



# Combining vessel-based surveys and tracking data to identify key marine areas for seabirds

M. Louzao<sup>1,5,\*</sup>, J. Bécares<sup>2</sup>, B. Rodríguez<sup>2</sup>, K.D. Hyrenbach<sup>3</sup>, A. Ruiz<sup>4</sup>, J. M. Arcos<sup>2</sup>

<sup>1</sup>Centre d'Etudes Biologiques de Chizé, CNRS, 79360 Villiers en Bois, France

<sup>2</sup>Sociedad Española de Ornitología/BirdLife, C/Múrcia 2-8, local 13, 08026 Barcelona, Spain

<sup>3</sup>Duke University Marine Laboratory, 135 Duke Marine Laboratory Road, Beaufort, North Carolina 28516, USA

<sup>4</sup>Sociedad Española de Ornitología/BirdLife, C/Melquíades Biencinto 34, 28053 Madrid, Spain

<sup>5</sup>Present address: Helmholtz Centre for Environmental Research-UFZ, Permoserstraße 15, 04318 Leipzig, Germany

**ABSTRACT:** An understanding of the distribution and habitat associations of far-ranging marine predators is being increasingly applied to protect these species from anthropogenic threats at sea (e.g. oil spills and fisheries bycatch). Within this framework, this research on Cory's shearwater *Calonectris diomedea* integrates vessel-based survey and tracking data to provide 2 distinct, yet complementary, perspectives of the habitats of this species in the western Mediterranean during incubation in June and chick rearing in August of 2007. We used a hierarchical modelling approach to (1) delineate the foraging habitat of the species using vessel-based surveys and (2) identify its feeding habitat based on tracking data within the Information-Theoretic framework. Our habitat modelling analyses suggest that shearwaters respond to complex bio-physical coupling, illustrated by their association with frontal features and elevated ocean productivity. Our models yielded moderate predictions of Cory's shearwater habitats within 2 distinct spatial scales. At the mesoscale, the foraging range of the species comprised the continental and insular shelf-slope waters of the Iberian Peninsula and the Balearic Islands, between the Gulf of Lions to the north and Cape Palos to the south. At the coarse scale, the tracking data highlighted important feeding areas within this larger foraging range: 3 continental shelf-slope 'hotspots' — (1) Gulf of Lions, (2) Cape Creus–Barcelona–Ebro Delta and (3) Cape La Nao–Cape Palos; from north to south — as well as the insular shelf-slope areas around the Balearic Islands. These results match previous observations of the foraging range and feeding patterns of the species, and are consistent with the interpretation of the regional oceanography. This study highlights how the integration of tracking and vessel-based survey data can provide a wider understanding of the predictability of aggregation (i.e. hotspots) and the key oceanographic habitats of far-ranging seabirds at multiple spatial scales. Thus, complementary data integration is a step forward in conservation studies of far-ranging marine top predators.

**KEY WORDS:** *Calonectris diomedea* · Cory's shearwater · Feeding hotspot · Foraging range · Habitat modelling · Important Bird Area (IBA) · Information-Theoretic approach · Ocean remote sensing · Conservation · Western Mediterranean

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

## INTRODUCTION

The distribution of marine top predators and their prey is influenced by dynamic and complex physical and biological processes (e.g. Hunt et al. 1999). Seabirds forage within heterogeneous and dynamic

environments, where the vertical and horizontal distributions of prey resources are governed by diverse oceanographic processes operating at a range of spatial scales, from macro–mega (1000s of kilometres) to coarse–meso (10s to 100s of kilometres) (Hunt & Schneider 1987). Dense prey patches of drifting and

\*Email: maite.louzao@ufz.de

weakly swimming prey become concentrated at specific features with enhanced vertical (upwelling/downwelling) and horizontal (convergence/divergence) water flow (e.g. Hammer & Schneider 1986, Franks 1992, Haury et al. 1978). In turn, mobile predators concentrate and forage at these same features in response to elevated localized productivity and dense prey patches (e.g. Rodhouse et al. 1996, Johnston et al. 2005).

The increasing awareness of the serious threats that marine top predators face at sea, such as fisheries bycatch and oil spills, has triggered the development of conservation measures to ensure the protection of important key marine areas (e.g. foraging areas and migration corridors; see BirdLife International 2004, Hooker & Gerber 2004). Understanding the distribution patterns and habitat associations of these highly mobile organisms is critical to effectively monitor and protect them. In particular, the large ranges of most marine top predators make spatially explicit conservation approaches targeting high-use areas (hereafter termed 'hot spots') especially attractive (e.g. Hyrenbach et al. 2000, Alpine & Hobday 2007). Following this principle, BirdLife International is expanding the Important Bird Areas (IBAs) programme to the marine environment with the aim of creating a network of marine protected areas (MPAs) that will provide effective protection for seabirds and their underlying habitats (BirdLife International 2004). This initiative has lagged behind similar initiatives in terrestrial ecosystems due to the difficulty of obtaining systematic seabird distribution and movement data at sea. However, this situation has dramatically changed in the last 2 decades, with the advent of systematic seabird surveys and the development of miniaturised tracking technologies (Boyd et al. 2006). Taking advantage of these advances and benefiting from EC-LIFE Natura funding, 2 BirdLife partners, SEO/BirdLife in Spain and SPEA in Portugal, have recently provided comprehensive inventories of marine IBAs for their whole territorial waters (Ramírez et al. 2008, Arcos et al. 2009).

Within the context of the Spanish marine IBA project, the present study showcases our research to identify key marine areas for Cory's shearwater *Calonectris diomedea* in the western Mediterranean, where the local subspecies (*C. diomedea diomedea*) is listed as Threatened in Spain (Carboneras 2004, Red List of Spanish birds at: [www.seo.org/media/docs/LR%20completo%20para%20web.pdf](http://www.seo.org/media/docs/LR%20completo%20para%20web.pdf)), and therefore requires rapid conservation action. This study focuses on vessel-based surveys and GPS-tracking data to assess the predictability of aggregation hot-spots and the key supporting oceanographic habitats. These approaches are complementary, since vessel-based surveys pro-

vide a broad scale (10s to 100s of kilometres) perspective of population-level distribution and habitat associations, whereas the tracking data provide detailed information on the fine-scale (0.1s to 10s of kilometres) use of the marine environment by individual foragers. While both methodologies have been proved to be useful for mapping and prioritizing the critical migration routes and foraging grounds of protected seabirds (e.g. Hyrenbach et al. 2002, Louzao et al. 2006, González-Solís et al. 2007, Harris et al. 2007), few studies have integrated these 2 disparate methodologies (but see Rodhouse et al. 1996, Hyrenbach & Dotson 2003, Hyrenbach et al. 2006).

Herein, we develop habitat suitability models for Cory's shearwater in the western Mediterranean during incubation in June and chick rearing in August, 2007, on the basis of concurrent tracking of individual birds and vessel-based surveys. Habitat suitability techniques use information on species records (time/space) and concurrent environmental factors to generate statistical predictions of potentially suitable species habitats (see review by Guisan & Zimmermann 2000), critical for conservation planning (Gray et al. 2007). In a first step, we used a hierarchical modelling approach to identify those environmental variables that most accurately reflected the oceanographic habitat of Cory's shearwater by (1) delineating the foraging habitat of the birds using vessel-based surveys (i.e. where the birds search for food) and (2) identifying the feeding habitat of the species using tracking data (i.e. where the species feed). After developing the habitat suitability models accounting for the peculiarities of the 2 disparate methodologies, the second step entailed predicting the suitable foraging and feeding habitats of Cory's shearwater within the Information-Theoretic approach. The third step involved evaluating the predictive performance of the models using resampling techniques.

This paper reports the results of this modelling exercise and discusses the limitations and opportunities of both contrasting datasets for the development of improved habitat suitability models. We also examine the conservation implications of this integrated habitat modelling approach and its findings for identifying key areas for marine birds.

## MATERIALS AND METHODS

**Study site.** In spring–summer of 2007, we conducted vessel-based surveys along the Iberian coast and around the Balearic Islands, whereas Cory's shearwaters breeding at the Balearic archipelago were tracked with global positioning system (GPS) loggers (western Mediterranean; Fig. 1). The Balearic Sea, a sub-basin

of the western Mediterranean located between the Iberian Peninsula and the Balearic Islands, is considered a key transition zone between the Gulf of Lions and the Algerian basin (see Supplement 1 available in MEPS Supplementary Material at: [www.int-res.com/articles/suppl/m391p183\\_app.pdf](http://www.int-res.com/articles/suppl/m391p183_app.pdf)).

**Vessel-based data.** At-sea seabird surveys were conducted onboard the R/V 'Cornide de Saavedra', taking advantage of the annual Mediterranean International Trawl Survey (MEDITS; Bertrand et al. 2002); this survey coincided with the incubation period of Cory's shearwater (May 30 to June 29, 2007). Seabird counts followed the methodology proposed by Tasker et al. (1984), adapted to the study area (Louzao et al. 2006): birds were counted within a 300 m strip-transect band, on 1 or both sides ahead of the vessel according to census conditions; snap-shot counts were used to census flying birds. Shearwater observations were summed into 10 min survey bins.

