Vol. 329: 289–299, 2007

MARINE ECOLOGY PROGRESS SERIES
Mar Ecol Prog Ser

Published January 11

Comparing cetacean abundance estimates derived from spatial models and design-based line transect methods

Amaia Gómez de Segura[1,*], Philip S. Hammond[2], Ana Cañadas[3], Juan A. Raga[1]

[1]Zoología Marina, Instituto Cavanilles de Biodiversidad y Biología Evolutiva, Universidad de Valencia, PO Box 22085, 46071 Valencia, Spain
[2]Sea Mammal Research Unit, Gatty Marine Laboratory, University of St Andrews, St Andrews, Fife KY16 8LB, UK
[3]ALNITAK, Nalón 16, 28240 Hoyo de Manzanares, Madrid, Spain

ABSTRACT: Spatial modelling is increasingly being used as an alternative to conventional design-based line transect sampling to estimate cetacean abundance. This new method combines line transect sampling with spatial analysis to predict animal abundance based on the relationship of animals observed to environmental factors. It presents several advantages including: (1) the ability to use data collected from 'platforms of opportunity', (2) the ability to estimate abundance for any defined subarea within the study area, and (3) the possibility for increased precision if covariates explain sufficient variability in the data. One study has been conducted to compare spatial modelling with conventional line transect methods, but the use of covariates in the detection function and the inclusion of school size have not previously been investigated. In the present study, the density of striped dolphins *Stenella coeruleoalba* was estimated in western Mediterranean waters using spatial distance sampling models applying generalised additive models (GAMs). This estimate was compared with density values previously estimated in the same area using conventional line transect methods. The densities estimated were very similar: 0.494 animals km^{-2} (coefficient of variation, CV = 0.16) using spatial models and 0.489 animals km^{-2} (CV = 0.19) using conventional line transect methods. Densities were also similar when they were calculated in stratified areas defined during the original line transect study. The precision of the estimates from spatial modelling was higher than that of the estimates obtained from conventional line transect analysis, particularly in the subareas. The results confirm that spatial modelling is a good approach for estimating cetacean abundance.

KEY WORDS: Population size estimation · Spatial models · Distance sampling · Generalised additive models · GAMs · Striped dolphin

INTRODUCTION

Successful strategies for the management and conservation of vulnerable species, such as cetaceans, require reliable information on abundance and population trends. Information on abundance allows evaluation of the impact of threats affecting the population, which is a necessary step in the establishment of effective conservation measures (for example, to determine the sustainable number of animals that could be removed from a population as by-catch in fishing gear) (Reeves et al. 2003). Monitoring trends in abundance allow the conservation status of a particular species to be determined.

A standard technique used to estimate the abundance of wild animal populations is distance sampling (Buckland et al. 2001, 2004). This method is conventionally implemented by selecting transect lines that achieve equal coverage probability in a survey area, and is generally known as design-based line transect sampling. The method estimates density along a set of transects via a detection function fitted to the perpen-

*Email: amaia.gomez@uv.es

dicular distances of observed groups of animals to estimate the effective width of the strip searched on each side of the transect. The density estimated is extrapolated to the whole study area on the basis that placement of the transects provides equal coverage probability throughout the area. Design-based line transect surveys have been used successfully all over the world to estimate cetacean abundance (Wade & Gerrodette 1993, Forcada & Hammond 1998, Forney & Barlow 1998, Hammond et al. 2002, Mullin & Fulling 2003, Slooten et al. 2004).

An alternative way to estimate animal abundance using line transect sampling is the recently developed method of spatial modelling (Hedley et al. 1999, Buckland et al. 2004). This method combines line transect sampling with spatial analysis to predict animal abundance based on the relationship of animals observed to environmental factors, as well as taking into account the probability of detecting the animals. This methodology has recently been used to estimate cetacean abundance (Williams 2003, Cañadas & Hammond 2006) because it has a number of advantages over the methodology previously used (Buckland et al. 2004, Hedley & Buckland 2004).

One advantage of this new method is that it does not require transect lines to be located in order to achieve equal coverage probability for the whole area. It is thus an appropriate method for analysing data collected from dedicated surveys with non-systematic design or from the so-called 'platforms of opportunity', which may include oceanographic, fisheries, or other survey vessels, navy ships and ferries (Buckland et al. 2004). Another advantage is that spatial distance sampling models allow the estimation of abundance in any subset of the study area and the creation of surface maps of animal abundance. This is very useful for application of results to the conservation of wildlife populations, because it allows the delimitation of areas of high density that may be suitable as marine protected areas. Conventional line transect methods allow abundance to be estimated in subareas or strata of the study area as long as these strata have been surveyed to provide equal coverage probability; strata must therefore be defined at the design stage of the survey. Typically, the precision of the estimates decreases markedly with the area of the strata. An additional advantage of spatial distance sampling models is that the inclusion of environmental features when predicting abundance may increase the precision of the estimate. If some heterogeneity in density observed along the trackline represents real spatial variation rather than sampling error, variances can be reduced because the model will explain some of the heterogeneity (Hedley et al. 1999, Forney 2000, Williams 2003).

In the present study we estimated the density of the striped dolphins *Stenella coeruleoalba* in central Spanish Mediterranean waters using spatial distance sampling models applying generalised additive models (GAMs) (Hastie & Tibshirani 1990). These estimates (value, precision and spatial differences) are compared to results obtained in a previous study in the same area (Gómez de Segura et al. 2006) but which used conventional design-based line transect methods to assess the potential advantages of spatial modelling in estimating cetacean abundance. Hedley et al. (1999) and Hedley & Buckland (2004) conducted a similar comparison of these 2 methods with regard to whale abundance in the Antarctic. In our study, we extended this comparison to include covariates in estimation of the detection function and to account for school size and availability bias in the spatial models.

In the present study, spatial models were used as predictive models to estimate dolphin abundance, but not as explanatory models to make general inferences about the ecology of the species. Distribution and habitat use of striped dolphins (and other cetacean species) in the same area will be considered in a separate study.

MATERIALS AND METHODS

Data collection. The study area comprised the waters off eastern central Spain (western Mediterranean) from 40°41′N, 0°53′E to 37°22′N, 1°38′W, with depths ranging from 10 to 2800 m. The overall area was approximately 34 200 km^2 (Fig. 1).

