Journal of Biogeography (J. Biogeogr.) (2015) **42**, 1267–1280

ORIGINAL ARTICLE

Extrapolating cetacean densities beyond surveyed regions: habitat-based predictions in the circumtropical belt

Laura Mannocci[1]*, Pascal Monestiez[1,2], Jérôme Spitz[3,4] and Vincent Ridoux[1,3]

[1]*Centre d'Études Biologiques de Chizé et La Rochelle, UMR 7372 Université de La Rochelle-CNRS, La Rochelle F-17000, France,* [2]*INRA, UR0546, Unité Biostatistiques et Processus Spatiaux, Domaine Saint-Paul 84914, Avignon, France,* [3]*Observatoire PELAGIS, UMS 3462 Université de La Rochelle-CNRS, Systèmes d'Observation pour la Conservation des Mammifères et des Oiseaux Marins, La Rochelle 17000, France,* [4]*Marine Mammal Research Unit, Fisheries Center, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada*

*Correspondence and current address: Laura Mannocci, Marine Geospatial Ecology Lab, Nicholas School of the Environment, Duke University, Durham, NC 27708, USA.
E-mail: laura.mannocci@duke.edu

ABSTRACT

Aim Our knowledge of cetacean distributions is impeded by large data-gaps worldwide, particularly at tropical latitudes. This study aims to (1) find generic relationships between cetaceans and their habitats in a range of tropical waters, and (2) extrapolate cetacean densities in a circumtropical belt extending far beyond surveyed regions.

Location Pelagic, circumtropical.

Methods Aerial surveys were conducted over three regions in the tropical Atlantic (132,000 km^2), Indian (1.4 million km^2) and Pacific (1.4 million km^2) oceans. Three cetacean guilds were studied (Delphininae, Globicephalinae and sperm and beaked whales). For each guild, a generalized additive model was fitted using sightings recorded in all three regions and 14 candidate environmental predictors. Cetacean densities were tentatively extrapolated over a circumtropical belt, excluding waters where environmental characteristics departed from those encountered in the surveyed regions.

Results Each cetacean guild exhibited a relationship with the primary production and depth of the minimum dissolved oxygen concentration. Delphininae also showed a relationship with the dominant phytoplankton group. The prediction envelopes were primarily constrained by water temperature. Circumtropical extrapolations of Delphininae and Globicephalinae were contrasted between ocean basins, with high densities predicted in the equatorial waters of the three ocean basins. The predicted densities of sperm and beaked whales were lower and more uniform across the circumtropical belt than for the other two guilds.

Main conclusions Our modelling approach represents a good analytical solution to predicting cetacean population densities in poorly documented tropical waters. Future data collection should concentrate on areas where environmental characteristics were not encountered in our survey regions and where the predicted densities were the most uncertain. By highlighting cetacean hotspots far beyond waters under national jurisdiction, this study can provide guidance for the delimitation of Ecologically and Biologically Significant Marine Areas.

Keywords

Cetaceans, circumtropical, conservation biogeography, density, extrapolation, generalized additive models, pelagic waters.

INTRODUCTION

Cetaceans and other pelagic megafauna are charismatic species of major importance for biological conservation owing to several key characteristics. First, most are exposed to a variety of anthropogenic pressures that have driven many populations towards extinction (Schipper *et al.*, 2008). Second, the conservation of these species requires their use of

http://wileyonlinelibrary.com/journal/jbi
doi:10.1111/jbi.12530

resources and high-seas habitats to be taken into account when managing human activities (Hammond *et al.*, 2013). Finally, their distributions and population densities reflect the oceanographical processes that are associated with biological productivity and diversity 'hotspots' (Worm *et al.*, 2005), as well as anthropogenic pressures on marine ecosystems (Pompa *et al.*, 2011).

The monitoring of cetaceans is motivated by legal requirements and measures implemented by particular countries, such as the US Marine Mammal Protection Act and more recently the EU's Habitat Directive and Marine Strategy Framework Directive. Techniques for monitoring cetacean distributions include visual surveys (most commonly line-transect surveys carried out from boats or aircraft), acoustic surveys and tagging studies (Redfern *et al.*, 2006). Line-transect surveys have mainly been conducted within countries' exclusive economic zones (EEZs), which extend 200 nautical miles from their coastal baselines; the great proportion of the ocean that lies beyond these zones is known as the high seas. Kaschner *et al.* (2012) noted that only a quarter of the world's ocean surface had been covered by line-transect surveys, with most data having been collected on the continental shelves and slopes of the USA and northern Europe, as well as around the Antarctic continent. Our knowledge of cetacean distributions is thus impeded by large data-gaps, especially at tropical latitudes.

Two approaches could be pursued to fill these data-gaps. First, line transects could be implemented in regions that have not yet been surveyed. Although the collection of new data is the optimal solution for filling data-gaps, it would be very costly and time-consuming and is hardly achievable globally, because the autonomy of boats or aircrafts would limit the surveys to waters close to continents and inhabited islands. Alternatively, statistical models could be developed that relate cetacean sightings to environmental data. These habitat models would allow cetacean densities to be extrapolated into unsurveyed areas, provided the predictions are made within the range of environmental conditions encountered in the surveyed regions (Elith & Leathwick, 2009). Indeed, making predictions outside the range of environmental covariates used when fitting a model can lead to unreasonable results (Becker *et al.*, 2014).

We believe that robust statistical models based on high-quality sighting data collected in some representative regions, along with relevant environmental predictors, are valuable tools to fill these data gaps. In contrast to cetacean sightings, broad-scale environmental data are readily available. Satellite remote sensing has become instrumental for collecting environmental data at the ocean surface (e.g. chlorophyll-*a* concentration, sea-surface temperature) (Chassot *et al.*, 2011). *In situ* oceanographical measurements reveal environmental conditions along the vertical axis (e.g. temperature and nutrient profiles) and have been combined and interpolated worldwide into global climatologies (e.g. *CSIRO Atlas of Regional Seas*; Condie & Dunn, 2006). Thanks to depth soundings and satellite-derived gravity data, the topography of the sea floor is also well known. When assembled together, these environmental data constitute a large set of indicators of productivity and prey aggregation, ultimately determining the quality of pelagic habitats (Longhurst & Pauly, 1987).

Statistical models have been developed for cetaceans in many areas and have revealed relationships between cetacean species and their habitats in various ecosystems (e.g. Jaquet & Whitehead, 1996; Becker *et al.*, 2010; Forney *et al.*, 2012). These relationships remain specific to the studied ecosystem and cannot be generalized to other locations. Instead of investigating cetacean–habitat relationships in a particular location, it would be valuable to search for generic properties of cetacean habitats. Generic properties are used here to mean cetacean–habitat relationships that would be valid within a broad range of environmental conditions that characterize the surveyed regions. Because cetaceans face a suite of anthropogenic pressures, posed by fisheries (Read, 2008), collisions (Laist *et al.*, 2001), anthropogenic noise (Weilgart, 2007), pollution (Aguilar *et al.*, 2002) and global warming (Alter *et al.*, 2010), the knowledge of such generic properties would be helpful for implementing spatially explicit conservation measures at a broad scale, such as Ecologically and Biologically Significant Marine Areas (Dunn *et al.*, 2014).

In 2008, the French Agency for Marine Protected Areas launched a large programme aimed at providing marine conservation agencies with baseline information on the distributions of megafauna throughout the tropical waters of the French EEZ. As part of this programme, aerial surveys were conducted in three regions of the tropical Atlantic, Indian and Pacific oceans (Mannocci *et al.*, 2013, 2014a,b). In previous studies, we fitted habitat models to each tropical region and predicted the densities of three cetacean guilds characterized by increasing costs of living. We found that the Delphininae and Globicephalinae had to choose the highest-productivity habitats to meet their high energetic requirements, whereas sperm and beaked whales satisfied their needs by exploiting habitats of either high or low productivity (Mannocci *et al.*, 2013, 2014a,b). In this study, we merged the three regional datasets and fitted a generic habitat model to each cetacean guild. We first searched for generic properties of cetacean habitats in tropical waters – cetacean–habitat relationships that would be valid within the full range of environmental conditions encountered in the surveyed regions of all three tropical oceans. We then carefully extrapolated cetacean densities over a large circumtropical belt, excluding areas where environmental characteristics departed from those of the surveyed regions. Our extrapolations encompass tropical pelagic areas in which no cetacean survey has ever been conducted and represent a relevant analytical solution for filling the data-gaps. Validating these extrapolations would be the next important step when more data become available in these poorly known areas.

