distance from mudflat grid with 100 × 100 m cell size. Distance values ranged from 0.00 to 40.49 km.

**Inlet locations and flow accumulation.** We obtained hydrographic Arc/Info coverages for the Cook Inlet watershed from 2 sources: (1) the United States Geological Survey (USGS) seamless data distribution system (USGS 2004<sup>2</sup>) and (2) the Alaskan Department of Fish and Game Fish Distribution Database (ADFG 2004, see [www.sf.adfg.state.ak.us/SARR/FishDistrib/FDD\\_gisdata.cfm](http://www.sf.adfg.state.ak.us/SARR/FishDistrib/FDD_gisdata.cfm)). Next, we used a 2-arcsecond (1:63 360) digital elevation model (DEM) (USGS 2004<sup>2</sup>) for Alaska, in combination with the spatial hydrographic data, to create a point coverage depicting the locations where each river and stream entered Cook Inlet.

Accumulated flow to each cell was then calculated by summing the weight for all cells that flow into each down-slope cell. To compensate for spatial inaccuracies between the flow accumulation grid and hydrographic coverages, we buffered all inlet points 150 m and created unique identities for each inlet buffer. We then assigned the highest flow accumulation value within the 150 m buffer to each inlet. Based on the spread of flow accumulation values, inlet points were

classified as high, medium, or low flow inlets. Inlets with flow accumulation values less than the first quartile were categorized as low flow accumulation inlets; values between the first and third quartiles were classified as medium; and those inlets with values greater than the third quartile were classified as high. Using the 'spatial analyst' extension, we created three 100 × 100 m cell resolution Euclidean distance grids: distance from low (0 to 82.26 km), medium (0 to 83.48 km), and high (0 to 110.31 km) flow accumulation inlets.

**Spatial sampling.** For each beluga sighting and random point location, each environmental GIS layer was sampled to determine the 5 corresponding habitat variables (BATHDEPTH, DISTMUD, DISTLOW, DISTMED, DISTHIGH; Table 1). The resulting data set was examined for co-variation using a correlation matrix. While some correlation was evident between variables, all correlations were below 0.75. For the purposes of building a habitat model we chose to retain all the variables.

**Modeling beluga habitat. Classification and regression tree:** We used a CART model to reveal structure within the data and to identify relationships between beluga presence/absence (the response variable) and 3 explanatory environmental variables. CART models are both flexible and robust; they can interpret missing values and capture interactions present among the predictor variables (De'Ath & Fabricius 2000). This method produces a binary tree from the recursive partitioning of data into increasingly homogenous subgroups (Redfern et al. 2006). Unlike linear models, CART models do not assume an *a priori* relationship between the response and predictor variables; instead each predictor variable is considered individually among subgroups. CART uses an algorithm based on a single best predictor variable to resolve relationships within a complex data set (Torres et al. 2003). Each split in the tree is based on the maximum deviance in the response variable (Redfern et al. 2006). The tree ends with a set of terminal nodes representing final classifications. Without a rule to deter-

Table 1. Variables used to model beluga habitat

Variable	Unit	Abbreviation
Bathymetric depth	m	BATHDEPTH
Distance from mudflats	km	DISTMUD
Distance from low flow accumulation river inlets	km	DISTLOW
Distance from medium flow accumulation river inlets	km	DISTMED
Distance from high flow accumulation river inlets	km	DISTHIGH

<sup>2</sup> Available at <http://seamless.usgs.gov/website/Seamless>

mine when to stop the binary partitioning, splitting will continue until terminal nodes contain only 1 data point.

In order to avoid over-fitting the data, we used a 10-fold cross validation method to determine an appropriate stopping point—the highest prediction accuracy for a data set. This technique randomly splits the data into 10 subsets. A model is built from 9 of the subsets and the 10th subset is used to evaluate the model. This process is repeated 10 times, holding one of the subsets aside for evaluation (Chambers & Hastie 1992). The cross validation averages the deviance from the 10 functions. Deviance per subtree was plotted to visually determine the appropriate cut-off point for the original tree. Habitat area resulting from classification rules was mapped using Arc/Info (Version 9.1).