The field methods used for this study to estimate cetacean abundance by means of conventional design-based distance sampling method are described in detail in Gómez de Segura et al. (2006). Seasonal surveys were conducted from May 2001 to March 2003 with a high-wing aircraft ('push-pull' Cessna 337) flying at an altitude of 152 m, at a groundspeed of approximately 166 km h^{-1} (90 knots). Transects followed a systematic saw-tooth pattern and were designed to provide equal coverage probability of the area (Buckland et al. 2001). Two observers, one on each side of the aircraft, scanned the surface of the sea and a third researcher recorded the data: species, number of animals, location (obtained from a global positioning system, GPS), time, observer making the sighting, angle between the horizon and the target when perpendicular to the aircraft (to estimate distance to the school), and environmental conditions, including Beaufort sea state. Environmental conditions were updated whenever changes occurred, and the GPS provided a continuous record of position (updated every few seconds). Surveys were conducted only in good sighting conditions (i.e. Beaufort sea state ≤3).

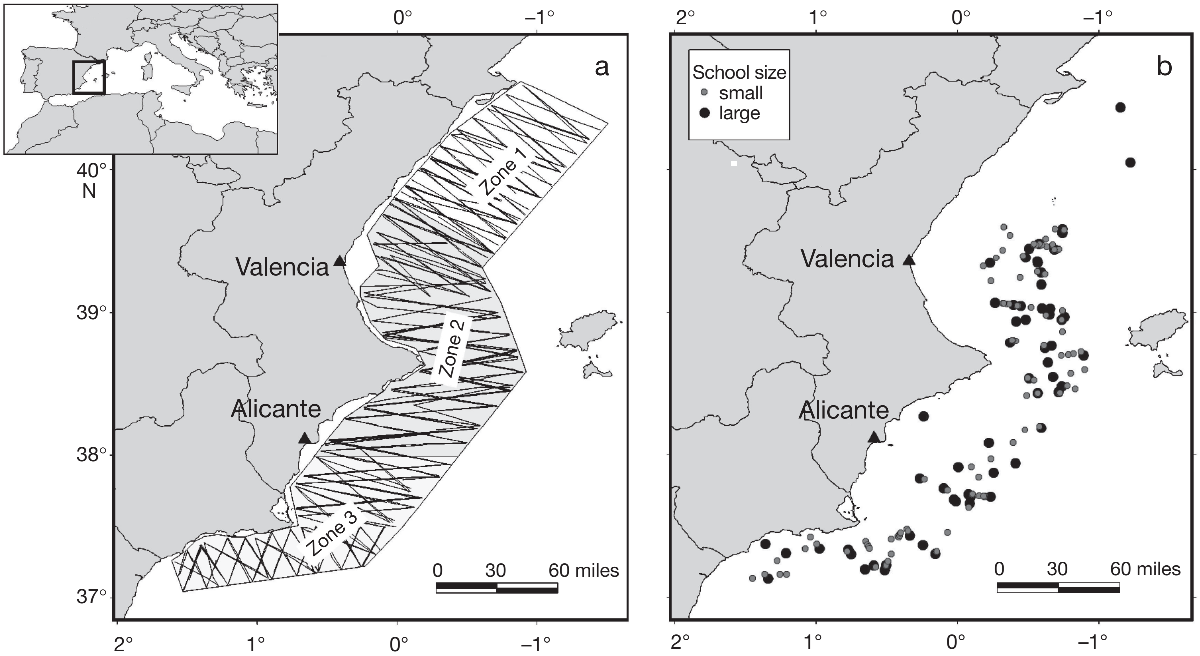


Fig. 1. *Stenella coeruleoalba.* (a) Study area surveyed, divided into the 3 subareas (zones) (see 'Materials and methods'), with transect lines searched on effort during the aerial surveys conducted from 2001 to 2003. (b) Distribution of the observed sightings of striped dolphins during the surveys. Small schools: ≤15 individuals; large schools: >15 individuals

Based on the available data in the area and previous studies of striped dolphin habitat use in the western Mediterranean (Cañadas et al. 2002, 2005, A. Gómez de Segura et al. unpubl. data), the following environmental variables were used to predict dolphin abundance: (1) water depth, derived from the General Bathymetric Chart of the Oceans (GEBCO); (2) slope, calculated as maximum depth – minimum depth/distance between them; (3) annual average of sea-surface temperature (SST); (4) temporal variability in SST (standard deviation of the monthly average SST over the year); (5) latitude; and (6) longitude. Data on SST were extracted from satellite images with a pixel resolution of 2 km^2. Images were obtained from the CREPAD service (Centro de Recepción, Proceso, Archivo y Distribución de Imágenes de Observación de la Tierra) received from the NOAA sensor AVHRR.

Data analysis. All on-effort legs were divided into 5 nautical mile (n mile, 9.26 km) segments using a geographic information system (GIS) (ArcMap 8.2). This length was chosen to balance minimising the number of segments without sightings (too many makes model fitting problematic) without losing too much resolution in the environmental covariates, which can reduce explanatory power. Segments at the ends of the transect that were <5 n miles in length were not used in the analysis. Each segment was characterised by the number of dolphins, the number of schools and the arithmetic mean value of each environmental variable analysed. It was assumed that there was little variability in the environmental covariates and in the sighting conditions within a segment. Models were fitted to the data in these segments to predict dolphin abundance along the transect line as a function of the environmental covariates.

The study area was divided into grid cells with a resolution of 5′ latitude (5 n miles) by 5′ longitude (3.8 n miles), which were characterised by the mean of each environmental variable. Density predicted by the model was extrapolated to each grid cell as a function of these values to predict the abundance of dolphins over the whole area.

Hedley et al. (1999) proposed various options for estimating abundance of cetaceans using spatial models, the simplest of which is to model the number of animals observed in each segment. However, the overdispersion of our data, due to the high variability in school size (from 1 to 100 individuals), was too great to allow the robust fitting of any function. Therefore, we decided to model separately the number of schools and school size and then multiply both values to obtain dolphin abundance (Hedley & Buckland 2004). The method can be summarised by, first, estimating the number of schools in each segment by correcting for schools missed as a function of perpendicular distance, poor visibility conditions, or because animals were submerged. Then, the estimated number of schools in the segments was modelled as a function of the environmental covariates using GAMs, and, based on this, a prediction of the number of schools in each grid cell of the study area was made. Separate models were run for 2 subsets of data, small and large schools, because the probability of detection differed between them (see following subsection). The number of schools predicted in each cell for small and large schools was multiplied by the mean school size for each subset. Finally, both surface maps were summed to obtain a map of the total abundance of striped dolphins.