MATERIALS AND METHODS

Cetacean data

Cetacean data were obtained from aerial surveys in the EEZ of three tropical regions of French overseas territories: French Guiana in the western tropical Atlantic (132,000 km^2, surveyed in October 2008) (Mannocci *et al.*, 2013), the central South Pacific (1.75 million km^2, surveyed from January–May 2011) (Mannocci *et al.*, 2014a) and the south-western Indian Ocean (1.4 million km^2, surveyed from December 2009 to April 2010) (Mannocci *et al.*, 2014b) (Fig. 1). Aerial surveys were conducted in the summer to avoid the trade-wind season and hence to benefit from the best detection conditions. The aerial protocol was consistent across the three tropical regions and followed the general SCANS methodology developed for small cetaceans and adapted to aircraft (Hammond *et al.*, 2013). Cetacean sightings were collected according to the distance-sampling protocol (Buckland *et al.*, 2001). The recorded information included group size, an identification to the lowest possible taxonomic level, and the angle of declination to the group when it passed at right angles to the aircraft. Together with the altitude of the aircraft, these angles provide perpendicular distances, which allowed us to model the decrease in detection probability with distance from the transect line (Buckland *et al.*, 2001). Glare severity, turbidity, cloud coverage, the state of the sea on the Beaufort scale, and an overall subjective assessment of detection conditions (good, moderate or poor, relative to the detection of delphinids) were recorded as detection covariates. Details of the aerial protocol can be found in Mannocci *et al.* (2013, 2014a,b).

Within each tropical region, we studied three cetacean guilds characterized by increasing costs of living (sperm and beaked whales, Globicephalinae and Delphininae), as indicated by their muscle mitochondrial densities and lipid contents (Spitz *et al.*, 2012) and inferred from available information on their dive patterns (mostly active, short, shallow dives *versus* mostly passive, long, deep dives). Delphininae includes all species of the genera *Stenella*, *Delphinus*, *Sousa*, *Sotalia*, *Steno*, *Lagenodelphis* and *Tursiops*. Globicephalinae includes the genera *Globicephala*, *Orca*, *Pseudorca*, *Feresa*, *Grampus* and *Peponocephala*. Finally, the sperm and beaked whales comprise all the members of the families Physeteridae, Ziphiidae and Kogiidae (for details, see Mannocci *et al.*, 2013, 2014a,b). Nomenclature follows the World Cetacea database (Perrin, 2015)

The aerial surveys covered 7149 km in French Guiana, 83,737 km in the south-western Indian Ocean and 97,277 km in the central South Pacific. Overall, Delphininae was the most frequently seen guild (839 sightings, 12,137 individuals), followed by sperm and beaked whales (252 sightings, 490 individuals) and Globicephalinae (219 sightings, 8835 individuals).

Environmental data

We described the quality of pelagic habitats using 14 physiographic, physical, chemical and biological variables, derived from a combination of satellite and *in situ* measurements of surface, sea-floor and water-column properties. These variables are indicators of primary production and/or prey aggregation and constitute potential predictors of cetacean densities (Table 1). We used the climatology of the survey

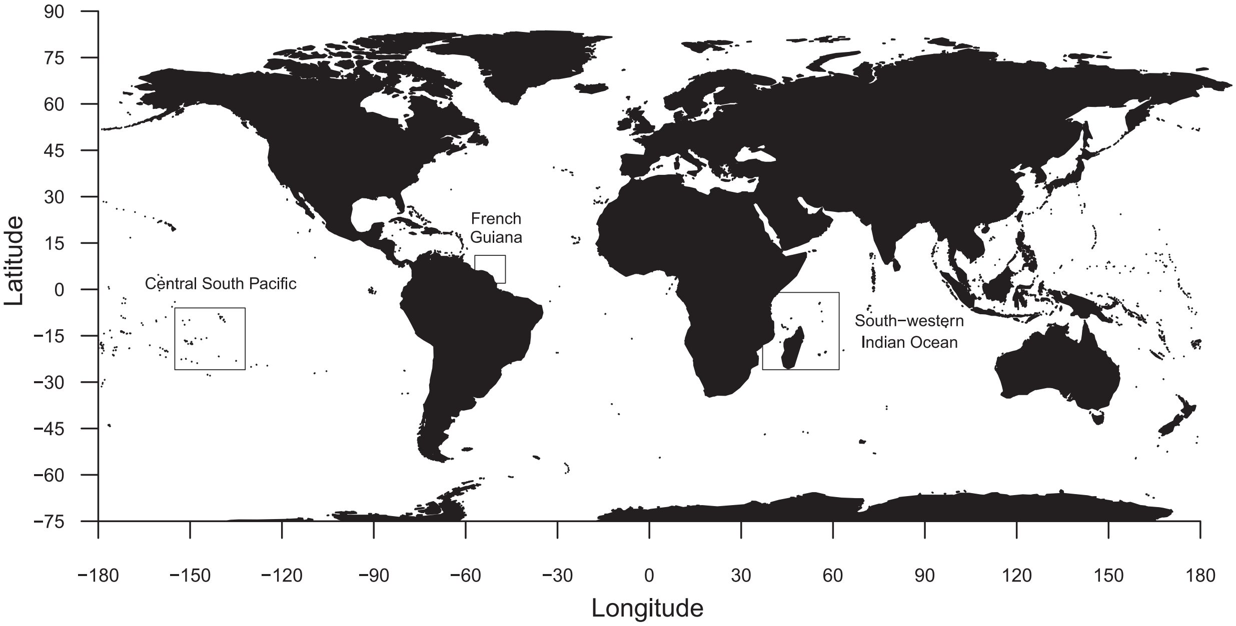

Figure 1 Locations of the three tropical surveyed regions: the central South Pacific, French Guiana and the south-western Indian Ocean. More details on the aerial surveys in each region can be found in Mannocci *et al.* (2013, 2014a,b).

Table 1 Candidate environmental predictors of cetacean habitat models.

Environmental predictors	Data source*	Underlying oceanographical or ecological process
Physiographic		
Depth (m)	1	Shallow waters are associated with higher primary production.
Slope (°)	1	High slope is associated with enhanced primary production and/or prey aggregation.
Physical		
Wind speed (m s^{-1})	2	High wind speed increases the depth of the mixed layer, which results in enhanced primary production.
Wind stress curl (N $m^{-2}/10^{-4}$ km)	2	Positive wind stress curl lifts the thermocline, which results in enhanced primary production.
Standard error of sea level anomaly (cm)	3	High standard error of sea level anomaly indicates a high mesoscale activity, enhancing primary production and/or prey aggregation.
Temperature (0–100 m) (°C)	4	Temperature influences the structure of the upper water column and has physiological effects on cetacean prey.
Surface salinity (PSU)	4	Low surface salinity indicates the proximity of a large river, associated with nutrient inputs and enhanced primary production.
Mixed layer depth (m)	4	Deep mixed layer is associated with nutrient inputs in the upper layer, enhancing primary production.
Chemical		
Silicate : nitrate ratio (0–100 m)	4	A silicate : nitrate ratio greater than 1:1 favours the growth of diatoms which leads to high fish abundance, supporting top predators.
Silicate : nitrate ratio (100–400 m)	4	A silicate : nitrate ratio greater than 1:1 favours the growth of diatoms which leads to high fish abundance, supporting top predators.
Depth of the minimum dissolved oxygen concentration (m)	4	A shallow depth of the minimum dissolved oxygen concentration reflects decomposition processes following high productivity and has physiological effects on cetacean prey (causing lethargic behaviour).
Biological		
Chlorophyll-*a* concentration (mg m^{-3})	5	High chlorophyll-*a* concentration indicates a high biomass of phytoplankton.
Net primary production (mg C m^{-3} d^{-1})	6	High net primary production indicates a high rate of primary production.
Dominant phytoplankton group	7	The dominant phytoplankton group informs the nature of primary production.