**Resource selection function:** In addition to examining the structure within the data, we also employed an RSF model to assess the significance of each environmental variable. An RSF model is any model that yields the estimated probability of habitat use of an organism (Manly et al. 1993). The logistic equation was used to calculate the coefficients for the RSF. The form of the RSF is:

$$w(x) = \exp(\beta_1 x_1 + \dots + \beta_p x_p)$$

where  $\beta_1 \dots \beta_p$  are the regression coefficients,  $x_1 \dots x_p$  are the independent variables, and ( $w$ ) is the estimated probability of a beluga sighting. This type of analysis describes the presence or absence of belugas as a function of several environmental or explanatory variables. To determine the significance of several GIS-derived habitat variables, we used automated stepwise model fitting with both backward and forward selection. Models were evaluated with the Akaike information criterion (AIC), which seeks to find the best compromise between maximizing the likelihood and minimizing the number of model parameters. The stepwise procedure helps reduce problems with multicollinearity by removing variables that have high redundancy with other variables already in the model (Thuiller et al. 2003).

The 'receiver operating characteristic' (ROC) curve was used to measure the diagnostic accuracy of the model. ROC analysis originated from signal detection theory as a model of how well a receiver is able to detect a signal in the presence of noise. The ROC curve is a graphical method that represents the relationship between the false positive fraction (1-specificity) and the true positive fraction (sensitivity) for a range of thresholds. The dependent variable is either presence or absence, and the model predicts the threshold of a given variable at which a species is present (Legendre & Legendre 1998). The optimal threshold, ranging from 0 to 1, optimizes errors of omission versus errors of commission. Good model performance is character-

ized by a curve that maximizes sensitivity for low values of 1-specificity. Heavily over-sampled random points versus actual whale sightings will result in a lower optimal threshold without changing the overall outcome of predicted beluga habitat area.

The area between the 45° line and the ROC curve (AUC) measures the ability of the model to correctly classify a species as present or absent (Thuiller et al. 2003). The AUC value ranges from a value of 0.5, indicating no discrimination ability, to 1.0 for models with perfect discrimination capability (Pearce & Ferrier 2000). Models with AUC values greater than 0.9 are considered very good because the true positive rate is high relative to the false positive rate (Swets 1988).

Results from the final RSF model were mapped into geographic space using ArcGIS (Version 9.1). All explanatory variables in the model were represented by grids and were multiplied with the appropriate calculated parameter coefficient to form a visual representation of habitat suitability. Based on the calculated threshold value from the ROC curve analysis, we classified habitat suitability as either beluga habitat or non-habitat areas.

## RESULTS

### Classification and regression tree model

After examining the cross-validated plot, we pruned the tree to 3 nodes to avoid over-fitting the data. The CART model correctly classified 88 % of the samples as beluga sightings or random points (misclassification error = 0.12) using only DISTMUD and DISTMED (see Table 1).

Terminal Node 1, farther than 2.7 km from mudflats, was the only node to contain unmixed results. This node classified 162 non-beluga sites as random points. The remaining 2 terminal nodes contained mixed data (Table 2). Terminal Node 2, within 2.7 km of mudflats and less than 11.5 km from medium flow accumulation

Table 2. Classification of beluga sightings and random locations ( $n = 452$ ) within the 3 terminal nodes created from CART analysis. Terminal Node 1: farther than 2.72 km from mudflats; Terminal Node 2: within 2.72 km of mudflats and <11.50 km from medium flow accumulation inlets; Terminal Node 3: within 2.72 km of mudflats and >11.50 km from medium flow accumulation inlets

Terminal node	Beluga sightings (n)	Random locations (n)	Total (n)
1	0	162	162
2	207	37	244
3	19	27	46