Estimation of the number of schools per segment: The estimated number of schools was calculated, taking into account schools missed during the survey due to factors that decrease the probability of detection, i.e. perpendicular distance to the observer, school size, environmental conditions, etc. In addition, the proportion of schools diving, and thus not available to the observer, was also considered.

The estimated number of schools in segment i, N_i, was calculated using the Horvitz Thompson estimator (Horvitz & Thompson 1952):

$$N_i = \sum_{j=1}^{n_i} \frac{1}{P_{ij}} \qquad (1)$$

where n_i is the number of detected schools in the ith segment, and P_{ij} is the estimated probability of detection of the jth detected school in the segment i.

The estimated probability of detection was obtained by fitting a function to the distance frequency histogram using multiple covariates distance sampling (MCDS) implemented in DISTANCE 4.2 software (Thomas et al. 2004). With this new method, P_{ij} depends, not only on the perpendicular distance to the observer, as in conventional distance sampling, but also on other covariates that affect the probability of detecting a school (Buckland et al. 2004). Three potential functions were considered to fit the perpendicular distance data: uniform, half-normal and hazard-rate, together with various adjustment terms (Buckland et al. 2001). The following covariates were included in the detection function: school size (treated as a continuous variable), sea state (treated as a factor) and time of the day (treated as a factor: morning, midday, afternoon). A range of detection functions (combining the functional forms and the covariates) were investigated, and the best function was selected using Akaike's information criterion (AIC) (Buckland et al. 2004).

To correct for submerged schools that were not available for detection, N_i was divided by $\hat{a}(S,0)$, the estimated availability probability on the trackline, after Laake et al. (1997) and Forcada et al. (2004). Gómez de Segura et al. (2006) estimated the values of $\hat{a}(S,0)$ for striped dolphin schools in the study area derived from shipboard observations of diving behaviour. Data from this study were used for our analyses. The $\hat{a}(S,0)$ values estimated varied according to school size, with $\hat{a}(S,0)$ = 0.67 for small schools and $\hat{a}(S,0)$ = 1 for large schools, see 'Estimation of school size' below. Although we could correct the estimated number of schools separately for both groups and sum the total estimated number of schools, this would not take into account the different variances of $\hat{a}(S,0)$ when estimating the precision of estimated abundance. Because of this, 2 separate school density models were constructed, 1 for small schools (≤15 individuals) and 1 for large schools (>15 individuals).

Spatial modelling of the number of schools: GAMs (Hastie & Tibshirani 1990) were used to model the abundance of schools based on the relationship of animals observed to environmental factors. The estimated number of schools (separated for small and large schools) was used as the response variable, and environmental parameters were used as the predictor covariates. Although the response variable was derived from count data, the Poisson error distribution was not considered appropriate because of over-dispersion. Instead, a quasi-Poisson family was used, with variance proportional to the mean (Hedley et al. 1999). Applying a logarithmic link function, the general structure of the model was:

$$N_i = \exp\left[\ln(a_i) + \theta_0 + \sum_k f_k(z_{ik})\right] \qquad (2)$$

where the offset variable a_i is the area surveyed along the ith segment (calculated as the length of the segment multiplied by twice the perpendicular distance at which data were truncated); θ_0 is a parameter to be estimated (commonly termed the 'intercept'); f_k are smoothed functions of the explanatory covariates; and z_{ik} is the value of the kth explanatory covariate in the ith segment. We also considered model interactions between pairs of explanatory variables (such as depth with latitude), but no significant interaction was observed.

The 'mgcv' package (Version 1.0-5) (Wood 2001) from R software (Version 1.9.0) (http://www.r-project.org) was used to fit the models. Two indicators were used to select the best-fitting model: (1) the general cross validation (GCV) score, which is, in practice, an approximation to AIC (Wood 2000) in which smoothing parameters (in terms of number of knots and degrees of freedom, df) are chosen by the software to minimise the GCV score of the model and (2) the percentage of explained deviance.

Finally, the number of small and large schools was predicted for each cell of the study area using the values of the environmental variables in each cell and the model selected. The results were displayed on a map using GIS.

Estimation of school size: Estimation of spatial school size surfaces using spatial modelling was attempted

(Borchers & Burt 2002, Hedley & Buckland 2004, Cañadas & Hammond 2006), because school size was expected to vary spatially across the study area. In this case, the response variable for the GAMs was the mean number of individuals per school and the predictor variables were the environmental factors. However, school size did not vary spatially, and the number of individuals was not related to any covariate. Thus, the estimated mean school size (E_s) of each school size subset was used instead (Buckland et al. 2001). E_s accounts for the relationship between school size and detectability that often occurs in cetacean line transect surveys (Forcada & Hammond 1998, Forney & Barlow 1998, Gómez de Segura et al. 2006). A regression of log school size against the estimated probability of detection [$g(x)$] was computed. If the regression was significant, mean school size was estimated as the predicted mean size of schools detected in the region around the trackline, where detection is assumed to be certain. If the regression was not significant, the mean school size was used as the E_s value (Buckland et al. 2001).

Estimation of abundance: The predicted surface maps of the number of small and large schools were multiplied by the estimated mean school size for each subset. The total number of dolphins in each grid cell was then estimated by summing the estimated dolphin abundance of small and large schools. Finally, the total abundance of dolphins in the study area was obtained by summing the abundance of all grid cells, and density was calculated by dividing this abundance by the total area.

Estimation of variance: The variance of school density was estimated by means of non-parametric bootstrap (400 replicates with replacement), with the R software using transect lines as sampling units (Hedley & Buckland 2004). For each replacement, the same number of transects as for the original data were sampled. The final variance of dolphin abundance was estimated using the delta method (Seber 1982), combining the variance of school density with the variance of the detection function, the variance of $\hat{a}(S,0)$ and the variance of the estimated mean school size (Hedley & Buckland 2004). The values of the coefficients of variation were plotted as surface maps of variability.

RESULTS

Altogether 10 surveys were conducted in the study area during the field work: in May, July and October 2001, March, June, August and December 2002, and March 2003. A total of 20 200 km were searched on effort (Fig. 1a) during which 162 striped dolphin *Stenella coeruleoalba* schools (49 small and 113 large schools) were observed (Fig. 1b).