*1, General Bathymetric Chart of the Oceans (http://www.gebco.net/); 2, Scatterometer Climatology of Ocean Winds (Risien & Chelton, 2008); 3, AVISO (http://www.aviso.oceanobs.com/duacs/); 4, CSIRO Atlas of Regional Seas (Condie & Dunn, 2006); 5, Ocean Color (http://oceancolor.gsfc.nasa.gov/); 6, Ocean productivity (http://www.science.oregonstate.edu/ocean.productivity/); 7, PHYSAT model (Ben Mustapha *et al.*, 2013).

periods as our temporal resolution because it has been shown that climatological variables (averaged over several years) yield higher explained deviances than the same variables measured at the time of individual surveys (Mannocci *et al.*, 2014b). In addition, the extrapolation beyond surveyed regions would have made the selection of a contemporaneous resolution difficult. This use of climatologies is in line with our goal of identifying generic cetacean–habitat relationships. Our selection of time periods and spatial resolutions for environmental variables was primarily based on data availability. Time periods were nevertheless overlapping and their differences were smoothed out by the use of climatologies. The differences in spatial resolution were smoothed out by subsequent spatial aggregation. Our physiographical variables were depth and slope.

Bathymetric data were obtained from the General Bathymetric Chart of the Ocean (GEBCO; available at: http://www.gebco.net/), providing a one-minute global bathymetric surface. Slope was derived from GEBCO using the RASTER package in R (Hijmans *et al.*, 2012).

A broad set of physical variables were included. Monthly wind speed and wind stress curl were obtained at 0.25° spatial resolution from the Scatterometer Climatology of Ocean Winds, averaged over 8 years (1999–2007) of QuikSCAT observations (Risien & Chelton, 2008). The standard error of sea level anomaly (SLA) was used as an indicator of mesoscale activity (Mannocci *et al.*, 2014b). The standard errors of SLAs were calculated for the 2002–2012 period based on weekly SLAs at 0.25° resolution from AVISO (http://www.aviso.oceanobs.com/duacs/). Temperature in the upper water column (using 100 m as an approximation of the euphotic depth) and surface salinity were extracted from 2009 CSIRO Atlas of Regional Seas (CARS) at 0.5° resolution (Condie & Dunn, 2006). CARS is derived from historical archives of subsurface ocean property measurements (primarily research vessel instrument profiles and autonomous profiling buoys). It is created by interpolating the available oceanographical profile data, most of which were collected in the last 50 years. Mixed-layer depth was calculated from salinity and temperature profiles, as reported in Condie & Dunn (2006).

Chemical variables were also retrieved from CARS. The silicate: nitrate ratio ('Redfield ratio'; Redfield, 1958) was calculated in (1) the euphotic layer (0–100 m), and (2) a deeper layer (100–400 m), approximating the layer where nutrients are potentially brought to the surface by mesoscale eddies (Menkes *et al.*, 2002). The depth of the minimum dissolved oxygen concentration was also calculated.

Finally, three biological variables were considered. Remotely sensed chlorophyll-*a* concentration (CHL) was obtained at 9-km and monthly resolution from the MODIS sensor (http://oceancolor.gsfc.nasa.gov/) for the period 2002–2012. Monthly net primary production (NPP) was retrieved at 9-km resolution from Ocean Productivity (http://www.science.oregonstate.edu/ocean.productivity/) for 2002–2012, based on the vertically generalized production model (Behrenfeld & Falkowski, 1997). In addition, we used the monthly climatology of dominant phytoplankton groups derived from the PHYSAT model at 9-km resolution (categorical variable). This algorithm identified six phytoplankton groups (nanoeukaryotes, *Prochlorococcus*, *Synechococcus*, diatoms, coccolithophores and *Phaeocystis*-like) from the emission spectra measured by the Sea-Viewing Wide Field-of-View Sensor between 1997 and 2010 (Ben Mustapha *et al.*, 2013). PHYSAT products have been validated using *in situ* pigment data from a wide range of oceanographical conditions.

Rasters of all variables were aggregated to match the same 0.5° latitude × 0.5° longitude grid using the SDMTools (VanDerWal *et al.*, 2012) and raster (Hijmans *et al.*, 2012) packages in R.

Detection-function modelling

Detection functions were based on cetacean sightings pooled across all three surveyed regions. We used multiple covariate distance sampling (Marques & Buckland, 2004) to model the effect of detection covariates on detection probability, in addition to distance. If the covariate provided a significantly smaller Akaike information criterion (AIC) score (difference greater than 2), we estimated an effective strip width (ESW) for each of the covariate levels; if not, we estimated a unique ESW from a detection function without any covariate.

Data processing

We aggregated cetacean sightings and survey effort into grid cells at the spatial resolution of the environmental variables (0.5° × 0.5°) using ArcGIS 10 (ESRI, Redlands, CA, USA). The numbers of individuals and survey effort were summed in each grid cell. For each guild, the mean ESW was calculated as the arithmetic mean of ESWs in each grid cell.

Habitat modelling

We built generalized additive models (GAMs) to relate the numbers of cetaceans per grid cell to environmental predictors. GAMs constitute a flexible regression technique in which the inclusion of splines generates smooth functions of the covariates (Wood, 2006). A link function, $g()$, relates the mean of the response variable given the covariates, $\mu = \mathrm{E}(N \mid X_1, \ldots, X_p)$, to the additive predictor $\alpha + \sum f_i(X_i)$:

$$g(\mu) = \alpha + \sum f_i(X_i),$$

where f_i is a nonparametric smoothing function of covariate i, and α is an intercept term.

We used the Tweedie distribution to model the number of individuals (our response variable). Its three parameters (mean, dispersion and power) provide additional flexibility to model the high proportion of zeros that characterize species count data (Miller *et al.*, 2013). To account for non-constant effort, we used the logarithm of sampled area as the offset (survey effort multiplied by twice the mean ESW in each grid cell, as estimated above). GAMs were fitted using the R package mgcv 1.8-4, in which degrees of freedom are dictated by the smoothing parameters estimated along with the parameter values of smoothing functions during model fitting, and thin-plate regression splines are the default basis functions (Wood, 2006). We adopted restricted maximum likelihood as the smoothness selection criteria, which has been shown to penalize overfitting and lead to more pronounced optima (Wood, 2011). We set an upper limit of four on the degrees of freedom for each spline, because upper limits of 5–10 degrees of freedom led to negligible differences in the resulting functions. This appeared to be enough to derive ecologically sound relationships, but without leading to overfitted functions that would not extrapolate well outside the study regions.

For each cetacean guild, we fitted models containing all possible combinations of four environmental variables, excluding combinations of collinear variables (those with pairwise Spearman rank correlation coefficients > 0.6 or < -0.6). Four predictors were considered in the model to avoid excessive complexity while providing reasonable explanatory power. Indeed, the use of parsimonious models was in line with our objective of extrapolating beyond surveyed regions. The model with the lowest AIC was selected (Wood, 2006).

We predicted the relative densities of Delphininae, Globicephalinae and sperm and beaked whales in each grid cell, based on the climatology of the summer in each hemisphere (20 December–20 March in the Southern Hemisphere; 20 June–20 September in the Northern Hemisphere), because all aerial surveys were conducted during summer. We limited the predictions to a circumtropical envelope encompassing the range of environmental conditions encountered in the three surveyed regions (Elith & Leathwick, 2009). The prediction envelope was first bounded by the range of water temperatures and then by the range of the other environmental predictors of the model.

The probability of detecting an animal on the transect line, $g(0)$, is affected both by availability bias (when animals are submerged) and by perception bias (when observers fail to detect animals at the surface) (Pollock *et al.*, 2006). In aerial surveys, a higher proportion of animals is expected to be missed due to availability bias. We therefore applied correction factors based on diving data to account for the availability-bias portion of $g(0)$. Sperm and beaked whales spend on average 13% of their time at the surface (Papastavrou *et al.*, 1989; Barlow *et al.*, 1997; Hooker *et al.*, 2012), Globicephalinae 70% (Barlow *et al.*, 1997; Hooker *et al.*, 2012; Minamikawa *et al.*, 2013) and Delphininae 74% (Corkeron & Martin, 2004; Gómez de Segura *et al.*, 2006). Although these correction factors allowed standardization between guilds, the predicted densities may still be underestimates because there was no correction for perception bias.

We documented the uncertainty around the predictions by mapping coefficients of variation. Uncertainty was derived from the Bayesian covariance matrix of the model coefficients within the MGCV package (Wood, 2006).