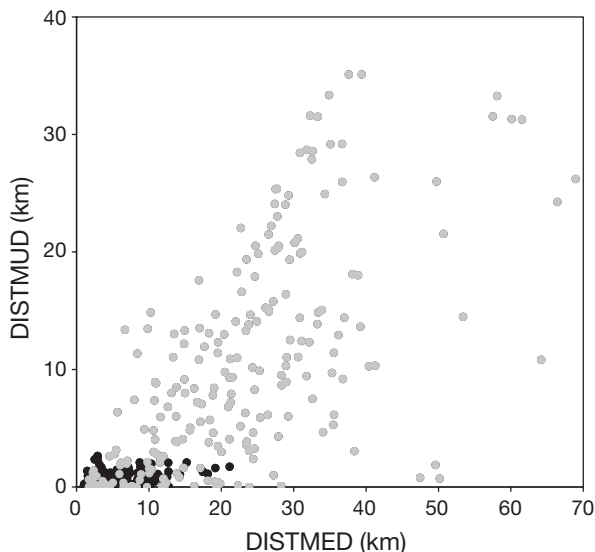


Fig. 2. CART-based analysis showing distribution of beluga sightings (●) and random, non-sighting locations (●) relative to distance from mudflats (DISTMUD) and distance from medium flow accumulation river inlets (DISTMED)

inlets, contained mixed results; of the 243 beluga sightings considered, 37 were misclassified as random points (misclassification error = 0.15). Within 2.7 km of mudflats and farther than 11.5 km from medium flow accumulation inlets, Terminal Node 3 also contained mixed results. This node had the highest level of uncertainty, with a misclassification error of 0.41.

Fig. 2 displays the distribution of points in relation to the CART partitioning of the data. Belugas were most commonly found within 2.7 km of mudflat areas and 11.5 km of medium flow accumulation inlets. Once mapped into GIS, the CART translated into approximately 2985 km<sup>2</sup> of habitat area (Fig. 3).

**Resource selection model**

We examined the full model including all measured environmental variables. All interaction terms were also examined, but they were insignificant and, due to the addition of unnecessary model complexity, excluded from the remainder of the analysis. Of the 5 environmental variables considered, neither BATHDEPTH (*p* = 0.26) nor DISTLOW (*p* = 0.93) contributed significantly to

the fitted model. AIC values from automated stepwise model fitting were ranked for 5 models including the full model (Table 3). Although the model containing DISTMUD, DISTMED, and DISTHIGH had the lowest AIC value, the second and third models were within 2 AIC units of the 'best' model, suggesting that further investigation is necessary to determine which model should be used for the remaining analysis. However, examination of the 2 lower ranked models resulted in less than a 1% difference in habitat. Therefore, we chose the most parsimonious model with parameters DISTMUD, DISTMED, and DISTHIGH as the final model:

$$w = -(7.99 \times 10^{-4} \text{ DISTMUD}) - (1.14 \times 10^{-4} \text{ DISTMED}) - (3.78 \times 10^{-5} \text{ DISTHIGH})$$

where *w* is the estimated probability of beluga detection within a 100 × 100 m cell. Model coefficients, standard errors, and significance values are reported in Table 4. The median odds ratios (Table 4) indicate that every 100 m of additional distance from mudflats reduces the odds of finding a beluga by 8%, every 500 m of additional distance from medium flow accumulation inlets reduces the odds by 6%, and every 1000 m of additional distance from high flow accumulation inlets reduces the odds by 4%. In other words,

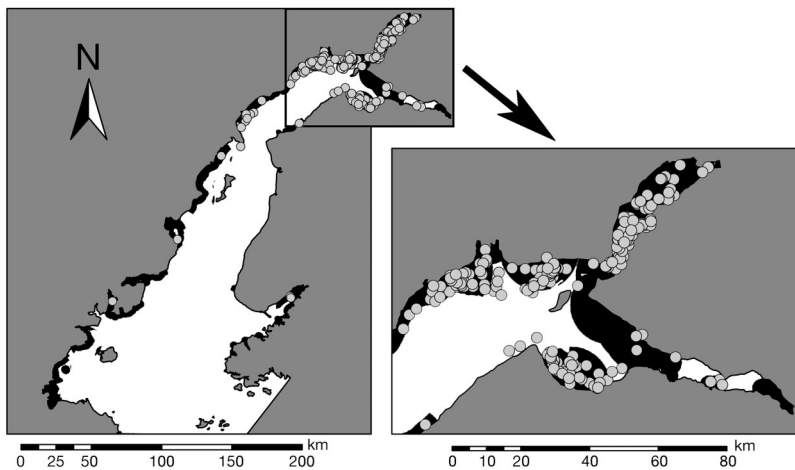


Fig. 3. Habitat (black) predicted by CART model with beluga sightings shown (●)