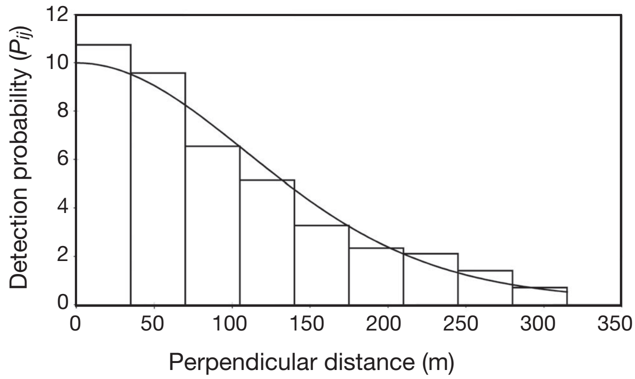


Fig. 2. *Stenella coeruleoalba.* Frequency distribution of perpendicular distances from the transect line pooled over all covariates. The continuous curve represents the fitted detection function. For details on estimation of detection probability see 'Results'

The half-normal function with school size and time of day as covariates was identified as the best-fitting model for the detection function according to AIC. Fig. 2 shows the perpendicular distance frequency histogram and the best fitted detection function. This function was used to estimate the detection probability of each school sighted, P_{ij}, from the following formula:

$$P(y) = \exp[-y^2/(2s^2)]$$
$$s = a1 \times \exp[f(a2) + f(a3)] \quad (3)$$

where y is the distance to the trackline, f is the function $a1$ is the intercept of the scale parameter s, $a2$ is the coefficient of covariate 'cluster size' and $a3$ is the coefficient of level 'midday' of factor covariate 'time day'.

GAMs were run with the estimated number of schools in each segment (both for small and large schools) and the environmental covariates. Various models were considered to fit equally well due to the similarity in the GCV and the percentage of deviance explained (Table 1). For small schools, the model in-

Table 1. *Stenella coeruleoalba.* Best-fitting models in the generalised additive model (GAM) analyses for small and large schools of striped dolphin, with indicators used to select the best-fitting model (lat: latitude; lon: longitude; GCV: general cross validation). Shading indicates the final model selected (see 'Materials and methods')

Schools	Model	% deviance explained	GCV score
Small schools	Depth–lat–lon	23.9	3.7479
	Depth–lon	23.1	3.8914
	Depth	19.4	4.4629
Large schools	Depth–slope	16.8	1.5464
	Depth	15.0	1.5901

Table 2. *Stenella coeruleoalba.* Environmental covariates selected in the best-fitting model for small and large schools of striped dolphin. The degrees of freedom (df) and the significance of (p) for the chi-squared test are also shown

Model	Term	df	Chi-squared	p-value
Small schools	Depth	5.91	28.95	<0.001
	Latitude	4.14	5.42	0.263
	Longitude	7.01	15.16	0.035
Large schools	Depth	4.73	38.63	<0.001
	Slope	2.89	4.51	0.199

Table 3. *Stenella coeruleoalba.* Estimates of abundance (N) and density (D) of striped dolphin, with the coefficient of variation (CV) obtained from spatial analysis. Parameters used to estimate abundance, estimated mean school size (E_s), probability of detection (P) and availability bias [$\hat{a}(S,0)$] are also shown

	Parameter	Estimate	CV
Small schools	$N_{schools}$	1266	0.22
	E_s	5.46	0.07
	P	0.46	0.06
	$\hat{a}(S,0)$	0.68	0.24
	$N_{ind.}$	6919	0.34
Large schools	$N_{schools}$	296	0.13
	E_s	33.7	0.08
	P	0.46	0.06
	$\hat{a}(S,0)$	1.00	0.00
	$N_{ind.}$	9973	0.16
Total	N	16 892	0.16
	D	0.494	0.16

cluding depth, latitude and longitude was the best-fitting model, but for large schools the best-fitting model included depth and slope. However, latitude and slope, respectively, were dropped from these models because they were not significantly related to dolphin distribution, and the GCV and deviance explained by the model including them did not improve significantly (Tables 1 & 2). The selected models therefore included depth and longitude for small schools and depth only for large schools. All models yielded quite similar abundance estimates (see below), indicating that the selected model is robust. Fig. 3 shows the shape of the functional forms for the smoothed covariates used in the selected models. Of these variables, depth was the most important for both small and large schools as shown by the p-values (Table 2).

The environmental covariates selected were used to predict the number of schools in each cell of the study area. The estimated number of schools in each cell [$E(N_i)$] was calculated from the selected model:

$$E(N_i) = \exp\left[\ln(a_i) + \theta_0 + s(\text{depth}, 6.4) + s(\text{lon}, 7.3)\right] \quad (4)$$

for small schools, and

$$E(N_i) = \exp\left[\ln(a_i) + \theta_0 + s(\text{depth}, 4.6)\right]$$

for large schools (5)

where the offset variable a_i is the area of the cell, θ_0 is the 'intercept' of the model, lon is longitude, and $s(\ldots)$ indicates the estimated smoothed function of the specified covariate and its estimated degrees of freedom.

The estimated numbers of schools in the study area were 1266 small and 296 large schools (Table 3). The regression of school size against detection probability was not significant in the case of small schools; thus,