**Tracking data.** We deployed GPS loggers on 29 Mediterranean Cory's shearwaters breeding at 3 Balearic Island colonies between early August and mid September 2007, coinciding with the chick-rearing period: 13 birds from Cala Morell and 6 from Illa de l'Aire, both locations in Menorca, and 10 from Pantaleu in Mallorca (see details in Table S1 in Supplement 2 — available in MEPS Supplementary Material at: [www.int-res.com/articles/suppl/m391p183\\_app.pdf](http://www.int-res.com/articles/suppl/m391p183_app.pdf) — and Fig. 1b for colony locations). Cala Morell is the main breeding colony of Cory's shearwater in the Balearic archipelago, with ca. 1000 to 6000 breeding pairs; Pantaleu holds ca. 200 pairs, and Illa de l'Aire ca. 35 to 40 pairs (Carboneras 2004).

The loggers weighed 25g with dimensions of 46.5 × 32 × 18.5 mm (Earth & Ocean Technologies); they were fixed to the back feathers of the birds with TESA tape, thus increasing total weight to almost 30 g. This represented slightly more than 3% of body mass of the species, which is the recommended threshold for instruments deployed on tubenose birds (Phillips et al. 2003). However, we made efforts to minimise the impact of loggers by deploying them for very short periods (4 to 17 d) (see Table S1 in Supplement 2). Whenever an instrument was not retrieved, the attachment system would guarantee the release of the logger after a few

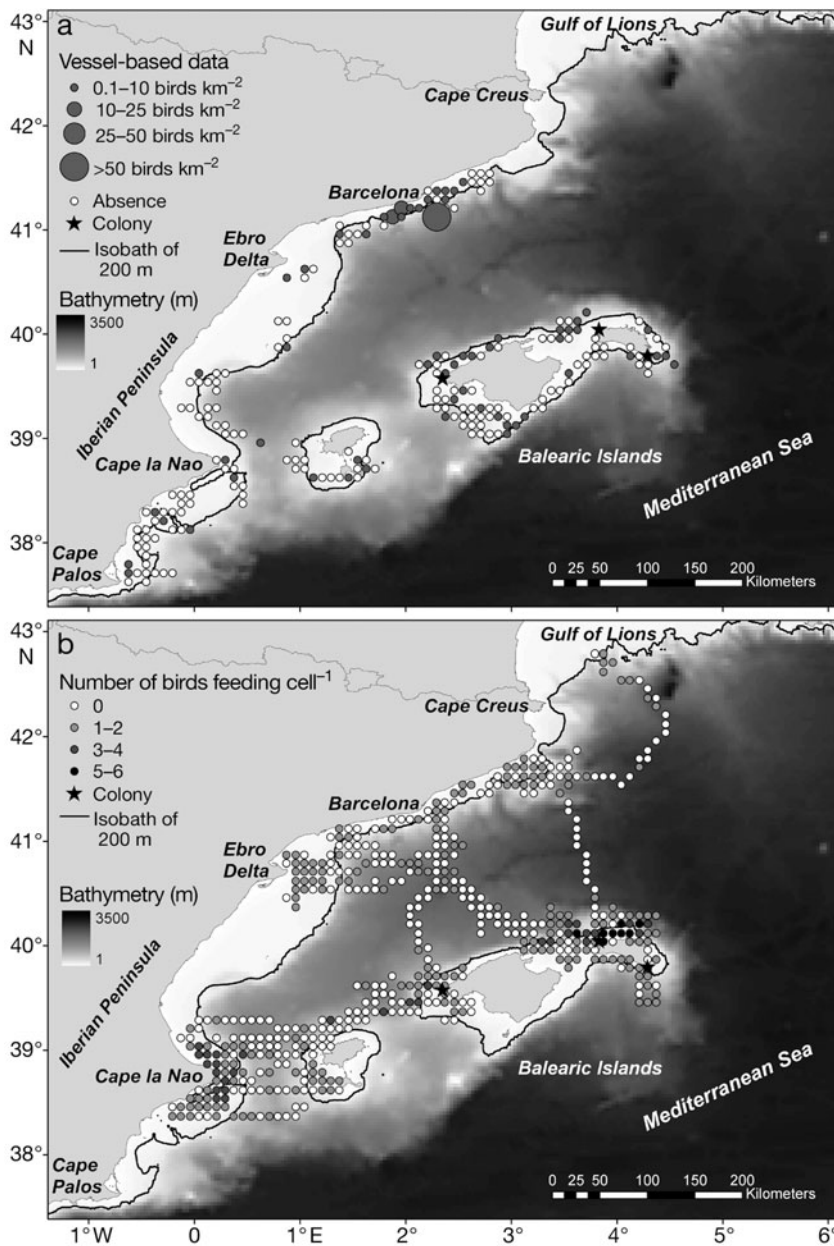


Fig. 1. *Calonectris diomedea*. (a) Foraging and (b) feeding patterns of Cory's shearwaters inferred from vessel-based surveys and tracking data after processing based on 5 n mile cell size, respectively. Vessel-based data corresponds to the presence (represented by density, with filled circles proportional in size to their values) and absence (open circles) of the species, whereas tracking data corresponds to the number of feeding birds by 5 n mile cell. Breeding colonies within the study area are also indicated, as well as the bathymetry (dark grey) and the 200 m isobath (i.e. the limit of the continental shelf). Note that some observations are over the coastline due to an artefact of binning the data in a standard grid

weeks of deployment. Moreover, the species does not seem to be particularly sensitive to carrying devices, as no detrimental effects were detected with smaller loggers (12 g GLS) attached for far longer periods (1 yr) (Iguál et al. 2005).

The programmable GPS recording interval was set at 5 min initially (first 12 deployments) and extended to 10 min thereafter (remaining 17 deployments) to increase battery performance (see Table S1 in Supplement 2). Average battery life-span was 3.7 d (range 2.0 to 4.4) and 8.3 d (range 6.0 to 10.5) for the 5 and 10 min sampling rates, respectively. At these 2 sampling intervals, we registered an average of 245 and 132 positions per trip.

Although 29 birds were equipped, we obtained tracking data for 19 of them. Due to the small positional error of the GPS loggers (< 20 m for 90% of locations), we were able to precisely assess fine-scale habitat use patterns by following the birds' trajectories in detail. Moreover, we inferred bird behaviour on the basis of the apparent flying speeds of the tracked birds: travelling, searching, feeding, and resting (sitting on the water or nesting). We classified the bird behaviour into 4 categories, on the basis of the movement rates calculated between successive positions ( $\text{km h}^{-1}$ ) and visual inspection of trips: resting on the water (< 2), feeding (2 to 10), searching (10 to 15), and travelling (> 15) (see example in Fig. S1 in Supplement 2). Then, we grouped these behaviours in 2 classes: 'feeding' or 'not feeding' (the latter corresponding to resting on the water, searching, and travelling) and assigned this categorical variable to each location.

**Data processing.** To compare the habitat modelling results from the vessel-based surveys and the tracking data, we first standardised these 2 disparate datasets using a common spatial scale of 5 n miles (~9.3 km) and a temporal resolution of 1 mo. This resolution, selected on the basis of the coarser of the inherent scales of the datasets, was previously used to characterize the oceanographic habitat of the Balearic shearwater *Puffinus mauretanicus* within the same area (Louzao et al. 2006). Thus, all subsequent analyses were performed using a grid of 5 × 5 n mile cells, with an extent delineated by the spatial distribution of the tracking and survey data.

For vessel-based surveys, 10 min Cory's shearwater observations within each cell were summed and then re-coded into a binary presence/absence variable, indicative of whether at least 1 bird was recorded within a given cell. Thus, a total of 660 bins (10 min bins) were aggregated into 212 grid cells corresponding to 30 survey d. Due to cloud cover, we discarded 0.94% of the surveyed grid cells with incomplete remote sensing information and we used the remaining grid cells (n = 210) as the training dataset for the

analysis of shearwater occurrence, containing 60 'presences' and 150 'absences'.

Regarding tracking data, we derived a binomial response variable for assessing the feeding habitat coding those cells containing at least 1 feeding event (GPS location characterized by 2 to 10  $\text{km h}^{-1}$ ) as 'feeding'. Conversely, those cells where no feeding events were observed were coded as 'not feeding'. To avoid pseudoreplication, we randomly selected a complete foraging trip per bird (tracking dataset: independent trips of 19 individuals). Thus, the initial training dataset of 7324 locations from 19 foraging trips was aggregated into 715 diurnal cells, containing 300 'feeding' and 415 'not feeding' events.