Table 3. Resource selection function models supported by AIC (*n* = 452) with *k* parameters

Model	<i>k</i>	AIC	ΔAIC
DISTMUD + DISTMED + DISTHIGH	3	277.72	0.00
BATHDEPTH + DISTMUD + DISTMED + DISTHIGH	4	278.42	0.70
DISTMUD + DISTLOW + DISTMED + DISTHIGH	4	279.68	1.96
BATHDEPTH + DISTMUD + DISTLOW + DISTMED + DISTHIGH	5	280.41	2.69
DISTMUD + DISTMED	2	281.52	3.80

Table 4. Resource selection function for final model

Variable	Coefficient	Odds ratio		p	95% Confidence interval	
		Median	SE		Upper	Lower
DISTMUD	$-7.991 \times 10^{-4}$	0.92 (100 m)	$1.550 \times 10^{-4}$	<0.001	$-4.944 \times 10^{-4}$	$-1.104 \times 10^{-3}$
DISTMED	$-1.144 \times 10^{-4}$	0.94 (500 m)	$3.248 \times 10^{-5}$	<0.001	$-5.058 \times 10^{-5}$	$-1.782 \times 10^{-4}$
DISTHIGH	$-3.785 \times 10^{-5}$	0.96 (1000 m)	$1.561 \times 10^{-5}$	0.016	$-7.169 \times 10^{-6}$	$-6.853 \times 10^{-5}$

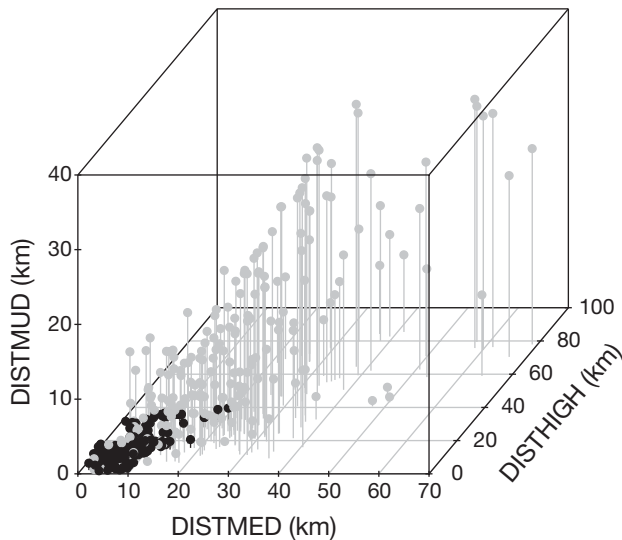


Fig. 4. RSF-based analysis showing beluga sightings (●) and random, non-sighting locations (●) relative to distance from mudflats (DISTMUD) and distance from medium (DISTMED) and high (DISTHIGH) flow accumulation river inlets

the odds of finding a beluga is significantly greater closer to mudflats, and medium and high flow accumulation inlets (Fig. 4).

The ROC curve resulted in a 0.09 probability threshold for the final model in which threshold values  $\geq 0.09$  were considered habitat and values  $< 0.09$  non-habitat. Habitat area was mapped and calculated in Arc/Info (Version 9.1) using the final model. For the Cook Inlet study area, this procedure estimates 2373 km<sup>2</sup> of beluga habitat (Fig. 5).

A 0.92 AUC value indicated that the model correctly discriminated between beluga presence and absence based on the 3 selected environmental variables 92% of the time. Therefore, if a pair of evaluation sites (beluga present vs. absent) is chosen at random, there is a 0.92 probability that the model will predict a higher likelihood of occurrence for beluga presence than absence.

## DISCUSSION

The CART and the RSF models predict similar size and location of beluga habitat in Cook Inlet (2985 and 2373 km<sup>2</sup>, respectively). Both models suggest that mudflats and flow accumulation are important environmental features for the summer distribution of this population. Belugas tend to prefer medium and high flow accumulation inlets, indicating a preference for larger river basins. However, without additional data, we cannot determine whether these preferences are being driven directly by the location of mudflats and higher flow accumulation inlets or indirectly as a result of prey availability and distribution during the summer months.