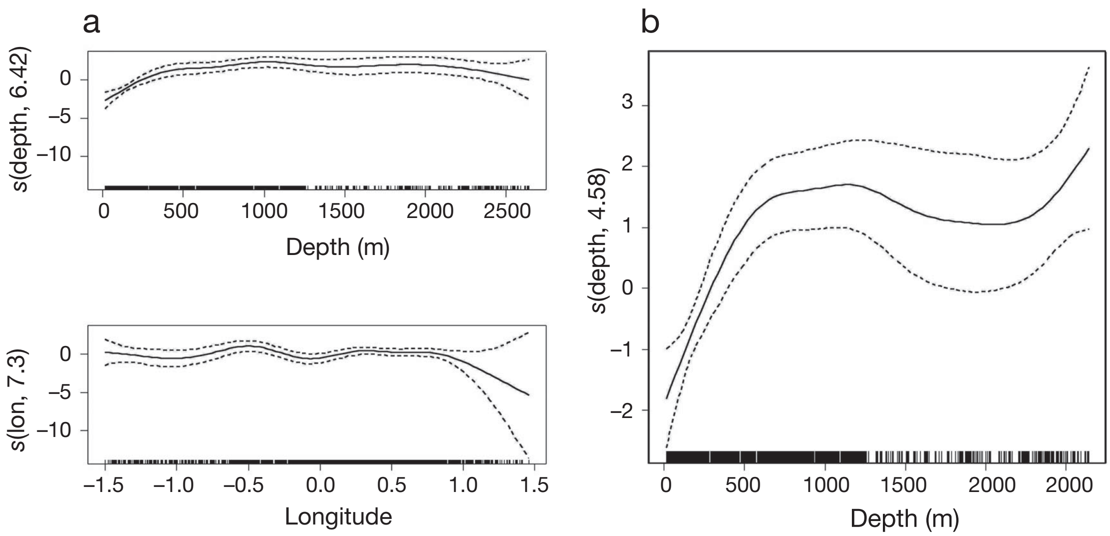


Fig. 3. *Stenella coeruleoalba.* Plots of the GAM smooth fits of the environmental covariates selected for (a) small and (b) large schools of striped dolphin. Solid lines represent the best fit, dashed lines represent the 2 SE confidence limits, and vertical lines on the *x*-axis are the observed data values

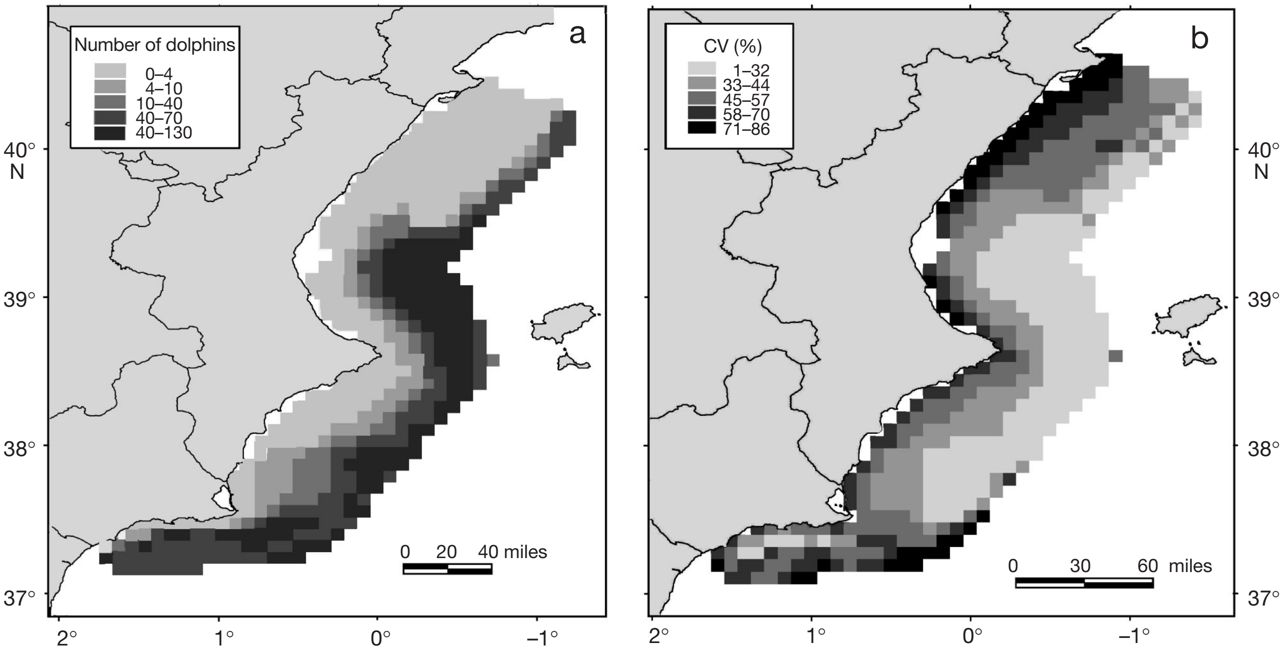


Fig. 4. *Stenella coeruleoalba.* Surface map of (a) predicted abundance of striped dolphins using the model-based method, and (b) the coefficient of variation (CV) of estimated abundance

the mean school size was used. However, regression was significant for large schools, so that E_s was used to estimate dolphin abundance. The mean school size for small schools was 5.5 individuals, and the estimated mean school size for large schools was 33.7 individuals (Table 3). The total abundance of dolphins in the study area was estimated at 16 892 animals (CV = 0.16), yielding a density of 0.494 animals km^{-2} (CV = 0.16) (Table 3).

Table 4 shows the mean density in the study area estimated by the spatial models (present study) and design-based line transect methods (extracted from Gómez de Segura et al. 2006). Densities estimated in the different subareas (zones), defined in Gómez de Segura et al. (2006), using both methods are also shown. Comparing the results of the 2 methodologies, both the mean density and the densities estimated in the subareas were very similar.

Fig. 4 shows the predicted surface of abundance of striped dolphins in the study area based on the spatial models and the variation associated with the abundance estimated. The prediction is well supported by the observed data (Fig. 1b), indicating that the model provides a reliable description of the spatial variation of striped dolphin abundance although it only explained 15 to 23 % of the deviance. The surface map showed 1 area of high density in front of the Gulf of Valencia (approximately 39.4° N, 0.6° E) and some other small areas in the south-eastern part of the study area.

DISCUSSION

Density obtained using spatial distance sampling models was very similar to that obtained by means of conventional design-based line transect sampling (Table 4). Comparing densities of *Stenella coeruleoalba* estimated in the different subareas, the values were also very similar (Table 4). Furthermore, the estimates obtained from different models with similar GCV and deviance explained were mostly quite similar, indicating that the prediction is quite robust (Table 5). This confirms that spatial distance sampling models are a good approach to estimate cetacean abundance. Hedley et al. (1999) and Hedley & Buckland (2004) reached the same conclusion, obtaining slightly higher abundance estimates with spatial

Table 4. *Stenella coeruleoalba.* Estimates of density of striped dolphins (*D*), with the coefficient of variation (CV), from conventional stratified design-based distance sampling (Gómez de Segura et al. 2006) and from the spatial analysis