To evaluate model fit, we provided diagnostic plots of deviance residuals, as well as randomized quantile residuals, which are better adapted to count models and easier to interpret than deviance residuals. They are obtained by inverting the fitted distribution for each response value and finding the equivalent standard normal quantile. A randomization procedure allows continuous residuals to be computed when the response is discrete (Dunn & Smyth, 1996).

RESULTS

Detection models

The best-fitting detection model was a half-normal distribution (a fold at the mean of an ordinary normal distribution with mean zero) for the three cetacean guilds (see Appendix S1 in Supporting Information). Cloud coverage significantly affected the detection of Delphininae. Sea state significantly affected the detection of sperm and beaked whales. No covariate had a significant effect on the detection of Globicephalinae.

Generic cetacean–habitat relationships in the tropics

The model results highlighted generic properties of cetacean habitats in tropical waters (Fig. 2). The explained deviances were 20% for sperm and beaked whales, 37% for Globicephalinae, and 45% for Delphininae. All the relationships were significant. Delphininae showed a generally increasing relationship with NPP, a linearly decreasing relationship with mixed-layer depth and a unimodal relationship with depth of the minimum dissolved oxygen concentration (DepMinOx; optimum around 700 m). The dominant phytoplankton group was also selected; densities of Delphininae were highest when nanoeukaryotes were the dominant phytoplankton group. Globicephalinae exhibited a unimodal relationship with CHL (optimum around 0.6 mg m^{-3}), a linearly decreasing relationship with wind speed and an increasing relationship with DepMinOx. Their densities also increased with slope. Sperm and beaked whales showed an increasing relationship with slope, a unimodal relationship with CHL (optimum around 0.5 mg m^{-3}), a linearly increasing relationship with DepMinOx and a slightly decreasing relationship with surface salinity.

Prediction envelopes

Water temperatures ranged from 22.8 to 28.8 °C in the study regions. The prediction envelopes therefore excluded the Pacific warm pool and the cold upwelling regions situated in the western margins of South America and Africa, as well as the equatorial Pacific. For each cetacean guild, the prediction envelope was constrained by water temperature, and was then restricted to the minimum extent envelope that encompassed all grid cells in which the four environmental predictors were within the range of values encountered in surveyed regions (see Appendix S2). Using quantiles (e.g. the 5% and 95% quantiles) of the covariates instead of the range would result in narrower environmental envelopes. For the three guilds, the prediction envelope excluded some areas of the central Atlantic, western Pacific and eastern Indian Ocean due to missing values of DepMinOx. For Delphininae, the prediction envelope excluded areas of the central Pacific where the mixed layer was deeper than in the study regions, as well as the north of the Arabian Sea, where there was no dominant phytoplankton group. For Globicephalinae, it excluded areas where wind speed was higher than in the study regions, notably the Arabian Sea and the Bay of Bengal. For sperm and beaked whales, the prediction envelope excluded areas of high (e.g. the offshore Atlantic) and low surface salinity (e.g. the Bay of Bengal).

Circumtropical extrapolations of cetacean densities

After standardizing by surface availability, the predicted densities were much lower for sperm and beaked whales than for Delphininae and Globicephalinae, which showed predicted densities of the same order of magnitude. The predicted densities of Delphininae were contrasted between ocean basins (Fig. 3a). They peaked in the western Pacific – notably in the South China Sea and Banda Sea – the equatorial Indian Ocean and the equatorial Atlantic. They were intermediate to high in the western North Atlantic, Gulf of Mexico and some areas of the eastern tropical Pacific. Conversely, the predicted densities were depressed in the central parts of ocean basins (notably in the South Pacific subtropical gyre). Globicephalinae also exhibited differences between ocean basins (Fig. 3b). Their predicted densities were the highest in the equatorial Indian Ocean, Mozambique Channel and western Pacific and the lowest in the central North and South Atlantic, as well as in the South Pacific subtropical gyre. For sperm and beaked whales, the more homogeneous prediction map indicated less contrasting densities between and within ocean basins (Fig. 3c). Densities were predicted to be the highest in the western Pacific (e.g. the South China Sea), north-eastern Indian Ocean (e.g. the Andaman Sea) and in some areas of the eastern tropical Pacific.

Uncertainty maps

In general, coefficients of variation (CVs) were higher in areas of low predicted density and at the edges of environ-

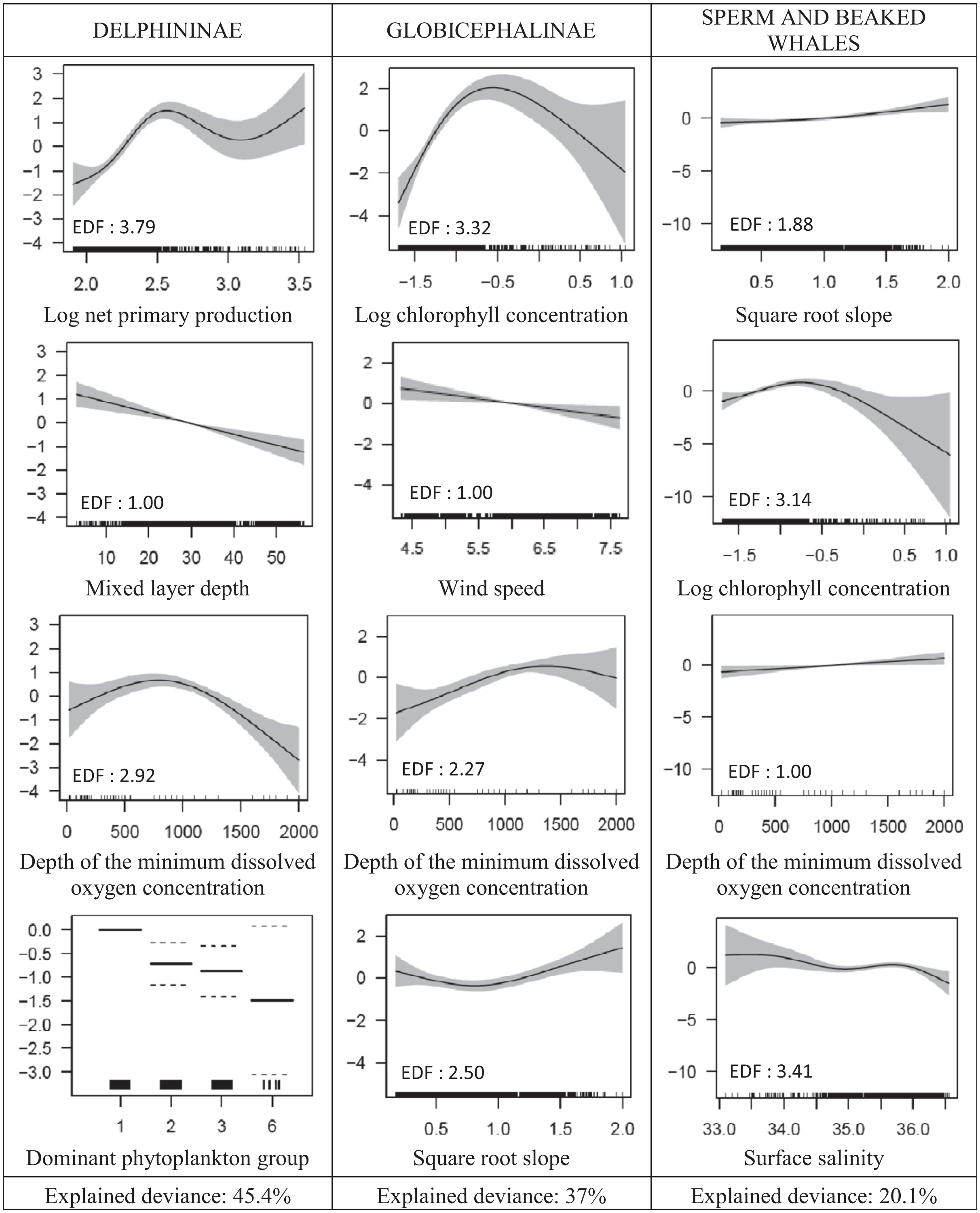

Figure 2 Cetacean–habitat relationships for the three cetacean guilds. Smooth terms are ranked according to decreasing *F*-scores (smooth terms with higher *F*-scores at the top). The *x*-axis represents the environmental predictor. Chlorophyll-*a* concentration and net primary production were log-transformed, slope was square-root-transformed to achieve a more even spread of data and reduce the chance of a few points being over influential on the data, as recommended by Wood (2006). All predictors were continuous except the dominant phytoplankton group (1, nanoeukaryotes; 2, *Prochlorococcus*; 3, *Synechococcus*; 6, diatoms) which was categorical (no smooth term). The *y*-axis represents the response variable in log scale. Estimated degrees of freedom (EDF) are provided for each smooth function. The solid line is the smooth function estimate and shaded regions represent approximate 95% confidence intervals. Zero on the *y*-axis indicates no effect of the environmental predictor on cetacean density (given that the other predictors are included in the model).