In contrast, low flow accumulation inlets and bathymetry were not significant predictors of beluga habitat in either of the 2 models. Because of the dynamic tidal cycle in Cook Inlet, water depths at the time of a sighting were not available and only depths relative to MLLW (+9 m) were included in the analysis. Therefore, even if belugas moved with the tides and maintained a fairly constant depth, bathymetric projections without compensating for tidal influence would show whales at varying depths. Alternatively, bathymetry may be more important to Cook Inlet belugas during the winter months when there are no Pacific salmon

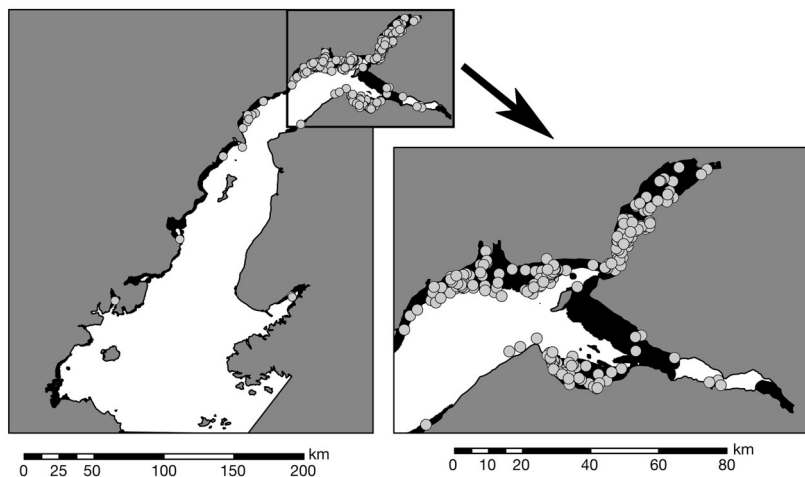


Fig. 5. Habitat (black) predicted by RSF model with beluga sightings shown (●)

runs attracting belugas into shallow waters and sea ice prevents access to tidal flats.

Although sighting data in this study were primarily collected in June, other aerial surveys (Rugh et al. 2000, 2004), opportunistic observations (NMFS unpubl. data), and tagged whales (Hobbs et al. 2005) have shown that the distribution documented in early June is fairly representative of the distribution throughout the summer. However, in winter, the distribution changes (Rugh et al. 2004) primarily because sea ice—which often scrapes the ground while moving across tidal flats—makes inhabiting shallow waters too hazardous for marine mammals.

The summer distribution of belugas indicated in this study agrees with previous qualitative studies that report belugas near coastal mudflats and river mouths (Calkins 1989, Smith & Martin 1994, Moore et al. 2000, Rugh et al. 2000). Although this study is the first to apply quantitative measures that correlate beluga distribution to habitat features, causative factors are not yet adequately documented. For example, other studies suggest that belugas may be attracted to near-shore environments for reasons such as prey availability (Calkins 1989, Huntington 2000, Moore et al. 2000), breeding (Calkins 1989), calving (Sergeant & Brodie 1975, Calkins 1989), molting (Calkins 1989, St. Aubin et al. 1990, Smith et al. 1992), and shelter from predators (Shelden et al. 2003). When detailed information on these biological variables is available, future habitat analyses can connect beluga behavior more directly to the factors that actually drive their distribution.

Even though the habitat predicted by CART and RSF modeling includes coastal areas extending the entire length of Cook Inlet (Figs. 3 & 5), only 3% of the beluga sightings were recorded south of the east and west Forelands (Fig. 1). Historically, belugas inhabited both the northern and southern reaches of the inlet (Rugh et al. 2000). However, since NMFS began conducting summer aerial surveys in 1993, belugas have been primarily sighted along the northern perimeter (Rugh et al. 2000, 2005). A significantly reduced population in combination with beluga preference for estuarine waters may explain the current distribution of whales in the core areas of their range. As the population recovers, their distribution will likely expand into the unused portions of the predicted habitat.