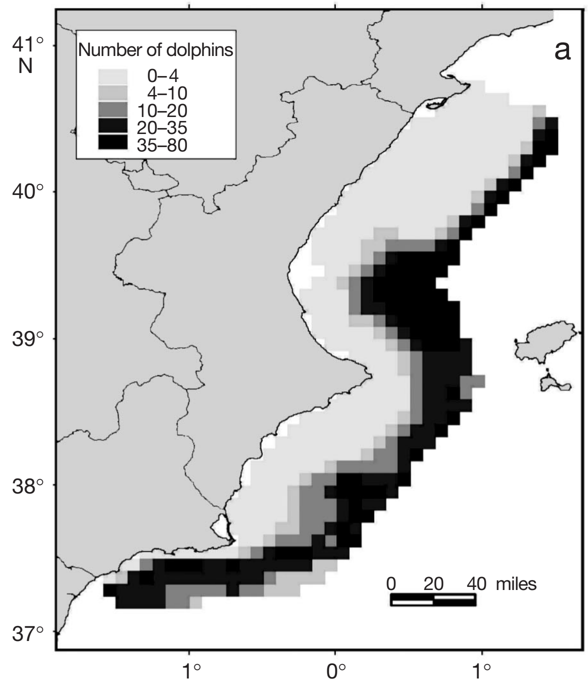
Zone	Spatial models		Line transect	
	D	CV	*D*	CV
1	0.182	0.15	0.268	0.31
2	0.551	0.15	0.529	0.22
3	0.706	0.22	0.672	0.23
Total	0.494	0.16	0.489	0.19

models than with the conventional distance sampling method.

In our study, group size, covariates and availability bias were included so that we could obtain not only the abundance of schools at the surface (as did Hedley et al. 1999 and Hedley & Buckland 2004), but also the absolute abundance of animals. The use of both $\hat{a}(S,0)$ and covariates to estimate the number of schools, and the use of the estimated mean school size to calculate the number of animals, have proven to be a good procedure to estimate absolute density in spatial distance sampling models.

Despite the good results obtained in the present study, it is recommended that spatial modelling be used with caution because GAMs are very flexible models and model selection in GAMs is still an area of active research (Hedley & Buckland 2004). In particular, overfitting and 'edge effects' could yield unrealistic densities and surface maps, and these should be considered carefully during model selection.

On the other hand, one of the potential advantages of spatial models is that they allow estimation of abundance at any spatial resolution (Buckland et al. 2004). In this study, although all models selected in the GAM analysis resulted in quite similar stratified abundances (Table 5), estimated abundance at very small scales can vary depending on the model selected (see Figs. 5 & 6). Hedley & Buckland (2004) also detected this lack of robustness at very small scales when comparing the density maps predicted by different models. Furthermore, density estimates in small areas are not very informative from an ecological perspective. Most cetacean species can move rapidly over large distances, and density estimated in a small area is likely to change over short periods of time. In addition, our estimates of density in small areas are derived from the



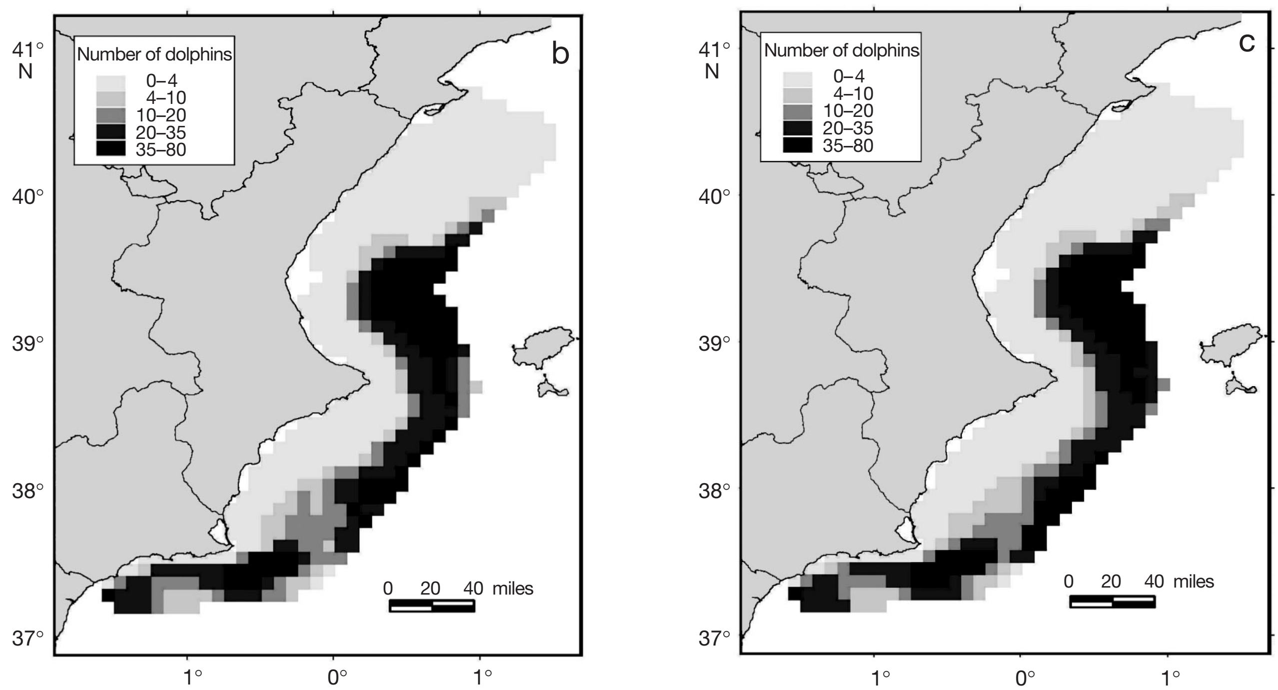


Fig. 5. *Stenella coeruleoalba.* Surface map of predicted abundance of small schools using (a) the model including only depth, (b) the model including depth and longitude and (c) the model including depth, longitude and latitude

Table 5. *Stenella coeruleoalba.* Abundance estimates of striped dolphins for the 3 strata obtained from the best fitted models in the spatial analysis for small and large schools (lat: latitude; lon: longitude)

Model	Small schools			Large schools	
	Depth	Depth–lon	Depth–lon–lat	Depth	Depth–slope
Zone 1	986	396	363	1321	1203
Zone 2	3409	3776	3766	4523	4558
Zone 3	2879	2747	2903	4130	2959
Total	7274	6919	7032	9973	8720

model fitted to all the data with the associated assumption that factors determining abundance at small scales are the same as those at large scales. Therefore, care should be taken when making inferences about estimates of abundance in small areas.