**Concurrent environmental variables.** We selected environmental variables on the basis of possible biological relevance and the availability of data (see Table 1 for a complete list). Bathymetry (BAT), sea surface temperature (SST) and chlorophyll *a* concentration (CHL, as a proxy of biological production) were extracted for the study area containing all locations (Table 1). Bathymetric data were obtained from NOAA's ETOPO 2-minute dataset ([www.ngdc.noaa.gov/mgg/gdas/gd\\_designagrid.html?dbase=GRDET2](http://www.ngdc.noaa.gov/mgg/gdas/gd_designagrid.html?dbase=GRDET2)). We derived monthly composites of SST (night-time) and CHL from MODIS/Aqua (available at ca. 0.04° spatial resolution in <http://poet.jpl.nasa.gov/> and at ca. 0.05° in <http://coastwatch.pfel.noaa.gov/coastwatch/CWBrowserWW180.jsp>, respectively). For dynamic variables such as SST and CHL, it is unlikely that marine top predators distribution responds instantaneously to changes in oceanographic variables (Redfern et al. 2006). Thus, we used the integrated value of those variables for the period from February to April preceding the surveys as a proxy of oceanographic patterns (see Fig. S2 in Supplement 3, available in MEPS Supplementary Material at: [www.int-res.com/articles/suppl/m391p183\\_app.pdf](http://www.int-res.com/articles/suppl/m391p183_app.pdf), for details of the analysis). Since these 3 habitat variables (SST, CHL, and BAT) were not normally distributed, we used the median as a central tendency statistic instead of the mean within each 5 n mile cell grid. The median is less strongly influenced by outliers, thus minimizing potential biases in the dataset (Whitmire et al. 2007). Additionally, we estimated their spatial gradients by estimating their proportional change (PC) within a surrounding 3 × 3 cell (15 × 15 n mile) grid using a moving window as follows:  $PC = [(maximum\ value - minimum\ value) \times 100] / maximum\ value$ . This dimensionless metric expresses the magnitude of change in each habitat variable, scaled to the maximum value (e.g. Louzao et al. 2006). Dynamic variables such as SST and CHL corresponding to vessel-based surveys and tracking data are shown in Fig. S3a–d, as well as static variables such as BAT and its gradient



(Fig. S3g,h, available in MEPS Supplementary Material at: [www.int-res.com/articles/suppl/m391p183\\_app.pdf](http://www.int-res.com/articles/suppl/m391p183_app.pdf)).

Distance to oceanographic fronts (FRONT) was estimated in a monthly basis using the analysis of satellite-derived SST fields (monthly MODIS/Aqua from <http://oceancolor.gsfc.nasa.gov>) under a geographic information system (GIS). Our approach was based on the Cayula-Cornillon algorithm (Cayula & Cornillon 1992), which identifies fronts by detecting the edge of adjacent water masses of different surface water temperature. We used the algorithm implemented in the Marine Geospatial Ecology Tools (available at <http://code.env.duke.edu/projects/mget>) for front detection and estimated the distance to the closest oceanographic front for each analysis grid cell (Fig. S3e,f).

To account for the influence of central-place foraging shearwaters (Orians & Pearson 1979), we included the distance between each grid cell and the nearest breeding colony in the case of vessel-based surveys (unknown origin of shearwaters) and to the colony of origin in the case of tracking data (COLONY, hereafter). The distance from each grid cell to the nearest shoreline (COAST) was also included in the model to take into account onshore-offshore distribution patterns. Finally, we also estimated the distance to the limit of the continental shelf (SHELF, delimited by the 200 m isobath) since the shelf-break is a productive bathymetric feature exploited by Cory's shearwater (Abelló et al. 2003). All distances were calculated using the Nearest Features extension of ArcView 3.2 (Jenness 2004).

**Statistical analysis.** We applied a quantitative habitat modelling procedure in order to identify key areas for Cory's shearwater, which involved 3 steps: analysis of spatial correlation patterns, development of habitat suitability models within the Information-Theoretic framework, and model evaluation using resampling techniques.

**Spatial autocorrelation:** Species distributional data are characterised by spatial autocorrelation, which occurs when adjacent observations are more similar than would be expected in randomly distributed data (Sokal et al. 1998). When aggregations are present, such as seabird flocks, significant positive autocorrelation can cause the false rejection of the null hypothesis (Type I error), increasing the probability of finding spurious significant habitat relationships (Hurlbert 1984).

Therefore, prior to habitat modelling, we checked the independence of the foraging and feeding patterns of shearwaters by means of the Moran's  $I$  coefficient, which measures the similarity of the response variables by relating the values for all pair-wise combina-

tions of cells as a function of their spatial distance. Moran's  $I$  values range from  $-1$  (negative autocorrelation) to  $+1$  (positive autocorrelation), and the resulting correlograms typically show a decrease of spatial autocorrelation to a value of 0, indicating no spatial autocorrelation at increasing distance between locations (Sokal et al. 1998). We considered up to 15 lags at 1 grid cell (9.3 km) intervals in all directions (i.e. isotropic spatial autocorrelation structure). Then, significance of the Moran's  $I$  values was assessed by performing 1000 permutations at each lag distance and contrasting the observed Moran's  $I$  values with those predicted by Monte Carlo randomization tests, after applying the Bonferroni correction for multiple testing (modified alpha = 0.05 /15; 15 lags for each dataset). All spatial autocorrelation analyses were performed using the excel add-in Rook Case (Sawada 1999).

**Habitat modelling:** We used a hierarchical modelling approach to identify those environmental variables that most accurately reflected the oceanographic habitat of Cory's shearwater by (1) delineating its foraging habitat using vessel-based surveys and (2) identifying its feeding habitat using tracking data of individual birds. Habitat suitability models were developed, accounting for the peculiarities of these 2 disparate methodologies.

Prior to modelling, all variables were tested for normality and, where necessary, were log-transformed (natural logarithms; in order to avoid the influence of outliers) and standardized to have a mean of 0 and an SD of 1 due to differing ranges of variables (Table 1) (Zuur et al. 2007). A preliminary screening of explanatory variables was undertaken to check for colinearity (cross-correlation analysis) by calculating all pairwise Spearman rank correlation coefficients. When pairs of predictor variables were strongly correlated ( $|r_s| > 0.5$ ), we discarded one of the redundant variables (e.g. Gray et al. 2007, McAlpine et al. 2008). We selected the 'non-correlated' predictors using the lowest Akaike Information Criteria (AIC) from univariate models of the 2 predictors (Burnham & Anderson 2002). This approach led to the removal of 6 predictors in the case of tracking data (SST, CHLG, BATG, COAST, COLONY, SHELF) and 2 predictors in the case of the vessel-based surveys (CHL, BAT) (see Table 2).

Habitat suitability models were developed by means of logistic regressions in order to generate a statistical function that would facilitate the prediction of the potential suitable habitat for the species (review in Guisan & Zimmermann 2000). We used generalized linear models (GLMs) for delineating the foraging habitat of species based on the 'glm' function, whereas generalized linear mixed models (GLMMs) were applied for identifying the feeding habitat using the 'lmer' function implemented in the 'lme4' package

(R Development Core Team 2008). Both logistic regressions were fitted with a binomial error distribution ('presence/absence' and 'feeding/not feeding') and a logit link function. In the case of GLMMs, we included the individual as random intercept term in order to account for individual effects. Models were built for all possible linear combination of 'non-correlated' explanatory variables ( $|r_S| < 0.5$ ) and no interaction terms were included.

**Model selection strategy:** Models were ranked based on their AIC value corrected for small sample sizes ( $AIC_c$ ) and we calculated the Akaike weight ( $w_i$ ) for each model, which represents the relative likelihood of candidate models (Burnham & Anderson 2002). If the model with the lowest  $AIC_c$  is not undoubtedly the 'best' (e.g.  $w_i > 0.90$ ), a model averaging procedure might be more appropriate to account for parameter uncertainty (Burnham & Anderson 2002). Therefore, we constructed a 95% confidence set of models where the sum of Akaike weights was  $> 95$ , starting with the model with the highest Akaike weight (Burnham & Anderson 2002). Accordingly, averaged coefficients were estimated from the 95% confidence set of models containing that variable, as well as the variance estimator in order to assess the precision of the estimates (Burnham & Anderson 2002, Johnson & Omland 2004).

Then, we calculated the probability of shearwater foraging and feeding for each grid cell using the averaged logistic model developed with 'presence/absence' and 'feeding/not feeding' data. We calculated the linear predictor (LP) using the intercept and the variable coefficients from the averaged models, and calculated the probability of models (Pr) as follows:  $Pr = e^{LP} \times (1 - e^{LP})^{-1}$ .

Finally, we could have ranked predictor importance summing the Akaike weights for all models containing an explanatory variable (e.g. Burnham & Anderson 2002, McAlpine et al. 2008), but Murray & Conner (2009) found that this approach was not sufficiently sensitive to correctly rank variable importance, suggesting alternative methods such as hierarchical partitioning. As an alternative, Burnham & Anderson (2002) recommended a randomization procedure to estimate the baseline value for  $w_i$ , denoted as  $w_{i0}$ , and then measure variable importance by computing the difference between  $w_i$  and  $w_{i0}$ . For those predictors with negligible predictive value, these differences should be close to zero. We used this approach for ranking predictor importance (see Supplement 5, available in MEPS Supplementary Material at: [www.int-res.com/articles/suppl/m391p183\\_app.pdf](http://www.int-res.com/articles/suppl/m391p183_app.pdf)).