While this study examined several environmental variables, it would be useful to incorporate other physical parameters such as sea surface temperature, turbidity, tidal cycles, and salinity into our habitat models as each of these parameters could play a role in explaining the distribution of more direct parameters such as prey availability. Unfortunately, these were not available for the Cook Inlet area. Prey distribution information is essential to determine whether beluga

preference for medium and high flow accumulation inlets is indeed a function of food availability. The occurrence of belugas near river mouths may reflect a feeding strategy whereby belugas take advantage of highly concentrated fish runs in shallow channels where they are easy to catch. Salmon and eulachon are common in stomach contents of beluga whales sampled (through subsistence hunts or whales found dead) during periods when the respective fish runs are underway in the area (Calkins 1989). The fauna in upper Cook Inlet include eulachon *Thaleichthys pacificus*, plus immigrating smolt and emigrating adults of 5 Pacific salmon species: chinook *Oncorhynchus tshawytscha*, pink *O. gorbuscha*, coho *O. kisutch*, sockeye *O. nerka*, and chum *O. keta* (Moulton 1997, Moore et al. 2000). During the months of June and July, when the surveys were conducted, many of these anadromous fish species were present and often at their peak availability (Moore et al. 2000).

Incorporating measures of prey distribution in general, and anadromous fish runs in particular, could offer greater predictive accuracy about the distribution of belugas. However, descriptions of fish abundance and distribution for the Cook Inlet area are generally lacking. Fish run data are currently biased toward commercially valuable fish stocks such as sockeye salmon, while escapement data for other fish species are recorded only on an opportunistic basis (D. Westerman pers. comm.). This trend of monitoring economically important fish species as well as the assorted, inconsistent data collection methods used to monitor fish runs— weir counts, sonar counts, ground surveys, and aerial surveys— have led to unreliable data within and across years. To date, data on prey distribution have not been collected for the purpose of beluga or other marine mammal research.

Finally, the predictive power of these habitat models could be further increased by accounting for anthropogenic factors such as fishing, illegal harvest, oil and gas activity, military action, and transportation. Using variables such as these can be difficult because of the amount of information needed to accurately model the effect of such anthropogenic issues. For example, in order to incorporate oil and gas activity in habitat models, researchers must account for the location and activities of individual oil rigs as well as the use of transportation to and from the rigs. Data on anthropogenic issues potentially influencing beluga distribution in Cook Inlet were not readily available and, consequently, were not included in this study.

The results from this study and others suggest that it is possible to use statistical modeling to predict species habitat based on environmental variables. Nevertheless, as with most modeling approaches, there are constraints. To accurately predict habitat, predictor vari-



ables must be available for surveyed and unsurveyed regions. As in the case of the present study, lack of prey distribution data may hinder the power of habitat modeling. Regardless, this study provides an example of how the interdisciplinary integration of GIS, biology, physical earth sciences, and statistical modeling can reveal important aspects of the habitat requirements of a cetacean species. In addition, the results indicate that sighting data can provide valuable information about beluga habitat use. Defining the extent of a species' range over time and space is essential for developing conservation strategies, and we are confident that the methods used to describe the spatiotemporal distribution of Cook Inlet belugas are applicable to other species with different environmental requirements. Such methods are particularly useful because of their potential to document aspects of habitat that may be subject to adverse impacts, a characteristic of important concern in any wildlife management plan.

*Acknowledgements.* We thank the National Marine Mammal Laboratory for providing support for this study and send our sincerest appreciation to the biologists and flight crew who flew many long days during the 1993 to 2004 aerial surveys (carried out under Scientific Research Permit No. 782-1438). Dr. N. Friday, J. Maresh, A. Friedlaender, D. Vos, and J. Carris reviewed and commented on earlier drafts. Dr. P. Halpin, B. Best, E. Trembl, and J. Benson offered their technical assistance with data interpretation. Dr. D. Urban, Dr. J. Laake, Dr. J. Ver Hoef, and K. Miclaus were all paramount in determining the appropriate use of statistical methods. B. Shavelson from Cook Inlet Keeper generously provided spatial data and communicated current legislative developments. K. Shelden and Dr. K. Laidre supplied constant support and enthusiasm along the way.