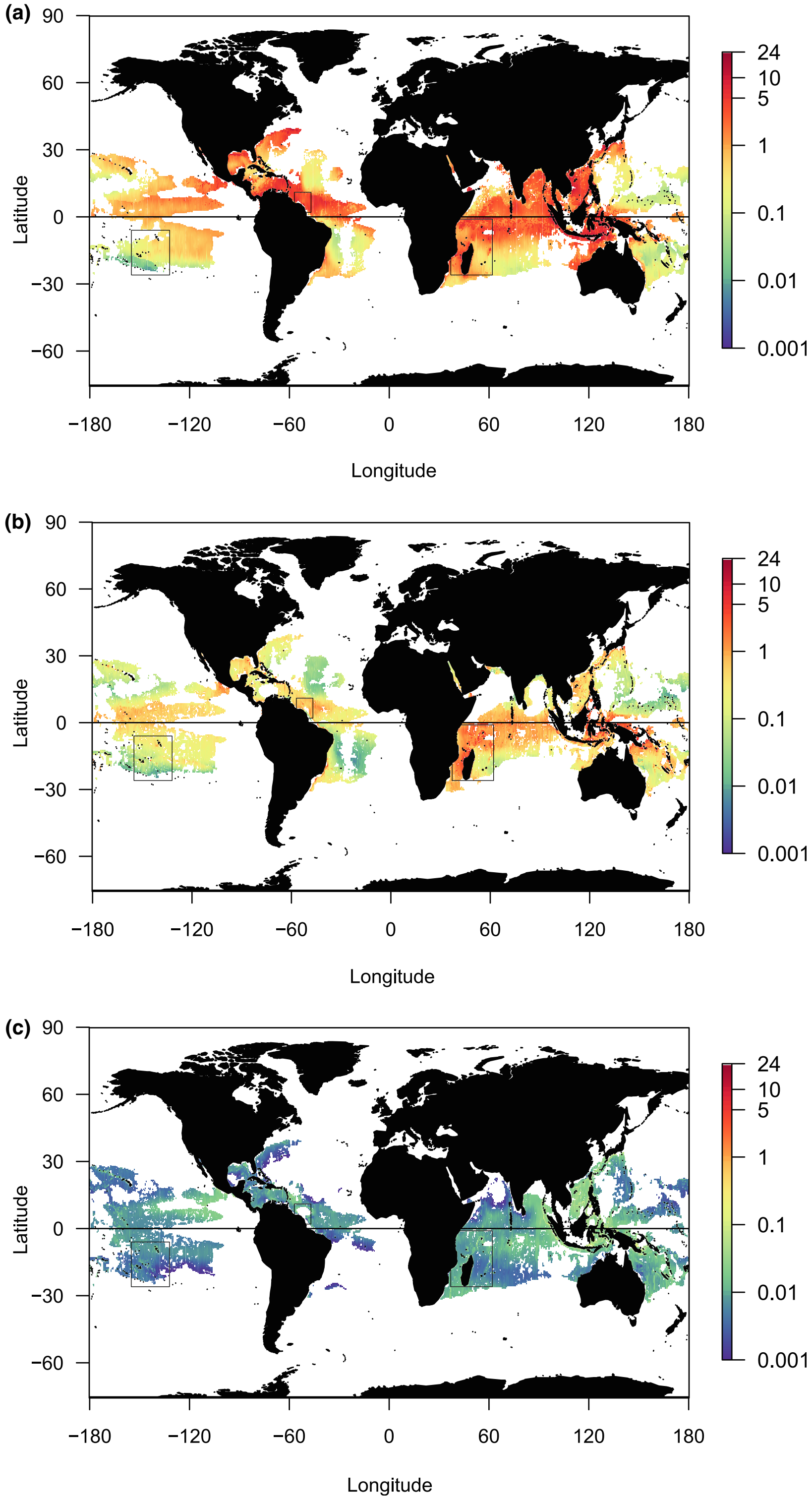

Figure 3 Predicted relative densities (individuals km^{-2}) of the three cetacean guilds: (a) Delphininae, (b) Globicephalinae, and (c) sperm and beaked whales. The three maps have the same colour scale. The models were fitted from cetacean sightings collected in all three tropical study regions (rectangles). Predictions were provided for the summer of each hemisphere since all surveys were conducted in the summer season. We standardized predicted densities between guilds based on their availability at the surface.

mental envelopes. For Delphininae, CVs were highest in the south-western and central south Pacific, where DepMinOx was high, as well as in areas of high NPP such as the Great Bahama Bank and the Gulf of Thailand (Fig. 4a). For Globicephalinae, CVs were highest in the high-CHL waters of French Guiana and the northern Arabian Sea (Fig. 4b) and at the edges of the DepMinOx envelope. For sperm and beaked whales, CVs peaked at the edge of the salinity envelope, notably in the Arabian Sea (high salinity) and in the South China Sea (low salinity) (Fig. 4c).

Model checking

For the three guilds, the deviance residuals reflected overdispersion in our data (indicated by the heavy-tailed histograms), but did not show any problematic patterns (see Appendix S3). The randomized quantile residuals showed no unusual patterns, indicating that our assumptions of a Tweedie distribution and log-link function were appropriate.