Table 1. *Calonectris diomedea*. Median (range) of the explanatory variables for Cory's shearwater vessel-based survey data (presence, n = 60, and absence, n = 150) and tracking data (feeding, n = 300, and not feeding events, n = 415) corresponding to 19 trips. An oceanography interpretation is also provided

Habitat variables	Vessel-based surveys		Tracking data		Indicative of the following processes
	Presence	Absence	Feeding	Not feeding	
Sea surface temperature (SST, °C)	13.97 (13.19–14.74)	14.2 (13.25–14.88)	14.31 (13.5–14.9)	14.3 (13.42–14.87)	Water mass distribution
SST gradient (SSTG)	1.65 (0.45–4.87)	1.66 (0–3.97)	0.86 (0–11.16)	1.02 (0–10.07)	Small-scale SST variability
Chlorophyll <i>a</i> (CHL, mg m <sup>-3</sup> )	0.37 (0.26–1.22)	0.38 (0.25–1.21)	0.38 (0.26–1.52)	0.37 (0.25–2.43)	Ocean productivity domains
CHL gradient (CHLG)	11.19 (1.49–75.84)	9.47 (0.83–66.25)	15.22 (1.12–62.26)	9.95 (0.22–75.44)	Small-scale CHL variability
Bathymetry (BAT, m)	154.12 (32.5–971.25)	149 (20.33–973.5)	153.5 (30–2511)	767 (35.5–2518.5)	Coastal vs. pelagic domains
BAT gradient (BATG)	90.39 (39.93–99.61)	84.08 (16.45–99.42)	81.12 (4.21–97.84)	70.77 (4.05–99.64)	Presence of topographic features (shelf-break, seamounts)
Distance to colony <sup>a</sup> (COLONY, km)	36.76 (1.67–196.25)	47.04 (0.94–198.41)	139.36 (0.94–313.65)	122.37 (0.94–310.41)	Colony influence on central place foragers
Distance to shoreline (COAST, km)	10.58 (0.03–40.59)	16.12 (0.45–43.30)	15.19 (0–86.02)	27.29 (0–101.91)	Onshore–offshore distribution patterns
Distance to oceanographic fronts (FRONT, km)	50.71 (0–84.21)	69.14 (0–197.54)	97.24 (0.78–184.36)	93.38 (0–186.62)	Mesoscale frontal systems
Distance to continental shelf (SHELF, km)	5.27 (0.12–39.05)	6 (0.06–25.04)	9.45 (0.003–74.05)	13.37 (0.03–79.76)	Proximity with shelf-break (slope currents, vertical mixing and prey concentration)

<sup>a</sup>Distance between the centroid of each grid cell and the nearest breeding colony for vessel-based surveys, and distance to the colony of origin for tracking data

**Model fit:** To assess the fit of the model with the lowest  $AIC_c$ , we used a Pearson  $\chi^2$  goodness-of-fit test (Crawley 1993); we checked for spatial autocorrelation in model residuals constructing Moran's  $I$  correlogram (see Supplement 5).

**Model evaluation:** Assessing the predictive ability of a model is a crucial step for allowing its proper use in ecological applications, especially when conservation issues are the main objective. The area under the curve (AUC) of a receiver operating characteristic (ROC) plot is widely utilised to assess habitat models developed through logistic regression (e.g. Brotons et al. 2004, Elith et al. 2006, Louzao et al. 2006, McAlpine et al. 2008), and was used here to validate our model predictions (Fielding & Bell 1997). This approach can also be applied to any model that produces estimates of probability for a binomially distributed response variable (e.g. feeding/not feeding). ROC curves are simply plots of sensitivity (the fraction of correctly predicted presences) against 1 – specificity (the fraction of correctly predicted absences) with changing critical values of threshold probability. AUC is a threshold-independent summary statistic that ranges from 0 to 1 (from negligible to perfect discriminatory power, respectively). AUC values represent the discriminatory ability of a model as follows: higher than 0.9, excellent; from 0.9 to 0.8, good; from 0.8 to 0.7, moderate; from 0.7 to 0.6, poor; and from 0.6 to 0.50, unsuccessful (Swets 1988).

We applied a cross-validation procedure to assess the predictive performance of the averaged model resulting from the Information-Theoretic approach using 2 different approaches: (1) use of an independent dataset for each of the 2 types of data and (2) resampling techniques (e.g. bootstrap) which provide an alternative approach for evaluating the model with the original data (Guisan & Zimmermann 2000, McAlpine et al. 2008). Concerning vessel-based survey data, we used the MEDITs surveys corresponding to 2006, which also covered the same geographic area comprising 216 cells (5 n mile cells; 51 presences and 165 absences). For tracking data, the independent dataset was built by randomly selecting a second foraging trip for each individual, which includes 15 foraging trips (15 individuals at least performed 2 foraging trips) with a total of 664 cells (5 n mile cells; 266 'feeding' and 398 'not feeding'). Both datasets should be preferably mentioned as quasi-independent data sets since they covered the same geographic extent of the original dataset (Guisan & Zimmermann 2000).

The cross-validation procedure was repeated 1000 times, and during each simulation we randomly assigning the 70% of the original data to the training dataset and 30% to the test dataset (e.g. Brotons et al. 2004). The best subset of models resulting from the Information-Theoretic approach (Table S2 in Supplement 5) was fit-

ted to the training dataset, the averaged coefficients extracted and the averaged model was fitted to the test dataset. AUC values were estimated in each simulation for both training and test datasets. After the 1000 simulations, the mean and upper and lower 95% CI of the AUC of both training and test dataset were used as a cross-validation measure of the predictive performance of the averaged model (McAlpine et al. 2008). If the lower 95% CI limit does not include the 0.5 value, then there is evidence that the model has an ability to discriminate between the 2 groups (Hanley & McNeil 1982). We also conducted the same cross-validation procedure to the quasi-independent dataset following the same random procedure as previously described.

**Mapping habitat probability.** We represented probabilities of suitable habitat for the vessel-based survey dataset (June 2007) and the tracking dataset (August 2007) with GIS. A regular lattice of 5 n mile cells was placed over the study area, and environmental variables considered in the averaged model were extracted. Model-averaged predictions of the foraging and feeding habitat of Cory's shearwater were estimated. We could have transformed probabilistic predictions into a binomial response variable based on threshold values in order to identify the suitable potential habitat; however, providing a continuous probability surface may be the most flexible and powerful method, thus allowing managers to select different thresholds depending on the model objective (Freeman & Moisen 2008). Additionally, the formal combination of both datasets was disregarded due to the different timing of the datasets, which corresponded to the Cory's shearwater incubation period for vessel-based surveys and chick-rearing for tracking data, respectively. Therefore, although multiplying both probabilities might be statistically correct, the most biologically correct approach might be multiplying probabilities of the 2 disparate datasets only when available for exactly the same period.

## RESULTS

### The seascape of Cory's shearwater

The seascape occupied by Cory's shearwater is characterised by the oceanographic patterns typical of the western Mediterranean Sea, including strong habitat gradients evidenced by the significant cross-correlations between several habitat variables ( $|r_S| > 0.5$ , Table 2). The west and north of the study area (the Ebro Delta and the Gulf of Lions, respectively) were characterised by a higher oceanographic variability at the small scale, reflected in both CHL and SST, and represent the most productive ( $8 \text{ mg m}^{-3}$ , maximum

integrated CHL values between February and April) and coldest waters of the study area (see Supplement 4, available in MEPS Supplementary Material at: [www.int-res.com/articles/suppl/m391p183\\_app.pdf](http://www.int-res.com/articles/suppl/m391p183_app.pdf) for oceanographic characterisation). Additionally, important mesoscale frontal systems were also identified within the study area, mainly at the north of the Balearic Islands (Fig. S3e,f).

### Spatial autocorrelation

We found no evidence of significant spatial autocorrelation in Cory's shearwater distributions for either the vessel-based survey data or the GPS tracking data at the selected scales of analysis (9.3 km cells, with lags from 9.3 to 139.5 km) (Fig. 2). Both datasets yielded small magnitude Moran's *I* values (from +0.2 to -0.1), suggestive of weak aggregated spatial patterns (Fig. 2). Thus, the correlograms revealed that the selected spatial scale of analysis yielded independent observations, suitable for performing the habitat modelling.

### Modelling foraging probability

For vessel-based data, the model with the lowest AIC<sub>c</sub> (AIC<sub>c</sub> = 241.581) contained SST, BATG and

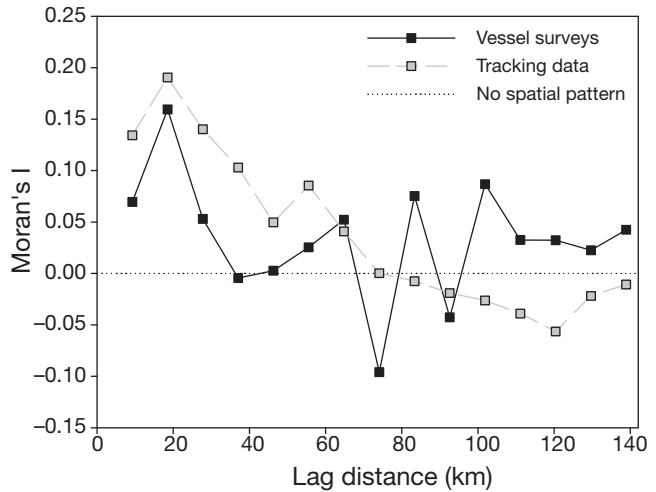


Fig. 2. *Calonectris diomedea*. Results of the spatial autocorrelation of vessel-based survey and tracking data within 5 distance lags (each lag increment corresponds to 5 n miles [9.3 km]). No evidence of significant spatial autocorrelation was found for any values of either dataset

COAST (see Table S2 in Supplement 5). Based on the correlogram analysis, we did not find any significant spatial autocorrelation in the residuals (see Fig. S4 in Supplement 5). Also, the Pearson  $\chi^2$  goodness-of-fit test revealed no evidence of significant lack of fit ( $\chi^2 = 212.29$ ,  $p = 0.366$ ,  $df = 206$ ). Finally, the area under