#### LITERATURE CITED

- Baumgartner MF, Cole TVN, Clapham PJ, Mate BR (2003) North Atlantic right whale habitat in the lower Bay of Fundy and on the SW Scotian Shelf during 1999–2001. *Mar Ecol Prog Ser* 264:137–154
- Brabets TP, Nelson GL, Dorava JM, Milner AM (1999) Water-quality assessment of the Cook Inlet basin, Alaska—environmental setting. Water-resources investigations report. 99–4025. US Geological Survey, Branch of Information Services, Denver, CO
- Calkins DG (1989) Status of belukha whales in Cook Inlet. In: Jarvela LE, Thorsteinson LK (eds) *Gulf of Alaska, Cook Inlet, and North Aleutian Basin information update meeting*. Anchorage, AK, Feb. 7–8, 1989, USDOC, NOAA, OCSEAP, Anchorage, AK, p 109–112
- Chambers JM, Hastie TJ (1992) *Statistical models*. S. Wadsworth & Brooks/Cole Advances Books & Software, Pacific Grove, CA
- Davis RW, Fargion GS, May N, Leming TD, Baumgartner M, Evans WE, Hansen LJ, Mullin K (1998) Physical habitat of cetaceans along the continental slope in the north-central and western Gulf of Mexico. *Mar Mamm Sci* 14: 490–507
- De'Ath G, Fabricius KE (2000) Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology* 81:3178–3192
- Engler R, Guisan A, Rechsteiner L (2004) An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. *J Appl Ecol* 41:263–274
- Frost KJ, Lowry LF (1990) Distribution, abundance and movements of beluga whales, *Delphinapterus leucas*, in coastal waters of western Alaska. *Can Bull Fish Aquat Sci* 224: 39–57
- Gregg EJ, Trites AW (2001) Predictions of critical habitat for five whale species in the waters of coastal British Columbia. *Can Bull Fish Aquat Sci* 58:1265–1285
- Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology. *Ecol Model* 135:147–186
- Hamazaki T (2002) Spatiotemporal prediction models of cetacean habitats in the mid-western north Atlantic Ocean (from Cape Hatteras, North Carolina, USA to Nova Scotia, Canada). *Mar Mamm Sci* 18:920–939
- Hazard K (1988) Beluga whale, *Delphinapterus leucas*. In: Leutfer JW (ed) *Selected marine mammals of Alaska: species accounts with research and management recommendations*. Marine Mammal Commission, Washington, DC, p 195–235
- Hobbs RC, Rugh DJ, DeMaster, DP (2000) Abundance of belugas, *Delphinapterus leucas*, in Cook Inlet, Alaska, 1994–2000. *Mar Fish Rev* 62(3):37–45
- Hobbs RC, Laidre KL, Vos DJ, Mahoney BA, Eagleton M (2005) Movements and area use of belugas, *Delphinapterus leucas*, in a subarctic Alaskan estuary. *Arctic* 58: 331–340
- Hooker SK, Whitehead H, Gowans S (1999) Marine protected area design and the spatial and temporal distribution of cetaceans in a submarine canyon. *Conserv Biol* 13: 592–602
- Huntington HP (2000) Traditional knowledge of the ecology of belugas, *Delphinapterus leucas*, in Cook Inlet, Alaska. *Mar Fish Rev* 62(3):134–140
- Laidre KL, Shelden KEW, Rugh DJ, Mahoney BA (2000) Beluga, *Delphinapterus leucas*, distribution and survey effort in the Gulf of Alaska. *Mar Fish Rev* 62(3):27–36
- Legendre P, Legendre L (1998) *Numerical ecology*, 2nd English edn. Elsevier, Amsterdam
- Manly BFJ, McDonald LL, Thomas DL (1993) *Resource selection by animals: statistical design and analysis for field studies*. Chapman & Hall, London
- Mendes S, Turrell W, Lütkebohle T, Thompson P (2002) Influence of the tidal cycle and a tidal intrusion front on the spatio-temporal distribution of coastal bottlenose dolphins. *Mar Ecol Prog Ser* 239:221–229
- Moore SE, Shelden KEW, Litzky LK, Mahoney BA, Rugh DJ (2000) Beluga, *Delphinapterus leucas*, habitat associations in Cook Inlet, Alaska. *Mar Fish Rev* 62(3):60–80
- Moulton LL (1997) Early marine residence, growth, and feeding by juvenile salmon in northern Cook Inlet, Alaska. *Alsk Fish Res Bull* 4:154–177
- Mulherin ND, Tucker WB III, Smith OP, Lee WJ (2001) *Marine ice atlas for Cook Inlet, Alaska*. ERDC/CRREL TR-01-10, Cold Regions Research and Engineering Laboratory, Hanover, NH, for National Oceanic and Atmospheric Administration, Office of Response and Restoration, Seattle, WA
- O'Corry-Crowe GM, Suydam RS, Rosenberg A, Frost KJ, Dizon AE (1997) Phylogeography, population structure and dispersal patterns of the beluga whale *Delphinapterus leucas* in the western Nearctic revealed by mitochondrial DNA. *Mol Ecol* 6:955–970
- Palma L, Beja P, Rodrigues M (1999) The use of sighting data