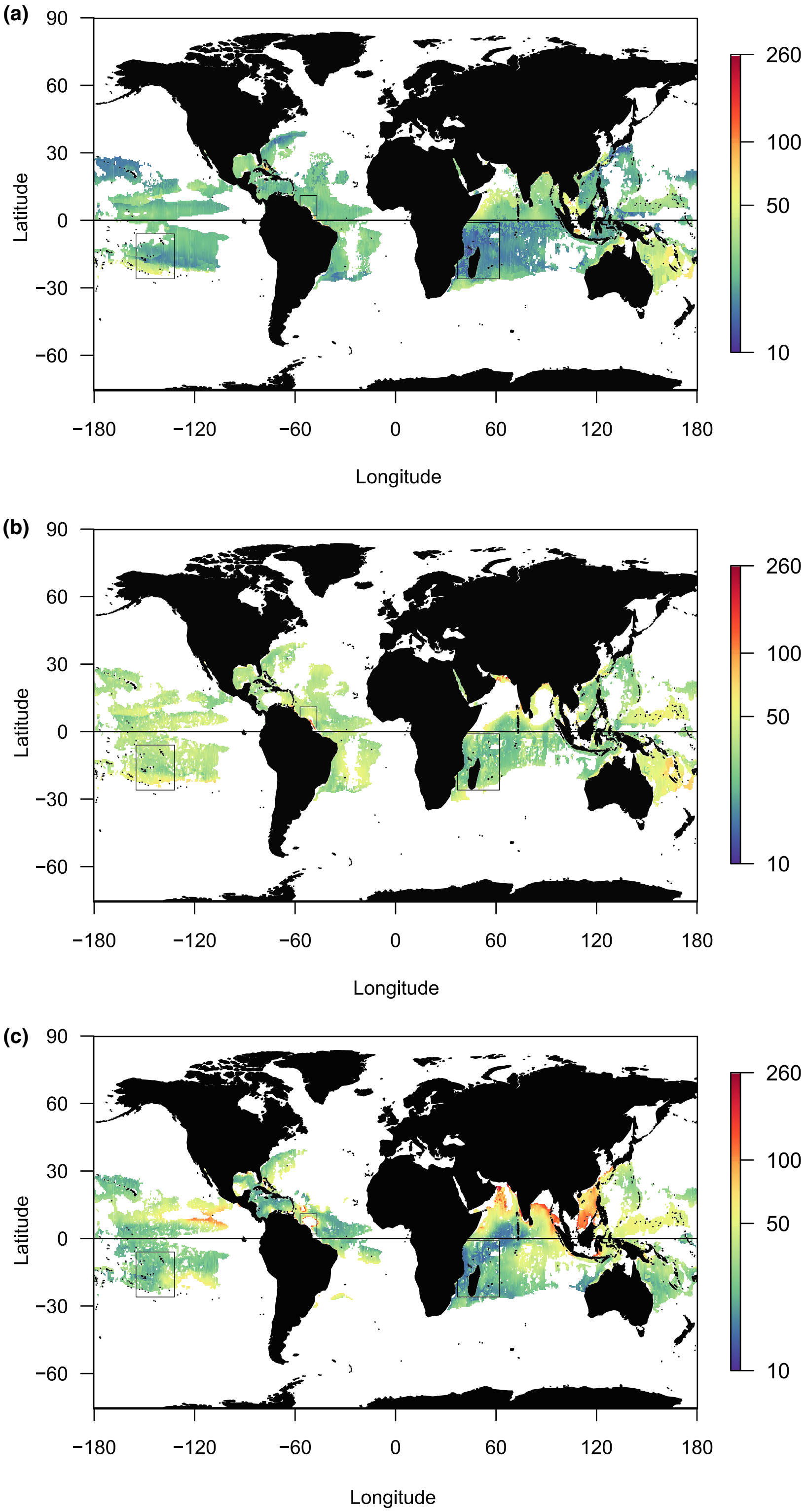

Figure 4 Coefficients of variation (%) associated with the predicted densities of the three cetacean guilds: (a) Delphininae, (b) Globicephalinae, and (c) sperm and beaked whales. The three maps have the same colour scale.

DISCUSSION

In this study, we extrapolated cetacean densities over a vast circumtropical belt that encompasses the environmental characteristics covered in three regions of the Atlantic, Pacific and Indian oceans. Our approach stands out from other studies that have predicted cetacean habitats at a global scale. Kaschner *et al.* (2006) relied on expert knowledge about the relationships between species presence and basic environmental parameters to build environmental suitability models. The resulting global maps described the relative environmental suitability for different species, which has been shown to correlate well with relative species occurrence, but do not represent predictions of species densities. In contrast, we relied on generic cetacean–habitat relationships derived from sightings collected in three contrasting tropical regions and meaningful environmental predictors to build density extrapolations.

Generic cetacean–habitat relationships in the tropics

We found cetacean relationships with a variety of indicators of productivity and/or prey aggregation (Table 1) and qualified them as generic because they emerged from all three tropical regions. Some of these relationships have been documented before, but others are presented here for the first time. For example, the three cetacean guilds exhibited generic relationships with the quantity of primary production (with either chlorophyll-*a* concentration or net primary production as predictors). This has been documented in various tropical and non-tropical systems (Jaquet & Whitehead, 1996; Ferguson *et al.*, 2006a,b; Becker *et al.*, 2010). Interestingly, we found a generic relationship of Delphininae with the dominant phytoplankton group, which characterizes the nature of primary production. Food chains based on large phytoplankton groups like diatoms, whose growth is favoured when the silicate: nitrate ratio is higher than 1:1, lead to high fish abundances, typically supporting top predators (Parsons & Lalli, 2002). We found that Delphininae were mostly associated with the small phytoplankton groups that are dominant in the tropics (Ben Mustapha *et al.*, 2013). We also documented for the first time generic relationships of cetaceans with the depth of the minimum dissolved oxygen concentration. At low concentrations of dissolved oxygen, squid and fish tend to reduce their metabolism and adopt lethargic behaviours (Seibel, 2011). This may facilitate their capture by deep-diving cetaceans and limit competition with other large predators like tuna, whose vertical distribution can be limited by dissolved oxygen concentration (Prince & Goodyear, 2006).

Even though the explained deviances ranged from 20% to 45%, which are good percentages for cetacean habitat modelling studies, a significant part of the cetacean distributions remained unexplained. The fine-scale physical and biological processes that cause prey to aggregate can vary between ecosystems (Benson *et al.*, 2011), and this may account for part of the unexplained deviance. Although we succeeded in finding generic cetacean–habitat relationships in tropical waters, our approach was not appropriate for capturing ecosystem-specific processes. Our extrapolations relied on the assumption that cetacean–habitat relationships would hold outside the study regions. They may be less reliable in the areas situated furthest from our study regions, where forcing mechanisms might be different.

Qualitative evaluation of circumtropical predictions

In order to evaluate the genericity of a model, it is recommended that external validation be performed using independent data collected in distinct regions (Fielding & Bell, 1997). Our aerial surveys are the only dedicated large-scale surveys in tropical waters, apart from the comprehensive US line-transect surveys conducted in the Gulf of Mexico (GOM), south-eastern United States (SE) and eastern tropical Pacific (ETP), areas which fall only marginally within our circumtropical prediction envelopes. This severely limits the potential for real external validation. Until they can be validated by independent datasets collected at the core of the circumtropical envelope, the extrapolations in the furthest areas remain tentative.

Nevertheless, we made visual comparisons of our extrapolations with model outputs available in the well-documented US waters. The high densities of Delphininae we predicted on the continental shelf of the GOM correspond fairly well to high-suitability areas previously identified for *Stenella* and *Tursiops* (Best *et al.*, 2012), the main genera in this guild. We also predicted intermediate densities of Delphininae on the SE continental shelf, where coastal waters are highly suitable for *Tursiops* (Best *et al.*, 2012). We predicted intermediate densities of Globicephalinae on the SE shelf edge, in accordance with high-suitability areas for pilot whales (Best *et al.*, 2012). Our intermediate densities of sperm and beaked whales in the oceanic waters of the GOM also appeared consistent with the results of Best *et al.* (2012).

Our prediction envelopes in the ETP encompassed the eastern Pacific warm pool and the countercurrent thermocline ridge, but excluded the Costa Rica dome and equatorial front, two areas of enhanced productivity that are important to many cetaceans (Ballance *et al.*, 2006). The countercurrent appeared as a high-density area for the spotted dolphin (*Stenella attenuata*) and the spinner dolphin (*Stenella longirostris*), and the eastern warm pool appeared as a high-density area for the spinner dolphin and the rough-toothed dolphin, *Steno bredanensis* (Redfern *et al.*, 2008; Forney *et al.*, 2012). This seems to be consistent with our intermediate predicted densities for Delphininae. However, the densities of the short-finned pilot whale (*Globicephala macrorhynchus*) and Risso's dolphin (*Grampus griseus*) were highest outside our prediction envelope. Forney *et al.* (2012) predicted high densities of the dwarf sperm whale (*Kogia sima*) and Cuvier's beaked whale (*Ziphius cavirostris*) in the eastern warm pool, as well as *Mesoplodon* spp. in the countercurrent. Furthermore, catch data in the Pacific indicated that sperm whales were abundant in equatorial waters (Jaquet *et al.*, 1996). This seems to accord with our predictions for sperm and beaked whales.

Conclusion and perspectives

Because of a very uneven allocation of survey effort, our knowledge of cetacean distributions and densities is hindered by large data-gaps (Kaschner *et al.*, 2012). Until new surveys are carried out, our modelling approach appears to be a good analytical solution for filling these gaps in tropical waters where data are scarce or non-existent. We extrapolated cetacean densities with care, limiting the extent of our predictions to a circumtropical envelope in which environmental conditions resembled those encountered in the study regions. We highlighted generic cetacean–habitat relationships in the tropics, notably with primary production and the depth of the minimum dissolved oxygen concentration. Circumtropical extrapolations for Delphininae and Globicephalinae revealed contrasts between ocean basins, with highest predicted densities in equatorial waters of the three ocean basins, whereas the predicted densities for sperm and beaked whales were comparatively lower and more uniform within the circumtropical belt than for the other two guilds.