Table 2. *Calonectris diomedea*. Results of the cross-correlation analysis of environmental variables for tracking data and vessel-based survey based on the Spearman-rank correlation coefficient  $r_s$  (in bold  $|r_s| > 0.5$ ) and corresponding significant levels (lower and upper diagonal, respectively). Significant level set at  $<0.05$ ,  $<0.01$  and  $<0.001$ ; NS: not significant. See Table 1 for abbreviations

	SSTG	SST	CHLG	CHL	BATG	BAT	COLONY	COAST	FRONT	SHELF
<b>Vessel-based survey data</b>										
SSTG	–	NS	NS	NS	NS	NS	NS	0.01	NS	0.05
SST	-0.080	–	NS	0.001	NS	NS	0.001	NS	0.05	NS
CHLG	-0.056	-0.068	–	0.001	NS	0.001	0.001	NS	NS	NS
CHL	0.065	-0.261	0.385	–	0.001	0.001	0.001	NS	NS	0.05
BATG	-0.095	0.005	-0.071	-0.318	–	NS	0.01	0.001	0.001	0.001
BAT	0.097	-0.069	-0.392	-0.288	-0.058	–	NS	0.001	NS	NS
COLONY	0.093	-0.494	0.356	<b>0.586</b>	-0.204	-0.022	–	0.001	0.001	NS
COAST	0.198	-0.135	-0.132	0.082	-0.480	<b>0.580</b>	0.303	–	0.05	NS
FRONT	-0.041	0.161	-0.001	0.099	-0.290	0.030	-0.308	0.167	–	NS
SHELF	0.153	-0.105	0.050	0.150	-0.493	-0.118	0.079	0.083	0.072	–
<b>Tracking data</b>										
SSTG	–	0.001	0.000	0.001	0.001	0.001	0.001	0.001	0.001	0.001
SST	-0.377	–	0.001	0.001	0.010	0.010	0.001	0.001	0.001	0.001
CHLG	0.048	-0.310	–	0.001	0.001	0.001	0.001	0.001	0.010	0.001
CHL	0.348	-0.436	0.319	–	0.001	0.001	0.001	0.010	0.001	0.001
BATG	-0.305	0.121	0.287	-0.290	–	0.001	0.001	0.001	0.001	0.001
BAT	0.227	-0.116	<b>-0.534</b>	-0.150	<b>-0.564</b>	–	0.001	0.001	0.001	0.001
COLONY	0.379	-0.492	0.281	<b>0.769</b>	-0.308	0.133	–	0.001	0.001	0.001
COAST	0.400	-0.242	-0.392	0.114	<b>-0.744</b>	<b>0.830</b>	0.193	–	0.001	0.001
FRONT	-0.326	<b>0.713</b>	-0.097	-0.248	0.258	-0.251	-0.353	-0.332	–	0.001
SHELF	0.241	-0.156	-0.262	0.181	<b>-0.788</b>	<b>0.582</b>	0.143	<b>0.667</b>	-0.207	–



the ROC curve was  $0.714 \pm 0.038$ , indicating moderate discriminating ability. However, this model had an Akaike weight of 0.133, indicating substantial model uncertainty (66 models in the 95% confidence set) and that a model averaging approach was appropriate.

All ‘non-correlated’ explanatory variables were included within the 95% model set, and the relationship between the response variable and predictors is based on the sign of the averaged coefficients (between brackets). The SST (-) showed the strongest negative effect on shearwaters occurrence, whereas the rest of predictors ranked from BATG (+) > COLONY (-) > COAST (-) > CHLG (+) > FRONT (-) > SHELF (-) > SSTG (+) (Fig. 3 and Table S2 in Supplement 5). At the wider scale, foraging might occur in cool waters characterised by high bathymetric variability close to the colonies and the coast. If within these areas, high chlorophyll variability and the presence of frontal systems are expected to increase the foraging probability of shearwaters both along the Iberian Peninsula (from the Gulf of Lions to Cape Palos) and around the Balearic Islands (Figs. 1 & 4, Supplement 4).

The AUC value ( $\pm$ SD) of the vessel-based surveys averaged model showed a moderate model performance ( $0.745 \pm 0.038$ ). The cross-validation of the original vessel-based survey data yielded moderate values of AUC: 0.747 for the training dataset (CI 95%: 0.700 to 0.806) and 0.676 for the test dataset (CI 95%: 0.570 to

0.780). Similarly, cross-validation of the quasi-independent dataset showed that our averaged model failed to provide reasonable foraging predictions: AUC of 0.684 for the training dataset (CI 95%: 0.627 to 0.746) and 0.600 for the test dataset (CI 95%: 0.510 to 0.710).

### Modelling feeding probability

For the tracking data, the model with the lowest  $AIC_c$  ( $AIC_c = 903.94$ ) contained BAT and FRONT (see Table S2 in Supplement 5). Based on the correlogram analysis, we did not find any significant spatial autocorrelation in the residuals (Fig. S4 in Supplement 5). Also, the Pearson  $\chi^2$  goodness-of-fit test revealed no evidence of significant lack of fit ( $\chi^2 = 712.86$ ,  $p = 0.473$ ,  $df = 711$ ). Finally, the area under the ROC curve was  $0.724 \pm 0.019$ , indicating moderate discriminating ability. However, this model has an Akaike weight of 0.44, indicating substantial model uncertainty, with 6 models in the 95% confidence set. This indicates that a model-averaging approach was appropriate.

All ‘non-correlated’ explanatory variables were included within the 95% model set with BAT (-) having the strongest negative effect on shearwaters feeding probability and occurring in all models in the 95% set (see Table S2 in Supplement 5). The relative importance of individual variables ranked additional variables as FRONT (-) > SSTG (-) > CHL (+). At the finer scale, feeding might occur in relatively shallow waters, close to frontal systems in areas of low fine-scale SST variability and highly productive areas along the Iberian Peninsula and around the Balearic Islands (Fig. 1, Supplement 4). Within this oceanographic context, tracking observations allowed us to identify 3 important feeding hotspots along the continental shelf-slope area of the Iberian Peninsula (Gulf of Lions, Cape Creus–Barcelona–Ebro Delta, and Cape La Nao–Cape Palos; from north to south), as well as the shelf-slope areas around the Balearic Islands (Figs. 1 & 4).

The AUC value ( $\pm$ SD) of the tracking averaged model showed a moderate model performance ( $0.700 \pm 0.019$ ). The cross-validation of the original tracking data demonstrated that our averaged model showed a moderate predictive performance: AUC of 0.701 for training dataset (CI 95%: 0.674 to 0.730) and 0.700 for test dataset (CI 95%: 0.640 to 0.748). However, the cross-validation of the quasi-independent dataset showed that our averaged model failed to provide reasonable feeding predictions: AUC of 0.583 for the training dataset (CI 95%: 0.548 to 0.615) and 0.561 for the test dataset (CI 95%: 0.51 to 0.623), but still discriminate between ‘feeding’ and ‘not feeding’

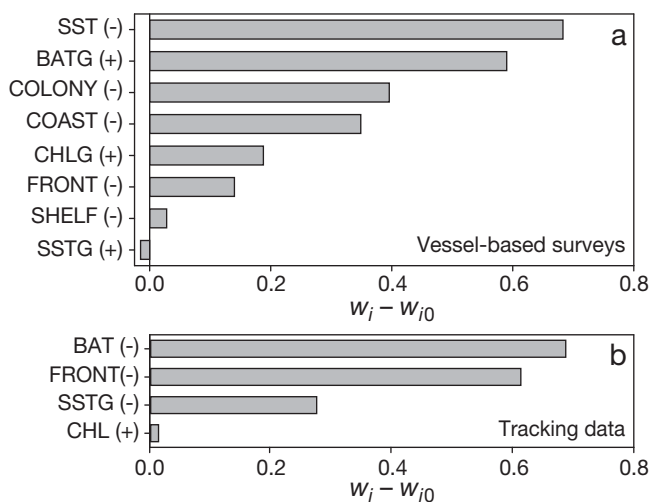


Fig. 3. *Calonectris diomedea*. Ranking of seascape variables based on a randomization procedure to estimate the baseline value for Akaike weight ( $w_i$ ), denoted as  $w_{i0}$ , for (a) vessel-based survey and (b) tracking data. If the difference between  $w_i$  and  $w_{i0}$  is close to zero, the predictor may not have any predictive value (Burnham & Anderson 2002). The sign within parentheses indicates the sign of the effect of the variable on the probability of Cory’s shearwaters foraging and feeding (see Table S2). See Table 1 for definition of variable abbreviation

