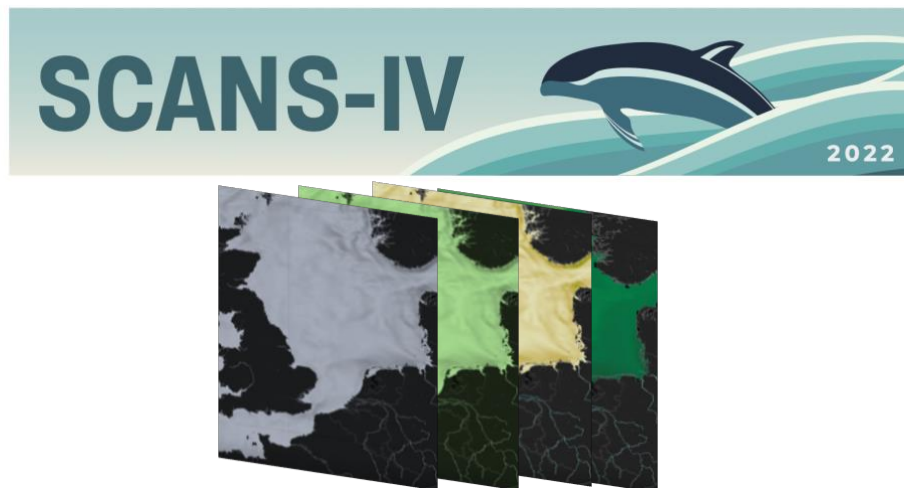


Spatial models of cetacean density in European Atlantic waters based on SCANS-IV summer 2022 survey data

A Gilles¹, M Authier², R Pigeault¹, N Ramirez-Martinez¹, V Benoit¹, J Carlström³, C Eira⁴, SCV Geelhoed⁵, S Laran², M Sequeira⁶, S Sveegaard⁷, NL Taylor⁸, C Saavedra⁹, JA Vázquez-Bonales⁹, PS Hammond¹⁰

1. Institute for Terrestrial and Aquatic Wildlife Research, University of Veterinary Medicine Hannover, Foundation, Buesum, Germany
2. Observatoire Pelagis, UAR 3462, CNRS-La Rochelle University, La Rochelle, France
3. Department of Environmental Research and Monitoring, Swedish Museum of Natural History, Stockholm, Sweden
4. Department of Biology & CESAM & ECOMARE, Aveiro University, Portugal
5. Wageningen Marine Research, Den Helder, The Netherlands
6. Instituto da Conservação da Natureza e Florestas (ICNF), Lisbon, Portugal
7. Department of Ecoscience, Aarhus University, Roskilde, Denmark
8. Joint Nature Conservation Committee, UK
9. Oceanographic Centre of Vigo, Spanish Institute of Oceanography, Spanish National Research Council (IEO-CSIC), Vigo, Spain
10. Sea Mammal Research Unit, University of St Andrews, UK



Please cite as:

Gilles, A, Authier, M, Pigeault, R, Ramirez-Martinez, NC, Benoit, V, Carlström, J, Eira, C, Geelhoed, SCV, Laran, S, Sequeira, M, Sveegaard, S, Taylor, NL, Saavedra, C, Vázquez-Bonales, JA, Hammond, PS (2025). Spatial models of cetacean density in European Atlantic waters based on SCANS-IV summer 2022 survey data. Final report published 14 May 2025. 31 pp plus Appendix. <https://tinyurl.com/3rv246v5>

Contents

INTRODUCTION	3
METHODS.....	4
Data preparation and processing.....	4
Data analysis	6
Model structure, fitting and selection	7
Model application	9
Model evaluation	10
RESULTS.....	11
Searching effort and sightings	11
Model results and predicted density surfaces	12
DISCUSSION	21
Harbour porpoise	21
Bottlenose dolphin	22
White-beaked dolphin	23
Common and striped dolphin	24
Beaked whales	24
Fin whale.....	25
Minke whale	25
ACKNOWLEDGMENTS	27
REFERENCES	28
APPENDIX.....	32
Deviance residuals.....	47
Rootograms	49
REFERENCES	51