The results of this study could be used in making recommendations on future data collection. Although we surveyed very different tropical regions, characterized by inherent environmental heterogeneity, most of the surveyed ecosystems were oligotrophic (with the exception of the coastal waters of French Guiana and Madagascar). In order to cover a broader range of tropical ecosystems, driven by various forcing mechanisms, additional surveys should be implemented in productive ecosystems, such as the Bay of Bengal and the South China Sea. This would improve the genericity of our habitat models and increase the extent of the extrapolation envelopes. Additional data should also be collected in the areas situated at the edges of environmental envelopes. This would provide a more even coverage of covariate values and reduce uncertainty at the edges. Finally, the collection of additional data at the core of the circumtropical belt remains crucial for validating the extrapolations with independent datasets. In this respect, the use of shipping lines as transoceanic platforms of opportunity appears to be promising, because it would allow new data to be acquired in remote oceanic regions at low cost.

In the context of identifying Ecologically or Biologically Significant Marine Areas (EBSAs), the parties to the Convention on Biological Diversity need to propose ocean areas that may require enhanced conservation and management measures, both within and beyond waters under national jurisdiction (Dunn *et al.*, 2014). These EBSAs must be based on the best available knowledge regarding the distributions of major components of marine biodiversity. Cetaceans are important both inherently and because they are considered indicators of less observable components of marine biodiversity. By providing predictions of cetacean hotspots at a global scale, in waters far beyond countries' EEZs, our results can help guide the delimitation of EBSAs and large-scale networks of marine protected areas.

ACKNOWLEDGEMENTS

The aerial surveys were funded by the French ministry of the environment (Ministère de l'Écologie, du Développement Durable et de l'Énergie) and the agency for marine protected areas (Agence des aires marines protégées et parcs naturels marins). We thank: Olivier Van Canneyt, Ghislain Doremus and Sophie Laran, who coordinated the surveys; all the observers who collected data in French Guiana, the southwestern Indian Ocean and Polynesia; Severine Alvain for providing the PHYSAT products (License: IDDN. FR.001.330003.000.S.P.2012.000.30300); and Lisa Ballance and Phil Hammond for their comments on a previous version of the manuscript. This study benefited from discussion during meetings of the PELAGIC project funded by the *Fondation pour la Recherche sur la Biodiversité* and *Fondation Total*, and hosted by the Centre for the Synthesis and Analysis of Biodiversity. Finally, we would like to thank David L. Miller and two anonymous referees for their constructive reviews.

REFERENCES

Aguilar, A., Borrell, A. & Reijnders, P.J.H. (2002) Geographical and temporal variation in levels of organochlorine contaminants in marine mammals. *Marine Environmental Research*, **53**, 425–452.

Alter, S.E., Simmonds, M.P. & Brandon, J.R. (2010) Forecasting the consequences of climate-driven shifts in human behavior on cetaceans. *Marine Policy*, **34**, 943–954.

Ballance, L.T., Pitman, R.L. & Fiedler, P.C. (2006) Oceanographic influences on seabirds and cetaceans of the eastern tropical Pacific: a review. *Progress in Oceanography*, **69**, 360–390.

Barlow, J., Forney, K.A., Von Saunder, A. & Urban-Ramirez, J. (1997) *A report of cetacean acoustic detection and dive interval studies (CADDIS) conducted in the southern Gulf of California, 1995*. NOAA Technical Memorandum NMFS NOAA-TM-NMFS-SWFSC-25. US Department of Commerce, Washington, DC.

Becker, E.A., Forney, K.A., Ferguson, M.C., Foley, D.G., Smith, R.C., Barlow, J. & Redfern, J.V. (2010) Comparing California Current cetacean–habitat models developed using *in situ* and remotely sensed sea surface temperature data. *Marine Ecology Progress Series*, **413**, 163–183.

Becker, E.A., Forney, K.A., Foley, D.G., Smith, R.C., Moore, T.J. & Barlow, J. (2014) Predicting seasonal density patterns of California cetaceans based on habitat models. *Endangered Species Research*, **23**, 1–22.

Behrenfeld, M.J. & Falkowski, P.G. (1997) Photosynthetic rates derived from satellite-based chlorophyll concentration. *Limnology and Oceanography*, **42**, 1–20.

Ben Mustapha, Z., Alvain, S., Jamet, C., Loisel, H. & Dessailly, D. (2013) Automatic classification of water-leaving radiance anomalies from global SeaWiFS imagery: application to the detection of phytoplankton groups in open ocean waters. *Remote Sensing of Environment*, **146**, 97–112.

Benson, S.R., Eguchi, T., Foley, D.G., Forney, K.A., Bailey, H., Hitipeuw, C., Samber, B.P., Tapilatu, R.F., Rei, V., Ramohia, P., Pita, J. & Dutton, P.H. (2011) Large-scale movements and high-use areas of western Pacific leatherback turtles, *Dermochelys coriacea*. *Ecosphere*, **2**, art84.

Best, B.D., Halpin, P.N., Read, A.J., Fujioka, E., Good, C.P., LaBrecque, E.A., Schick, R.S., Roberts, J.J., Hazen, L.J., Qian, S.S., Palka, D.L., Garrison, L.P. & McLellan, W.A. (2012) Online cetacean habitat modeling system for the US east coast and Gulf of Mexico. *Endangered Species Research*, **18**, 1–15.

Buckland, S.T., Anderson, D.R., Burnham, H.P., Laake, J.L., Borchers, D.L. & Thomas, L. (2001) *Introduction to distance sampling: estimating abundance of biological populations*. Oxford University Press, Oxford.

Chassot, E., Bonhommeau, S., Reygondeau, G., Nieto, K., Polovina, J.J., Huret, M., Dulvy, N.K. & Demarcq, H. (2011) Satellite remote sensing for an ecosystem approach to fisheries management. *ICES Journal of Marine Science*, **68**, 651–666.

Condie, S.A. & Dunn, J.R. (2006) Seasonal characteristics of the surface mixed layer in the Australasian region: implications for primary production regimes and biogeography. *Marine and Freshwater Research*, **57**, 569–590.

Corkeron, P.J. & Martin, A.R. (2004) Ranging and diving behaviour of two 'offshore' bottlenose dolphins, *Tursiops* sp., off eastern Australia. *Journal of the Marine Biological Association of the UK*, **84**, 465–468.

Dunn, P.K. & Smyth, G.K. (1996) Randomized quantile residuals. *Journal of Computational and Graphical Statistics*, **5**, 236–244.

Dunn, D.C., Ardron, J., Bax, N., Bernal, P., Cleary, J., Cresswell, I., Donnelly, B., Dunstan, P., Gjerde, K., Johnson, D., Kaschner, K., Lascelles, B., Rice, J., von Nordheim, H., Wood, L. & Halpin, P.N. (2014) The Convention on Biological Diversity's Ecologically or Biologically Significant Areas: origins, development, and current status. *Marine Policy*, **49**, 137–145.

Elith, J. & Leathwick, J.R. (2009) Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 677–697.

Ferguson, M.C., Barlow, J., Fiedler, P., Reilly, S.B. & Gerrodette, T. (2006a) Spatial models of delphinid (family Delphinidae) encounter rate and group size in the eastern tropical Pacific Ocean. *Ecological Modelling*, **193**, 645–662.

Ferguson, M.C., Barlow, J., Reilly, S.B. & Gerrodette, T. (2006b) Predicting Cuvier's (*Ziphius cavirostris*) and *Mesoplodon* beaked whale population density from habitat characteristics in the eastern tropical Pacific Ocean. *Journal of Cetacean Research and Management*, **7**, 287–299.

Fielding, A.H. & Bell, J.F. (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, **24**, 38–49.

Forney, K.A., Ferguson, M.C., Becker, E.A. & Fiedler, P.C. (2012) Habitat-based spatial models of cetacean density in the eastern Pacific Ocean. *Endangered Species Research*, **16**, 113–133.

Gómez de Segura, A., Crespo, E.A., Pedraza, S.N., Hammond, P.S. & Raga, J.A. (2006) Abundance of small cetaceans in waters of the central Spanish Mediterranean. *Marine Biology*, **150**, 149–160.

Hammond, P.S., Macleod, K., Berggren, P. *et al.* (2013) Cetacean abundance and distribution in European Atlantic shelf waters to inform conservation and management. *Biological Conservation*, **164**, 107–122.

Hijmans, R., van Etten, J. & Mattiuzzi, M. *et al.* (2012) *raster: geographic data analysis and modeling*. Available at: http://cran.r-project.org/web/packages/raster/.