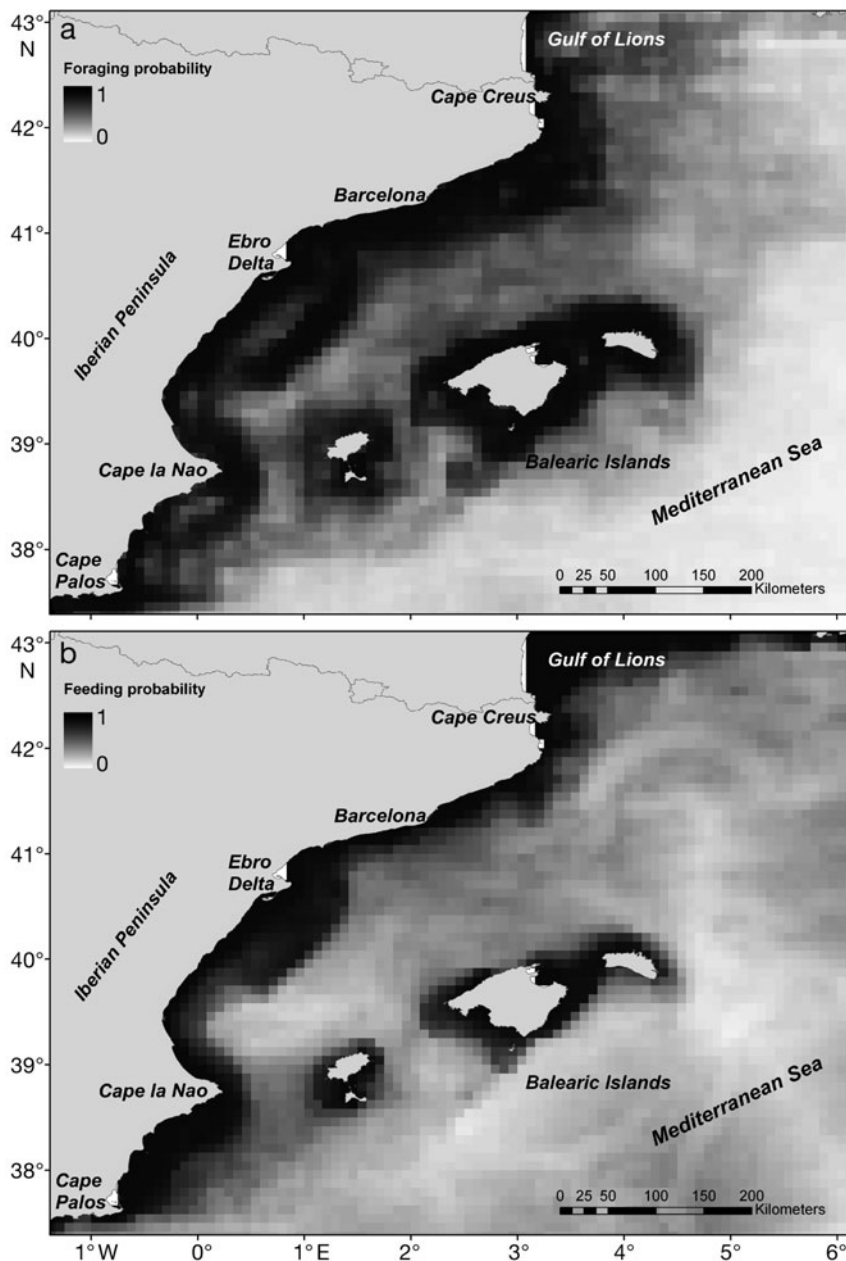


Fig. 4. *Calonectris diomedea*. Averaged models predicted (a) foraging probability (vessel-based survey data) and (b) feeding probability (tracking data) of Cory's shearwater

## DISCUSSION

We undertook a hierarchical analysis of the oceanographic features influencing the feeding and foraging habitat of a vulnerable marine top predator, Cory's shearwater, in the complex and dynamic seascape of the western Mediterranean by combining 2 disparate and complementary data sets. Our results provided new relevant insights for defining the oceanographic

habitat and for predicting the distribution of shearwaters during the breeding season. In the following sections, we discuss the conservation implications of this integrated habitat modelling approach and its findings for identifying key areas for marine birds.

### The seascape of Cory's shearwater: identifying key areas

The integration of tracking and vessel data provided a unique opportunity for identifying key marine areas for Cory's shearwaters by providing 2 distinct, yet complementary habitat perspectives that reinforced each other. Our habitat modelling analyses suggested that shearwaters respond to complex bio-physical coupling illustrated by their association with frontal features and elevated ocean productivity. Within a mesoscale perspective, the foraging range of the species comprised the continental and insular shelf-slope areas between the Gulf of Lions to the north and Cape Palos to the south. Within a coarser perspective, the tracking data highlighted 3 important feeding hotspots along this continental shelf-slope area of the Iberian Peninsula (Gulf of Lions, Cape Creus to Ebro Delta, and Cape La Nao to Cape Palos). Despite the more restricted spatial coverage of the vessel-based surveys, these data identified a further important feeding hotspot (showing densities  $>76$  birds  $\text{km}^{-2}$ ) off Barcelona (Fig. 1a). All 4 feeding hotspots might be influenced by the Northern Current, which interacts with different physical (e.g. submarine canyons in the Cape Creus) and chemical features (e.g. river run-off such as the Ebro) along its path, creating diverse meso- and coarse-scale oceanographic processes along the shelf-slope region and transporting rich nutrient waters from the Gulf of Lions southwards to the Cape Palos (Millot 1999). Our results agree with previous studies, which have underscored the shelf-slope distribution of Cory's shearwater and its tendency to occur along the Iberian continental shelf,

mainly between the Ebro Delta and Cape La Nao (Abelló et al. 2003). However, to our knowledge, this is the first study to quantitatively assess Cory's shearwaters marine habitat use on the basis of both static and dynamic habitat variables. Dynamic variables such as coastal water masses and frontal systems within productive continental shelves often delineate key feeding habitats for breeding seabirds (Skov et al. 2008). Contrary to the productive temperate and subpolar continental shelf habitats, the Mediterranean is considered an oligotrophic Sea, although remotely-sensed chlorophyll *a* values documented in the study area were generally characteristic of a mesotrophic regime (Louzao et al. 2006, present study).

Our tracking data also provided a larger ecological context for the interpretation of these feeding hotspots, by revealing that central-place foraging Cory's shearwaters repeatedly commuted between the less productive waters around the breeding colonies (Balearic Islands) and the highly productive waters of the shelf-slope areas of the Iberian Peninsula, thus exploiting 2 distinct shelf-slope foraging areas during the breeding period: continental (Iberian Peninsula) and insular (Balearic Islands). Commuting seems to be a common behaviour of pelagic birds within temperate and polar regions, and might suggest that breeding seabirds 'know' where to find food, probably from previous experience (Weimerskirch 2007). In fact, the predictability of marine resource dispersion is time-scale dependent, and likely varies across marine habitats (Weimerskirch 2007). For instance, the location of resources at large and intermediate scales appears reasonably predictable for seabirds over long time periods (e.g. seasons). The association of Mediterranean pelagic seabirds with the same mesoscale features along the Iberian continental shelf year after year exemplifies this long-term predictability (Arcos & Oro 2002, Louzao et al. 2006). Within the coarse spatial scale, prey patches are likely to be scattered within mesoscale features (Weimerskirch 2007), which might vary both spatially and temporally depending on the influence of physical and biological drivers, including riverine discharge, the timing and location of spawning stocks of small pelagic fish, and spatio-temporal dynamics of fisheries in the western Mediterranean (Lloret et al. 2004).

### Habitat modelling approach and limitations

While the Information-Theoretic approach has been widely applied to terrestrial ecosystems (Gray et al. 2007, McAlpine et al. 2008), it has not been used in the more dynamic marine ecosystem as a robust model selection strategy (but see Ribic et al. 2008). Within this framework, competing models are evaluated by assessing their relative support in relation to observed

data, rather than using the best single model approach (Buckland et al. 1997, Burnham & Anderson 2002, Johnson & Omland 2004). Furthermore, when models have similar levels of support, model averaging can be used to make robust predictions. Nevertheless, Richards (2005) raised some concerns related to model averaging with Akaike weights and stressed the need for continued research on the effectiveness of model averaging, as did Burnham & Anderson (2004). When comparing different methods for ranking variable importance, Murray & Conner (2009) found that summing the Akaike weights for all models containing  $x_i$  explanatory variables was not sufficiently sensitive to correctly rank variable importance. Burnham & Anderson (2002) also acknowledged some limitations of the summing of Akaike weights, since it cannot yield zero, even if some of the explanatory variables  $x_i$  have no contextual predictive value at all. They suggest (among others) a randomization method that we developed in the present study which might deserve further research (see Supplement 5).

We used logistic regressions (GLMs and GLMMs) to develop habitat suitability models using binary response variables. GLMs and GLMMs are robust tools for modelling species distributions and they allow the generation of statistical functions that allow predictions of potentially suitable habitat distribution for species within a GIS framework (Guisan & Zimmermann 2000, Brotons et al. 2004, McAlpine et al. 2008). However, presence/absence data rely on the need to obtain standardized absence data and the reliance on arbitrary criteria of the minimum amount of survey effort required to ascertain the absence of the species of interest in the sampling unit being considered (Redfern et al. 2006). Other methods, including envelope models and presence-only models, have been widely used in the literature to generate habitat suitability maps without requiring standardized effort and presence data (Brotons et al. 2004). These novel modelling approaches may be particularly suitable for tracking data, due to the lack of standardized effort and absence data (Skov et al. 2008).

Comparisons of these novel approaches with the more established presence/absence models underscore the disparity between models requiring high quality presence/absence data (e.g. GLMs) and models where absences are inadequate or unavailable (e.g. Ecological Niche Factor Analysis, ENFA; Brotons et al. 2004). Comparing the predictions of both methods revealed that the presence/absence predictions were more accurate than presence-only predictions, particularly when species were using available habitats proportionally to their suitability, thus making the absence data reliable and useful to enhance model calibration (Brotons et al. 2004). Because presence-only models

are continuously being developed, newly available approaches (e.g. boosted regression trees or maximum entropy methods) have been shown to out-perform more established methods, such as GLMs and generalised additive models (GAMs) (Elith et al. 2006, S. J. Phillips et al. 2006). Ultimately, the type and quality of the data (e.g. presence/ absence or presence-only, availability of standardized effort data), and the inherent characteristics of the focal species (e.g. habitat specificity, degree of aggregation) influences the selection and performance of habitat modelling approaches.