INTRODUCTION

Large-scale line-transect surveys of cetaceans in European Atlantic waters (known as SCANS) began in 1994, at that time covering the North Sea, parts of the Celtic Sea and adjacent waters (Hammond et al., 2002) with the primary goal to obtain the first comprehensive abundance estimate of harbour porpoise to contextualize bycatch estimates. These regional coordinated survey efforts continued in 2005 with SCANS-II, expanding the survey to all shelf waters (Hammond et al., 2013), and in 2007, extending to offshore waters (CODA, 2009). Subsequent SCANS surveys covering a much-expanded area followed in 2016 (Hammond et al., 2021) and 2022, with the ObSERVE project covering Irish waters in 2016 (Rogan et al., 2018) and 2022 (Giralt Paradell et al., 2024). By now, these surveys have yielded a time series of data on regularly occurring cetacean species, thus facilitating investigation of changes in distribution and abundance at an ecologically appropriate large spatial (European Atlantic) and temporal (almost three decades) scales for long-lived mobile species. These surveys have set high standards for data collection and analysis.

The SCANS surveys aim to provide critical information on population abundance and trend of cetacean species for statutory reporting and assessment needs across the NE Atlantic region, supporting EU Member States in reporting on Favourable Conservation Status (FCS) under the Habitats Directive (HD) and Good Environmental Status (GES) under the Marine Strategy Framework Directive (MSFD, Article 17) or national equivalent for non-EU Member States. The SCANS surveys also facilitate delivery of coherent and consistent marine mammal assessments under the Regional Sea Conventions OSPAR and HELCOM.

The SCANS projects have been implemented at approximately decadal intervals during the summer months. However, to inform relevant reporting and ecosystem assessments, the frequency was changed to six years, which is the reason why the fourth iteration of SCANS was conducted in the summer of 2022. SCANS-IV covered a 1.7 million km² study area stretching from the Strait of Gibraltar to southern Norway, providing robust abundance estimates and trends for regularly occurring cetacean species. The conventional design-based estimates of abundance are presented in Gilles et al. (2023), which are critical to inform assessments of the impact of anthropogenic activities, such as fisheries (e.g. by-catch) and offshore industries, especially renewable energy.

In the context of marine spatial planning, the outputs of habitat-based density models or species distribution models (SDMs) are increasingly used for marine management and conservation applications, including the assessment of potential impacts from a wide range of anthropogenic activities (Hammond et al., 2013; Gilles et al., 2016, Lacey et al., 2022, Pigeault et al., 2024a). The effectiveness of SDMs as conservation management tools is attributed to their capacity to predict spatial and temporal changes in species distribution patterns. Consequently, another major aim of SCANS-IV, and all preceding surveys, was to supply information on summer distribution by modelling the data in relation to environmentally-linked spatial features to generate density surface maps.

In this report, we summarise the development and results of predictive habitat-based models of cetacean density for seven species and two species groups using survey effort and sighting data from the SCANS-IV surveys conducted during summer 2022.

METHODS

Data preparation and processing

The SCANS-IV study area and survey design (see **Figure 1**) as well as the aerial and ship survey data collection methods are described in detail in Gilles et al. (2023). The survey effort and cetacean sighting data used for the species distribution modelling were the same as those used to derive the design-based estimates of abundance, as presented in the first SCANS-IV report (Gilles et al., 2023).