In addition, the possibility of spatial autocorrelation in the data should be assessed in cetacean-habitat modelling (Redfern et al. 2006): schools might be found in the same area for reasons unrelated to the modelled environmental covariates, e.g. because of the social behaviour of the animals. Ignoring autocorrelation can overestimate the correlation between animal distribution and environmental variables (Ferguson & Bester 2002). At present, no approach has been developed to address this issue using GAMs. Nevertheless, in our study, the autocorrelation is probably not a critical problem, since the aim of the study is to predict density, not to study the habitat use of animals.

The precision of estimated abundance using spatial distance sampling models improved over estimates obtained using conventional design-based line transect sampling. The coefficient of variation of density estimated by spatial models was slightly lower for the mean abundance and Zone 3, but much lower for the other 2 zones (Table 4). When spatial models are used, an increase in precision is expected if the model explains some of the spatial heterogeneity of animal distribution (Hedley et al. 1999, Forney 2000, Williams 2003). In our study, although the deviance explained by the models was low (15 to 23%; Table 1), the precision increased considerably using spatial analysis. When stratifying using conventional design-based distance sampling, the number of data available for analysis (number of samples and replicates) decreases, and therefore the coefficient of variation increases. This effect does not occur in spatial models, because data from outside a stratum provide relevant information to the estimation of density within that stratum. Hedley et al. (1999) and Hedley & Buckland (2004) also obtained an increase in precision using spatial models both in the mean and stratified abundance estimates.

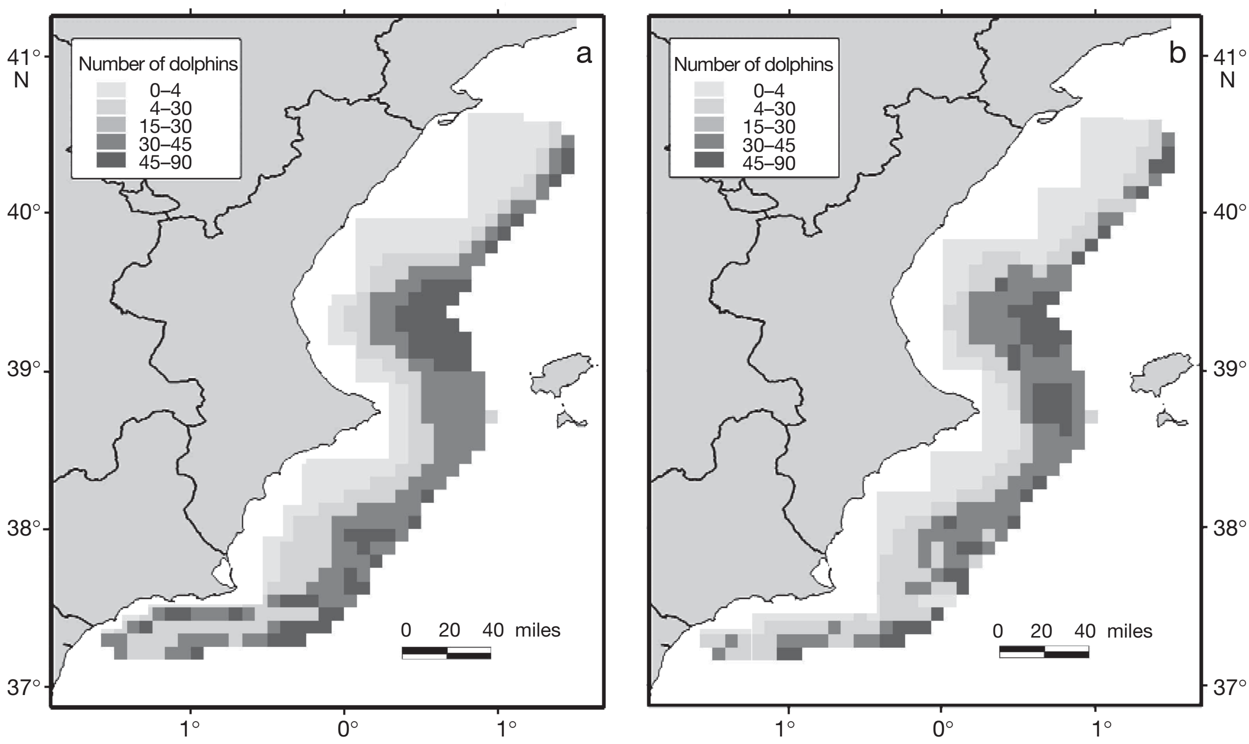


Fig. 6. *Stenella coeruleoalba.* Surface map of predicted abundance of large schools using (a) the model including only depth and (b) the model including depth and slope

The explanatory value of the models built for small and large schools was moderate; the adjusted R-square score was 0.04 and 0.02, and the deviance explained was 23 and 15%, respectively (Table 1). The low deviance explained by the models could be caused by some overdispersion of the data (such as 4 schools observed on the continental shelf), because, in general, the surface map predicted by the model fits the data observed during surveys quite well (compare Figs. 1b & 4). The low proportion of positive observations in the dataset, due to the low density of sightings, also contributes to the low deviance explained. Furthermore, the fact that data had to be separated in 2 subsets, small and large schools, and a model fitted for each one, further decreases the proportion of positive observations, and therefore the deviance explained.

In general, it is difficult to model the distribution of cetacean species due to their complex ecology and behaviour, because they live in a highly dynamic environment and because they can move rapidly over large distances (Redfern et al. 2006). Additionally, there might be other factors affecting the distribution of striped dolphins that were not considered in this study. For example, feeding ecology is likely the prime factor influencing cetacean distribution (Gaskin 1976), but spatial information on distribution and abundance of prey species is usually difficult to obtain or even unavailable, as was the case for the present analysis.

In conclusion, if the main goal of a study is simply to estimate animal abundance, conventional design-based line transect methods might be preferred, as they do not require environmental data (which can be difficult to obtain) and the analyses are less complicated. However, there is an increasing number of studies on cetaceans in which it has been possible to collect robust line transect data, but the platform has been unable to follow a design ensuring equal coverage probability (Cañadas & Hammond 2006, Williams et al. 2006). In these cases, if suitable environmental data are available, spatial distance sampling models are a good approach to estimate abundance. In addition, spatial models provide valuable information on habitat use of animals (Redfern et al. 2006), which can further understanding of processes that determine animal distribution, a fundamental problem in ecology. We will address these questions for striped dolphins and other cetacean species in our study area in a future paper. Furthermore, this information can be very useful for conservation and management strategies, e.g. it can be used to predict 'hot spots' of high density or areas important for feeding or calving that may be suitable for designation as marine protected areas (Hooker et al. 1999, Cañadas et al. 2005).