Also, species with less restricted ecological requirements and/or distribution ranges are modelled less accurately than species with more restricted requirements/ranges (Segurado & Araújo 2004, but see Elith et al. 2006), irrespective of the modelling methods employed (Brotons et al. 2004). Our work revealed a similar result in the western Mediterranean, where our habitat suitability model developed for the critically endangered Balearic shearwater performed better (AUC  $\pm$  SE:  $0.8 \pm 0.006$ ) than for the more widespread Cory's shearwater (Louzao et al. 2006). The Balearic shearwater is currently listed as Critically Endangered on the IUCN Red List due to its restricted breeding range (constrained to the Balearic Islands) and small (ca. 2000 breeding pairs) and declining population (7.4% decrease per year; BirdLife International 2008). Despite the moderate predictive performance of our models, their predictions matched the observed foraging and feeding patterns. Moreover, the interpretation of the modelling output was consistent with the regional oceanography. The moderate discrimination ability of the models might reflect the potential exclusion of an important explanatory variable or the non-linear relationship with a critical habitat variable (McAlpine et al. 2008). Including different seasons and years (when long-term data are available) will help to overcome some of the habitat modelling limitations by validating and refining current habitat suitability models. Future research should also focus on obtaining a mechanistic understanding of the small-scale interactions between local oceanographic conditions, prey distribution, and the aggregation patterns and behaviour of Cory's shearwaters within the high-use feeding hotspots identified in the present study.

#### **Combining vessel-based surveys and tracking data: conservation implications**

This study demonstrates how the combination of vessel-based surveys and tracking data provides a wider understanding of the predictability of aggregation (i.e. hotspots) and the key oceanographic habitats of far-

ranging seabirds at different spatial scales (Rodhouse et al. 1996, BirdLife International 2004). While seabird conservation initiatives have independently addressed seascape patterns from either vessel-based surveys (e.g. Louzao et al. 2006) or tracked individuals (e.g. R. A. Phillips et al. 2006, González-Solís et al. 2007, Skov et al. 2008), few studies have integrated these distinct seascape and individual perspectives due to both logistical limitations (e.g. lack of concurrent tracking and survey data; but see Hyrenbach & Dotson 2003, Hyrenbach et al. 2006) and specific characteristics of both disparate datasets. It is critical to acknowledge the advantages and disadvantages of both approaches in order to reinforce the lessons learnt from these 2 perspectives and to reconcile potential discrepancies.

Overall, vessel-based surveys provide a large-scale (100s to 1000s of kilometres) perspective of the population level distribution and habitat associations, although constrained, to a varying degree, by the ship-following behaviour of certain species and the inability to determine the origin, sex, and reproductive status of observed individuals, as well as by the small survey area covered by vessel-based surveys. On the other hand, tracking data provide fine-scale (0.1s to 10s of kilometres) and detailed information at the individual level, due to the advent of GPS-loggers with small positional errors (0.01s of kilometers) and a higher rate of data acquisition (sampling intervals of seconds to minutes), compared to satellite-linked platform transmitter terminals (PTTs) (errors in the order of 100s of metres to 10s of kilometres and a sampling interval of hours) (Weimerskirch 2007). Nevertheless, the tracking perspective is often limited to small sample sizes (few tagged individuals tracked over short time periods) due to the cost of both devices and fieldwork for deployment/retrieval of tags. Therefore, these distribution data may not be representative of the population as a whole, owing to variation in ranging behaviour according to specific characteristics such as colony of origin, age, sex, breeding status, and season (e.g. BirdLife International 2004).

The commonly used variables for key marine area identification include the overall number of birds in a given area (i.e. density) and their behaviour (i.e. proportion of foraging versus flying birds). While these metrics are easier to infer from landscape-level vessel-based surveys than from tracking studies of individual foragers, when loggers of low positional error and high rate of data acquisition are used, researchers can use animal behaviour to infer finer-scale habitat use patterns, such as migration corridors and feeding grounds (Weimerskirch 2007). Other available methodologies, such as land-based counts, are also key approaches for monitoring migration flows, mostly through narrow cor-



ridors defined by topographical features (e.g. the Strait of Gibraltar), which concentrate large numbers of migrating seabirds (del Hoyo et al. 1992). Yet, while these methods provide information on the timing and magnitude of the flow through, they cannot provide seascape level data on the distribution of these marine top predators (e.g. habitat associations, background densities).

Once the location and extent of key seabird habitat areas are identified, the next step is to ensure their legal protection. In the European Union, the Natura 2000 network provides the opportunity of effectively protecting the identified marine IBAs through their designation as Special Protection Areas (SPAs) within the EC Bird Council Directive 79/409/EEC for the protection of wild birds. Beyond the identification and designation process, further effort must be directed towards the establishment of long-term monitoring programmes for assessing significant threats affecting seabirds and to develop the appropriate management plans in the protected areas. Because some threats faced by seabirds are to some extent diffuse and widespread (e.g. oil spills), any spatially explicit conservation action (e.g. designation of SPAs) needs to be framed within a larger conservation approach, addressing the status and the threats to the species through its range and life cycle. For instance, Cory's shearwaters are highly susceptible to longline fisheries in the western Mediterranean (Belda & Sánchez 2001), where the implementation of both an observer programme and mitigation measures on the fishing vessels are critical for the conservation of the species. Both measures should be implemented within the whole range of the species, controlling longlining to a greater degree and intensifying the monitoring efforts within the key marine areas for the species. Ultimately, we believe that spatially explicit protective measures will help to catalyze a comprehensive ecosystem-based management approach to protect marine far-ranging species of conservation concern.

*Acknowledgements.* We thank many people their help in the development of this work: G. Peters and his team at Earth & Ocean Technologies for GPS design; L. Gil de Sola, M. García, E. Massutí and P. Abelló for help and support during the MEDITS cruise (Spanish Institute of Oceanography, IEO); R. Escandell (SOM), D. Oro and J. M. Igual (IMEDEA), C. Viada, D. Fernández, X. Larruy, A. Requejo, E. Miralles and J. Prieto for their collaboration with the GPS-tracking (GPSs); and D. Villero, L. Brotons and M. Plà (CTFC) for input on modelling. The MODIS/Aqua SST data were obtained through the online PO.DAAC Ocean ESIP Tool (POET) at the Physical Oceanography Distributed Active Archive Center (PO.DAAC), NASA Jet Propulsion Laboratory, Pasadena, CA and CHL data were obtained from the NOAA CoastWatch, West Coast Regional Node. Thanks also to J. Roberts (MGET) and to L. Spence for GIS assessment. K. Laneri, A. Martínez-Abraín, A. Goarant, C. Peron, M. Authier, P. Inchausti and D.

Pinaud provided valuable statistical advice. Fieldwork was conducted within the framework of Project LIFE04NAT/ES/000049, funded by the European Commission and the Spanish Ministry of Environment (MARM), and executed by SEO/BirdLife. KDH was supported by a Pew fellowship in marine conservation. ML was funded by a postdoctoral contract of the Spanish Ministry of Education and Science (Ref. EX-2007-1148) and a Marie Curie Individual Fellowship (PIEF-GA-2008-220063).

#### LITERATURE CITED

- Abelló P, Arcos JM, De Sola LG (2003) Geographical patterns of seabird attendance to a research trawler along the Iberian Mediterranean coast. *Sci Mar* 67:69–75
- Alpine JE, Hobday AJ (2007) Area requirements and pelagic protected areas: Is size an impediment to implementation? *Mar Freshw Res* 58:558–569
- Arcos JM, Oro D (2002) Significance of fisheries discards for a threatened Mediterranean seabird, the Balearic shearwater *Puffinus mauretanicus*. *Mar Ecol Prog Ser* 239:209–220
- Arcos JM, Bécarea J, Rodríguez B, Ruiz yA (2009) Áreas Importantes para la Conservación de las Aves marinas en España. LIFE04NAT/ES/000049-Sociedad Española de Ornitología (SEO/Bird-Life), Madrid
- Belda EJ, Sánchez A (2001) Seabird mortality on longline fisheries in the western Mediterranean: factors affecting bycatch and proposed mitigating measures. *Biol Conserv* 98:357–363
- Bertrand JA, Gil de Sola L, Papaconstantinou C, Relini G, Souplet A (2002) The general specifications of the MEDITS surveys. *Sci Mar* 66:9–17
- BirdLife International (2004) Tracking ocean wanderers: the global distribution of albatrosses and petrels. Results from the Global Procellariiform Tracking Workshop, 1 to 5 Sep, 2003, Gordon's bay, South Africa. BirdLife International, Cambridge
- BirdLife International (2008) Species factsheet: *Puffinus mauretanicus*. <http://www.birdlife.org>
- Boyd IL, Wanless S, Camphuysen CJ (2006) Top predators in marine ecosystems: their role in monitoring and management. Cambridge University Press, Cambridge
- Brotons L, Thuiller W, Araújo MB, Hirzel AH (2004) Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. *Ecography* 27:437–448
- Buckland ST, Burnham KP, Augustin NH (1997) Model selection: an integral part of inference. *Biometrics* 53:603–618
- Burnham KP, Anderson DR (2002) Model selection and multimodel inference: a practical Information-Theoretic approach. Springer Verlag, New York
- Burnham KP, Anderson DR (2004) Multimodel inference: understanding AIC and BIC in model selection. *Sociol Methods Res* 33:261–304
- Carboneras C (2004) Pardela Cenicienta *Calonectris diomedea diomedea*. In: Madroño A, González C, Atienza JC (eds) Libro Rojo de las Aves de España. Dirección General para la Biodiversidad-SEO/BirdLife, Madrid
- Cayula JF, Cornillon P (1992) Edge detection algorithm for SST images. *J Atmos Ocean Technol* 9:67–80
- Crawley MJ (1993) GLIM for ecologists. Blackwell Scientific, Oxford
- del Hoyo J, Elliott A, Sargatal J (1992) Handbook of the birds of the world. Lynx Edicions, Barcelona
- Elith J, Graham CH, Anderson RP, Dudik M and others (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29:129–151