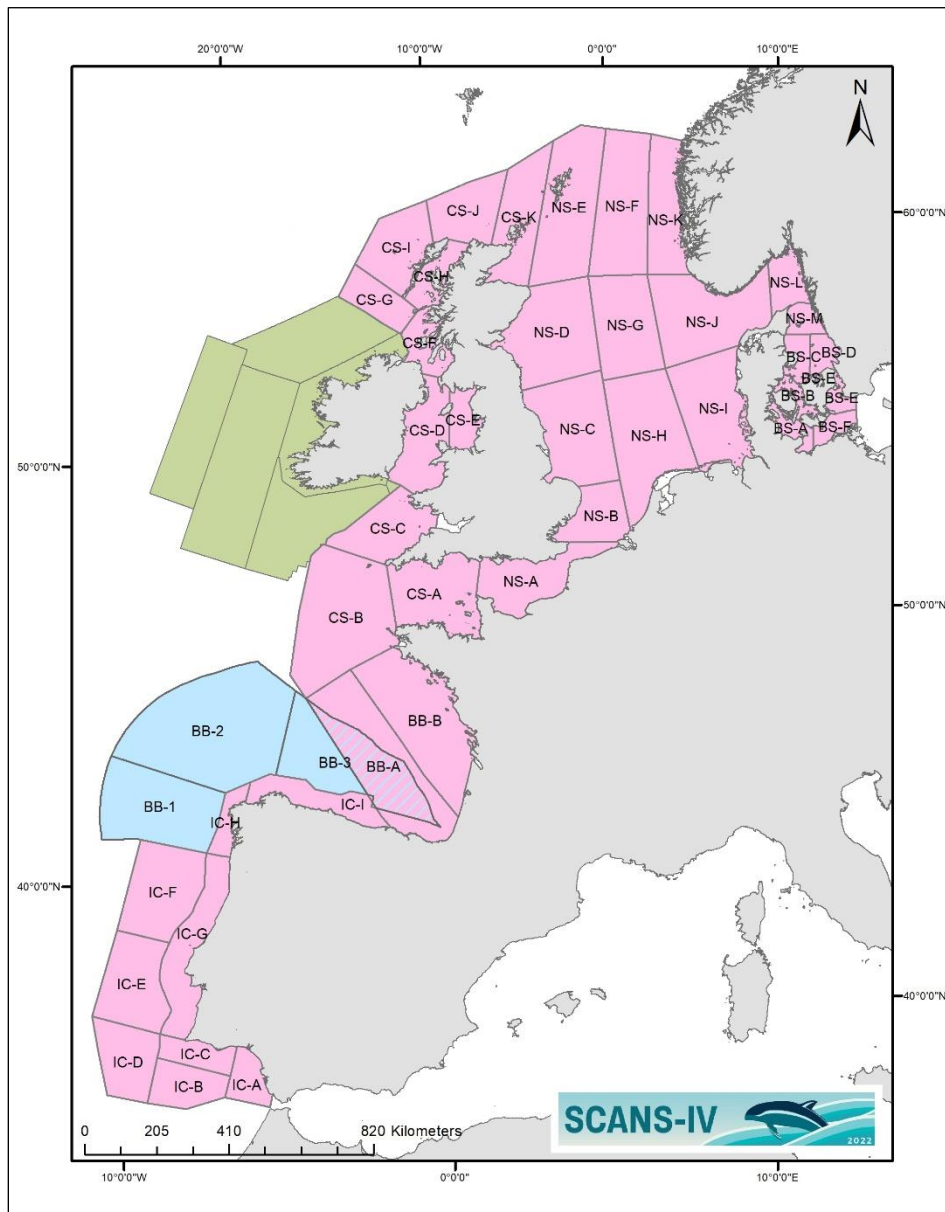


Figure 1. Area covered by SCANS-IV (see Gilles et al., 2023): pink blocks were surveyed by air and blue blocks were surveyed by ship. The cross-hatched area is where the ship survey BB-3 and aerial survey block BB-A overlapped in an area of 39,018 km². Blocks coloured green to the south and west of Ireland were surveyed by the Irish ObSERVE2 project (see Giralt Paradell et al., 2024).

SCANS-IV aerial and ship line-transect survey effort data (**Figure 1**; Gilles et al., 2023) were segmented into continuous portions of effort of approximate 10 km mean length, conforming with Becker et al. (2020), Gilles et al. (2016) and Virgili et al. (2019). A tolerance of 10 % was applied to the segment lengths as not all segments could be 10 km long (e.g. due to the end of the transect or short on-effort legs with a leg defined as a portion of a transect which was surveyed with homogenous sighting conditions). Segments were defined per transects and dates for on-effort legs in good and moderate sighting conditions only, and which were no more than 1 km from other legs. The minimum segment length was set at 5 km, and the remaining distance was randomly assigned to one of the segments if the length was less than this value. As some on-effort legs in good and moderate conditions were distant by more than 1 km from other legs and conducted over short distances (e.g. between two legs in poor conditions or between islands), some final segments were shorter than 5 km. Only segments longer than 2 km were retained for the modelling (i.e., 99.4 % of the segments), and the final mean length was 9.9 km (SD = 1.4 km), with lengths ranging from 2.0 to 18.9 km with a Gaussian-shaped distribution.