Hooker, S.K., Fahlman, A., Moore, M.J., de Soto, N.A., de Quirós, Y.B., Brubakk, A.O., Costa, D.P., Costidis, A.M., Dennison, S. & Falke, K.J. (2012) Deadly diving? Physiological and behavioural management of decompression stress in diving mammals. *Proceedings of the Royal Society B: Biological Sciences*, **279**, 1041–1050.

Jaquet, N. & Whitehead, H. (1996) Scale-dependent correlation of sperm whale distribution with environmental features and productivity in the South Pacific. *Marine Ecology Progress Series*, **135**, 1–9.

Jaquet, N., Whitehead, H. & Lewis, M. (1996) Coherence between 19th century sperm whale distributions and satellite-derived pigments in the tropical Pacific. *Marine Ecology Progress Series*, **145**, 1–10.

Kaschner, K., Watson, R., Trites, A.W. & Pauly, D. (2006) Mapping world-wide distributions of marine mammal species using a relative environmental suitability (RES) model. *Marine Ecology Progress Series*, **316**, 285–310.

Kaschner, K., Quick, N.J., Jewell, R., Williams, R. & Harris, C.M. (2012) Global coverage of cetacean line-transect surveys: status quo, data gaps and future challenges. *PLoS ONE*, **7**, e44075.

Laist, D.W., Knowlton, A.R., Mead, J.G., Collet, A.S. & Podesta, M. (2001) Collisions between ships and whales. *Marine Mammal Science*, **17**, 35–75.

Longhurst, A.R. & Pauly, D. (1987) *Ecology of tropical oceans*. Academic Press, Oxford.

Mannocci, L., Monestiez, P., Bolaños-Jiménez, J., Dorémus, G., Jeremie, S., Laran, S., Rinaldi, R., Van Canneyt, O. & Ridoux, V. (2013) Megavertebrate communities from two contrasting ecosystems in the western tropical Atlantic. *Journal of Marine Systems*, **111**, 208–222.

Mannocci, L., Catalogna, M., Dorémus, G., Laran, S., Lehodey, P., Massart, W., Monestiez, P., Van Canneyt, O., Watremez, P. & Ridoux, V. (2014a) Predicting cetacean and seabird habitats across a productivity gradient in the South Pacific gyre. *Progress in Oceanography*, **120**, 383–398.

Mannocci, L., Laran, S., Monestiez, P., Dorémus, G., Van Canneyt, O., Watremez, P. & Ridoux, V. (2014b) Predicting top predator habitats in the Southwest Indian Ocean. *Ecography*, **37**, 261–278.

Marques, F.F.C. & Buckland, S.T. (2004) Covariate models for the detection function. *Advanced distance sampling* (ed. by S.T. Buckland, D.R. Anderson, K.P. Burnham and L. Thomas), pp. 31–47. Oxford University Press, Oxford.

Menkes, C.E., Kennan, S.C., Flament, P., Dandonneau, Y., Masson, S., Biessy, B., Marchal, E., Eldin, G., Grelet, J., Montel, Y., Morlière, A., Lebourges-Dhaussy, A., Moulin, C., Champalbert, G. & Herbland, A. (2002) A whirling ecosystem in the equatorial Atlantic. *Geophysical Research Letters*, **29**, 48-1–48-4.

Miller, D.L., Burt, M.L., Rexstad, E.A. & Thomas, L. (2013) Spatial models for distance sampling data: recent developments and future directions. *Methods in Ecology and Evolution*, **4**, 1001–1010.

Minamikawa, S., Watanabe, H. & Iwasaki, T. (2013) Diving behavior of a false killer whale, *Pseudorca crassidens*, in the Kuroshio-Oyashio transition region and the Kuroshio front region of the western North Pacific. *Marine Mammal Science*, **29**, 177–185.

Papastavrou, V., Smith, S.C. & Whitehead, H. (1989) Diving behaviour of the sperm whale, *Physeter macrocephalus*, off the Galapagos Islands. *Canadian Journal of Zoology*, **67**, 839–846.

Parsons, T.R. & Lalli, C.M. (2002) Jellyfish population explosions: revisiting a hypothesis of possible causes. *La Mer*, **40**, 111–121.

Perrin, W.F. (2015) *World Cetacea database*. Available at: http://www.marinespecies.org/cetacea.

Pollock, K.H., Marsh, H.D., Lawler, I.R. & Alldredge, M. (2006) Estimating animal abundance in heterogeneous environments: an application to aerial surveys for dugongs. *Journal of Wildlife Management*, **70**, 255–262.

Pompa, S., Ehrlich, P.R. & Ceballos, G. (2011) Global distribution and conservation of marine mammals. *Proceedings of the National Academy of Sciences USA*, **108**, 13600–13605.

Prince, E.D. & Goodyear, C.P. (2006) Hypoxia-based habitat compression of tropical pelagic fishes. *Fisheries Oceanography*, **15**, 451–464.

Read, A.J. (2008) The looming crisis: interactions between marine mammals and fisheries. *Journal of Mammalogy*, **89**, 541–548.

Redfern, J.V., Ferguson, M.C., Becker, E.A., Hyrenbach, K.D., Good, C., Barlow, J., Kaschner, K., Baumgartner, M.F., Forney, K.A., Ballance, L.T., Fauchald, P., Halpin, P., Hamazaki, T., Pershing, A.J., Qian, S.S., Read, A., Reilly, S.B., Torres, L. & Werner, F. (2006) Techniques for cetacean–habitat modeling. *Marine Ecology Progress Series*, **310**, 271–295.

Redfern, J.V., Barlow, J., Ballance, L.T., Gerrodette, T. & Becker, E.A. (2008) Absence of scale dependence in dolphin–habitat models for the eastern tropical Pacific Ocean. *Marine Ecology Progress Series*, **363**, 1–14.

Redfield, A.C. (1958) The biological control of chemical factors in the environment. *American Scientist*, **46**, 205–221.

Risien, C.M. & Chelton, D.B. (2008) A global climatology of surface wind and wind stress fields from eight years of QuikSCAT scatterometer data. *Journal of Physical Oceanography*, **38**, 2379–2413.

Schipper, J., Chanson, J.S., Chiozza, F. *et al.* (2008) The status of the world's land and marine mammals: diversity, threat, and knowledge. *Science*, **322**, 225–230.

Seibel, B.A. (2011) Critical oxygen levels and metabolic suppression in oceanic oxygen minimum zones. *Journal of Experimental Biology*, **214**, 326–336.

Spitz, J., Trites, A.W., Becquet, V., Brind'Amour, A., Cherel, Y., Galois, R. & Ridoux, V. (2012) Cost of living dictates what whales, dolphins and porpoises eat: the importance of prey quality on predator foraging strategies. *PLoS ONE*, **7**, e50096.

VanDerWal, J., Falconi, L., Januchowski, S., Shoo, L. & Storlie, C. (2012) *SDMTools: Species Distribution Modelling Tools: tools for processing data associated with species distribution modelling exercises*. Available at: http://cran.r-project.org/web/packages/SDMTools/.

Weilgart, L.S. (2007) The impacts of anthropogenic ocean noise on cetaceans and implications for management. *Canadian Journal of Zoology*, **85**, 1091–1116.

Wood, S.N. (2006) *Generalized additive models: an introduction with R*. CRC, Boca Raton, FL.

Wood, S.N. (2011) Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society B: Statistical Methodology*, **73**, 3–36.

Worm, B., Sandow, M., Oschlies, A., Lotze, H.K. & Myers, R.A. (2005) Global patterns of predator diversity in the open oceans. *Science*, **309**, 1365.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Detection models and effective strip widths for the three cetacean guilds.

Appendix S2 Environmental envelopes for the selected environmental variables.

Appendix S3 Diagnostics of deviance and randomized quantile residuals for the three cetacean guilds.

BIOSKETCH

Laura Mannocci is interested in how cetaceans and other top predators occupy pelagic habitats at the macroecological scale. She uses spatial and habitat modelling tools to answer her research questions. During her PhD in the Centre d'Études Biologiques de Chizé et La Rochelle, she investigated cetacean and seabird strategies of habitat utilization in tropical waters with respect to their energetic costs of living. She is currently a postdoctoral associate in the Marine Geospatial Ecology Lab (Duke University).

Author contributions: V.R., P.M., J.S. and L.M. conceived the ideas; L.M. analysed the data; L.M. led the writing.

Editor: François Guilhaumon