- Fielding AH, Bell JF (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environ Conserv* 24:38–49
- Franks PJS (1992) Sink or swim: accumulation of biomass at fronts. *Mar Ecol Prog Ser* 82:1–12
- Freeman EA, Moisen GG (2008) A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and kappa. *Ecol Model* 217: 48–58
- González-Solís J, Croxall JP, Oro D, Ruiz X (2007) Trans-equatorial migration and mixing in the wintering areas of a pelagic seabird. *Front Ecol Environ* 5:297–301
- Gray TNE, Chamnan H, Borey B, Collar J, Dolman PM (2007) Habitat preferences of a globally threatened bustard provide support for community-based conservation in Cambodia. *Biol Conserv* 138:341–350
- Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology. *Ecol Model* 135:147–186
- Hammer WH, Schneider DC (1986) Regularly spaced rows of medusae in the Bering Sea: Role of Langmuir circulation. *Limnol Oceanogr* 31:171–177
- Hanley JA, McNeil BJ (1982) The meaning and use of the area under a receiver operating characteristics (ROC) curve. *Radiology* 143:29–36
- Harris J, Haward M, Jabour J, Woehler EJ (2007) A new approach to selecting Marine Protected Areas (MPAs) in the Southern Ocean. *Antarct Sci* 19:189–194
- Haury LR, McGowan JA, Wiebe PH (1978) Patterns and processes in the time-space scales of plankton distributions. In: Steele JH (ed) *Spatial patterns in plankton communities*. Plenum Press, New York
- Hooker SK, Gerber LR (2004) Marine reserves as a tool for ecosystem-based management: the potential importance of megafauna. *Bioscience* 54:27–39
- Hunt GL, Schneider DC (1987) Scale-dependent processes in the physical and biological environment of marine birds. In: Croxall JP (ed) *The feeding ecology of seabird and their role in marine ecosystems*. Cambridge University Press, Cambridge
- Hunt GL, Mehlum F, Russell RW, Irons D, Decker MB, Becker PH (1999) Physical processes, prey abundance, and the foraging ecology of seabirds. *Proc 22 Int Ornithol Cong* 22: 2040–2056
- Hurlbert SH (1984) Pseudoreplication and the design of ecological field experiments. *Ecol Monogr* 54:187–211
- Hyrenbach KD, Dotson RC (2003) Assessing the susceptibility of female black-footed albatross (*Phoebastria nigripes*) to longline fisheries during their post-breeding dispersal: an integrated approach. *Biol Conserv* 112:391–404
- Hyrenbach KD, Forney KA, Dayton PK (2000) Marine protected areas and ocean basin management. *Aquat Conserv: Mar Freshwat Ecosyst* 10:437–458
- Hyrenbach KD, Fernández P, Anderson DJ (2002) Oceanographic habitats of two sympatric North Pacific albatrosses during the breeding season. *Mar Ecol Prog Ser* 233: 283–301
- Hyrenbach KD, Keiper C, Allen SG, Ainley DG, Anderson DJ (2006) Use of marine sanctuaries by far-ranging predators: commuting flights to the California Current System by breeding Hawaiian albatrosses. *Fish Oceanogr* 15:95–103
- Igual JM, Forero MG, Tavecchia G, González-Solís J and others (2005) Short-term effects of data-loggers on Cory's shearwater (*Calonectris diomedea*). *Mar Biol* 146:619–624
- Jenness J (2004) Nearest features (nearfeat. avx) extension for ArcView 3.x, v. 3.8a. Available online at: [www.jennessent.com/arcview/nearest\\_features.htm](http://www.jennessent.com/arcview/nearest_features.htm)
- Johnson JB, Omland KS (2004) Model selection in ecology and evolution. *Trends Ecol Evol* 19:101–108
- Johnston DW, Westgate AJ, Read AJ (2005) Effects of fine-scale oceanographic features on the distribution and movements of harbour porpoises *Phocoena phocoena* in the Bay of Fundy. *Mar Ecol Prog Ser* 295:279–293
- Lloret J, Palomera I, Salat J, Sole I (2004) Impact of freshwater input and wind on landings of anchovy (*Engraulis encrasicolus*) and sardine (*Sardina pilchardus*) in shelf waters surrounding the Ebre (Ebro) River delta (north-western Mediterranean). *Fish Oceanogr* 13:102–110
- Louzao M, Hyrenbach KD, Arcos JM, Abelló P, Gil de Sola L, Oro D (2006) Oceanographic habitat of a critically endangered mediterranean procellariiform: implications for the design of marine protected areas. *Ecol Appl* 16:1683–1695
- McAlpine CA, Rhodes JR, Bowen ME, Lunney D, Callaghan JG, Mitchell DL, Possingham HP (2008) Can multiscale models of species' distribution be generalized from region to region? A case study of the koala. *J Appl Ecol* 45: 558–567
- Millot C (1999) Circulation in the Western Mediterranean Sea. *J Mar Syst* 20:423–442
- Murray K, Conner MM (2009) Methods to quantify variable importance: implications for the analysis of noisy ecological data. *Ecology* 90:348–355
- Orians GH, Pearson NE (1979) On the theory of central place foraging. In: Horn DJ, Stairs GR, Mitchell RD (eds) *Analysis of ecological systems*. Ohio State University Press, Columbus, OH, p 155–177
- Phillips RA, Xavier JC, Croxall JP (2003) Effects of satellite transmitters on albatrosses and petrels. *Auk* 120: 1082–1090
- Phillips RA, Silk JRD, Croxall JP, Afanasyev V (2006) Year-round distribution of white-chinned petrels from South Georgia: relationships with oceanography and fisheries. *Biol Conserv* 129:336–347
- Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecol Modell* 190:231–259
- Ramírez I, Geraldés P, Meirinho A, Amorim P, Paiva V (2008) Áreas Marinhas Importantes para as Aves em Portugal. Projecto LIFE04NAT/PT/000213-Sociedade Portuguesa Para o Estudo das Aves, Lisboa
- R Development Core Team (2008) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. [www.r-project.org](http://www.r-project.org)
- Redfern JV, Ferguson MC, Becker EA, Hyrenbach KD and others (2006) Techniques for cetacean-habitat modelling. *Mar Ecol Prog Ser* 310:271–295
- Ribic CA, Chapman E, Fraser WR, Lawson GL, Wiebe PH (2008) Top predators in relation to bathymetry, ice and krill during austral winter in Marguerite Bay, Antarctica. *Deep-Sea Res II* 55:485–499
- Richards SA (2005) Testing ecological theory using the Information-Theoretic approach: examples and cautionary results. *Ecology* 86:2805–2814
- Rodhouse PG, Prince PA, Trathan PN, Hatfield EMC and others (1996) Cephalopods and mesoscale oceanography at the Antarctic Polar Front: satellite tracked predators locate pelagic trophic interactions. *Mar Ecol Prog Ser* 136:37–50
- Sawada M (1999) ROOKCASE: an EXCEL 97/2000 VISUAL BASIC (VB) add-in for exploring global and local spatial autocorrelation. *Bull Ecol Soc Am* 80:231–234
- Segurado P, Araújo MB (2004) An evaluation of methods for modelling species distributions. *J Biogeogr* 31: 1555–1568

- Skov H, Humphreys E, Garthe S, Geitner K and others (2008) Application of habitat suitability modelling to tracking data of marine animals as a means of analyzing their feeding habitats. *Ecol Modell* 212:504–512
- Sokal RR, Oden NL, Thomson BA (1998) Local spatial autocorrelation in biological variables. *Biol J Linn Soc* 65: 41–62
- Swets JA (1988) Measuring the accuracy of diagnostic systems. *Science* 240:1285–1293
- Tasker ML, Jones PH, Dixon T, Blake BF (1984) Counting seabirds at sea from ships: a review of methods employed and a suggestion for a standardized approach. *Auk* 101: 567–577
- Weimerskirch H (2007) Are seabirds foraging for unpredictable resources? *Deep-Sea Res II* 54:211–223
- Whitmire AL, Boss E, Cowles TJ, Pegau WS (2007) Spectral variability of the particulate backscattering ratio. *Opt Express* 15:7019–7031
- Zuur AF, Ieno EN, Smith GM (2007) *Analysing ecological data*. Springer, New York

*Editorial responsibility: Jacob Gonz ales-Sol s,  
Barcelona, Spain*

*Submitted: November 17, 2008; Accepted: May 21, 2009  
Proofs received from author(s): August 27, 2009*