- to analyse Iberian lynx habitat and distribution. *J Appl Ecol* 36:812–824
- Pearce J, Ferrier S (2000) An evaluation of alternative algorithms for fitting species distribution models using logistic regression. *Ecol Model* 128:127–147
- Redfern JV, Ferguson MC, Becker EA, Hyrenbach KD and 15 others (2006) Techniques for cetacean habitat modeling. *Mar Ecol Prog Ser* 310:271–295
- Rugh DJ, Shelden KEW, Mahoney BA (2000) Distribution of belugas, *Delphinapterus leucas*, in Cook Inlet, Alaska, during June/July 1993–2000. *Mar Fish Rev* 62(3):6–21
- Rugh DJ, Mahoney BA, Smith BK (2004) Aerial surveys of beluga whales in Cook Inlet, Alaska, between June 2001 and June 2002. NOAA Tech Memo NMFS AFSC 145
- Rugh DJ, Shelden KEW, Sims CL, Mahoney BA, Smith BK, Litzky LK, Hobbs RC (2005) Aerial surveys of belugas in Cook Inlet, Alaska, June 2001, 2002, 2003, and 2004. NOAA Tech Memo NMFS AFSC 149
- Sergeant DE, Brodie PF (1975) Identity, abundance, and present status of populations of white whales, *Delphinapterus leucas*, in North America. *J Fish Res Board Can* 32: 1047–1054
- Shelden KEW, Rugh DJ, Mahoney BA, Dahlheim ME (2003) Killer whale predation on belugas in Cook Inlet, Alaska: implications for a depleted population. *Mar Mamm Sci* 19: 529–544
- Smith TG, Martin AR (1994) Distribution and movements of belugas, *Delphinapterus leucas*, in the Canadian high arctic. *Can J Fish Aquat Sci* 51:1953–1663
- Smith TG, St. Aubin DJ, Geraci JR (1990) Research on beluga whales, *Delphinapterus leucas*: introduction and overview. *Can Bull Fish Aquat Sci* 224:1–6
- Smith TG, St. Aubin DJ, Hammill MO (1992) Rubbing behaviour of belugas, *Delphinapterus leucas*, in a high arctic estuary. *Can J Zool* 70:2405–2409
- St. Aubin DJ, Smith TG, Geraci JR (1990) Seasonal epidermal molt in beluga whales (*Delphinapterus leucas*). *Can J Zool* 68:359–367
- Swets JA (1988) Measuring the accuracy of diagnostic systems. *Science* 240:1285–1293
- Thuiller W, Araujo MB, Lavorel S (2003) Generalized models vs. classification tree analysis: predicting spatial distributions of plant species at different scales. *J Veg Sci* 14:669–680
- Torres LG, Rosel PE, D'Agrosa C (2003) Improving management of overlapping bottlenose dolphin ecotypes through spatial analysis and genetics. *Mar Mamm Sci* 19:502–514
- Waring GT, Hamazaki T, Sheehan D, Wood G, Baker S (2001) Characterization of beaked whale (*Ziphiidae*) and sperm whale (*Physeter macrocephalus*) summer habitat in shelf-edge and deeper waters off the Northeast US. *Mar Mamm Sci* 17:703–717

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

Submitted: February 28, 2006; Accepted: May 27, 2006  
Proofs received from author(s): January 2, 2007