Acknowledgements. This study was conducted under the project 'Programa de Identificación de las Áreas de Especial Interés para la Conservación de los Cetáceos en el Mediterráneo Español', funded by the Spanish Environment Ministry. The support of the Conselleria de Territorio y Vivienda of the Generalitat Valenciana particularly to J. Jimenez is greatly appreciated. We thank the CREPAD service of the INTA (National Institute for Aerospace Technology) for the satellite images provided. We are very grateful to all the observers who assisted in data collection: J. Martínez, P. Gozalvez, J. Barona, K. Lehnert and D. Perdiguero, as well as all pilots from AVIALSA, especially to Jose Manuel Garcia. Thanks are due also to J. Tomás, C. Agustí and K. Crespo for their valuable contributions and support in this study. The comments of Megan Ferguson are greatly appreciated.

LITERATURE CITED

Borchers DL, Burt ML (2002) Generalized regression methods for estimating school size from line transect data. Paper SC/54/IA23, Scientific Committee of the International Whaling Commission, Shimonoseki, Japan

Buckland ST, Anderson DR, Burnham KP, Laake JL, Borchers DL, Thomas L (2001) Introduction to distance sampling: estimating abundance of biological populations. Oxford University Press, Oxford

Buckland ST, Anderson DR, Burnham KP, Laake JL, Borchers DL, Thomas L (2004) Advance distance sampling: estimating abundance of biological populations. Oxford University Press, Oxford

Cañadas A, Hammond PS (2006) Model-based abundance estimates for bottlenose dolphins (*Tursiops truncatus*) off southern Spain: implications for management. J Cetacean Res Manage 8(1):13–27

Cañadas A, Sagarminaga R, Garcia-Tiscar S (2002) Cetacean distribution related with depth and slope in the Mediterranean waters off southern Spain. Deep-Sea Res 49: 2053–2073

Cañadas A, Sagarminaga R, de Stephanis R, Urquiola E, Hammond PS (2005) Habitat selection modelling as a conservation tool: proposals for Marine Protected Areas for cetaceans in southern Spain. Aquat Conserv 15:495–521

Ferguson JWH, Bester MN (2002) The treatment of spatial autocorrelation in biological surveys: the case of line transect surveys. Antarct Sci 14(2):115–122

Forcada J, Hammond PS (1998) Geographical variation in abundance of striped and common dolphins of the western Mediterranean. J Sea Res 39:313–325

Forcada J, Gazo M, Aguilar A, Gonzalvo J, Fernandez-Contreras M (2004) Bottlenose dolphin abundance in the NW Mediterranean: addressing heterogeneity in distribution. Mar Ecol Prog Ser 275:275–287

Forney KA (2000) Environmental models of cetacean abundance: reducing uncertainty in population trends. Conserv Biol 14(5):1271–1286

Forney KA, Barlow J (1998) Seasonal patterns in the abundance and distribution of California cetaceans, 1991–1992. Mar Mamm Sci 14:460–489

Gaskin DE (1976) The evolution, zoogeography and ecology of Cetacea. Oceanogr Mar Biol Annu Rev 14:247–346

Gómez de Segura A, Crespo EA, Pedraza SN, Hammond PS, Raga JA (2006) Abundance of small cetaceans in the waters of the central Spanish Mediterranean. Mar Biol 150:149–160

Hammond PS, Berggren P, Benke H, Borchers DL and 6 others (2002) Abundance of harbour porpoises and other

cetaceans in the North Sea and adjacent waters. J Appl Ecol 39:361–376
Hastie TJ, Tibshirani RJ (1990) Generalized additive models. Chapman and Hall, London
Hedley SL, Buckland ST (2004) Spatial models for line transect sampling. J Agric Biol Environ Stat 9(2):181–199
Hedley SL, Buckland ST, Borchers DL (1999) Spatial modelling from line transect data. J Cetacean Res Manage 1(3): 255–264
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
Horvitz DG, Thompson DJ (1952) A generalisation of sampling without replacement from a finite universe. J Am Stat Assoc 47:663–685
Laake JL, Calambokidis J, Osmek SD, Rugh DJ (1997) Probability of detecting harbour porpoise from aerial surveys: estimating g(0). J Wildl Manag 61:63–75
Mullin KD, Fulling GL (2003) Abundance of cetaceans in the southern U.S. North Atlantic Ocean during summer 1998. Fish Bull 101:603–613
Redfern JV, Ferguson MC, Becker EA, Hyrenbach KD and 15 others (2006) Techniques for cetacean–habitat modelling: a review. Mar Ecol Prog Ser 310:271–295
Reeves RR, Smith B, Crespo EA, Notarbartolo di Sciara G (eds) (2003) Dolphins, whales and porpoises: 2002–2010. Conservation action plans for the world's cetaceans. IUCN/SSC Cetacean Specialist Group, IUCN, Gland
Seber GAF (1982) The estimation of animal abundance and related parameters. Macmillan, New York
Slooten E, Dawson SM, Rayment WJ (2004) Aerial surveys for coastal dolphins: abundance of Hector's dolphins off the South Island West Coast, New Zealand. Mar Mamm Sci 20:477–490
Thomas L, Laake JL, Strindberg S, Marques FFC and 7 others (2004) DISTANCE 4.1, Release 2. Research Unit for Wildlife Population Assessment, University of St. Andrews. Available at www.ruwpa.st-and.ac.uk/distance/
Wade PR, Gerrodette T (1993) Estimates of cetacean abundance and distribution in the eastern tropical Pacific. Rep Int Whal Comm 43:477–494
Williams R (2003) Cetacean studies using platforms of opportunity. PhD dissertation, University of St. Andrews
Williams R, Hedley SL, Hammond PS (2006) Modeling distribution and abundance of Antarctic baleen whales using ships of opportunity. Ecol Soc 11 (1): 1, available at www.ecologyandsociety.org/vol11/iss1/art1/
Wood SN (2000) Modeling and smoothing parameter estimation with multiple quadratic penalties. J R Stat Soc B 62(2): 413–428
Wood SN (2001) mgcv: GAMs and generalised ridge regression for R. R News 1(2):20–25

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

Submitted: February 14, 2006; Accepted: May 23, 2006
Proofs received from author(s): December 2, 2006