Species-specific sighting data were assigned to each segment. The effective area searched was estimated for each segment and for each species or species group, based on the species-specific effective strip widths (including $g(0)$) as reported in Gilles et al., 2023). The effective area searched was subsequently included as an offset in the model structure (see equation 1). This procedure accounts for both varying segment lengths and the different detection probabilities recorded during the surveys. The covariates were extracted at a daily resolution within a buffer of 5 km around the segments (see **Table 1** or candidate covariates). A suite of environmental covariates, i.e. spatial, static and dynamic covariates, were considered. It can be argued that the selected habitat predictors are most likely proxies for unmeasured underlying ecological processes driving species distributions rather than direct drivers.

In order to avoid leaking of density over land and smoothing over impassable boundaries for cetaceans, locational covariates X and Y were included in all models with a soap-film smoother (Wood, 2008). The number of knots in the soap-film proved challenging to determine, with little guidance in the published literature. Different values for the number of knots were thus tested and compared using various diagnostics (e.g. rootograms; see below). Following a comprehensive evaluation of the available options, it was determined that utilising 80 knots in the soap-film smoother would offer an optimal balance between good model performance and mitigating the risk of overfitting.

To create the soap-film smoother, the land contours of the study area were used and adapted to not cover any of the segments prepared for modelling. As complex land contours are difficult to fit and lead to complex and time-intensive spatial smooth functions, the land contours were simplified as far as possible, and only the contours of Great Britain and the Hebrides were retained as offshore islands within the soap smoother. The Hebrides were especially used as higher abundance and marine diversity is generally observed in this region. The knots were placed at equal distances over the area covered by the soap-film smooth.

The prediction grid was first delineated along the boundaries of the approximate 1.7 million km² study area to ensure the validity of predictions, which were not to be geographically extrapolated outside the region utilised for the development of the model. Daily prediction grids, comprising 10×10 km cells, were constructed from the initial day to the concluding day of the SCANS-IV survey, i.e. from 28 June to 22 October 2022. However, the primary survey period was from 28 June to 15 August 2022, during which a substantial proportion (83%) of the effort was accomplished (see Gilles et al., 2023 for details). The daily covariate means were calculated for each grid cell.

Table 1. Candidate environmental covariates used in the density surface models for SCANS-IV data, shown to be important in previous models (e.g. Gilles et al., 2016, Lacey et al., 2022).

Covariate	Description	Source
X	Longitude converted to ETRS89 (EPSG:3035)	
Y	Latitude converted to ETRS89 (EPSG:3035)	
Water depth (Depth)	Mean water depth (m)	EMODnet Digital Bathymetry (DTM 2020).
Slope	Slope of the seabed (°) calculated with R package <i>raster</i> , version 3.6-26 (Hijmans, 2023).	Derived from bathymetry.
Aspect	The direction that the slope faces (°) calculated with R package <i>raster</i> , version 3.6-26 (Hijmans, 2023).	
Distance from coast (dist_coast)	Distance at segment centroid (m)	EMODnet Bathymetry World Coastline, estimated at highest astronomical tide.
Distance from 50m isobath (dist_50m)	Distance at segment centroid (m)	Isobath contours were provided by EMODnet Bathymetry and generated based on the EMODnet DTM.
Distance from 200m isobath (dist_200m)*	Distance at segment centroid (m)	
Distance from 2000m isobath (dist_2000m)*	Distance at segment centroid (m)	
Mean sea surface temperature (mSST)	Daily temperature (°C)	Global Ocean Physics Reanalysis. E.U. Copernicus Marine Service Information (CMEMS).
Sea surface temperature gradient (gradSST)	Spatial gradient in daily temperature (°C) within the radius of one cell, calculated with R package <i>grec</i> , version 1.6.0 (Lau-Medrano, 2024)	
Eddy kinetic energy (EKE)	Eddies calculated as the current velocity (m/s)	
Net primary productivity (NPPV)	Expressed as carbon per unit volume in sea water (mg/m ³ /day)	Global Ocean Biogeochemistry Hindcast. E.U. Copernicus Marine Service Information (CMEMS)

*Not considered for harbour porpoise.

Data analysis

All data processing was undertaken in software R version 4.4.0 (R Core Team, 2024), and modelling was conducted using R package *mgcv*, version 1.9-1 (Wood, 2017). Density surfaces and coefficients of variation (CVs) were plotted on maps using software ArcGIS (Esri ArcGIS Pro 3.4.2). The same colour scale (min: blue; high: orange) was utilized for all species to facilitate comparisons of high versus low usage areas. However, due to the varying species-specific density values, the legend scales in the maps have different breakpoints.

Model structure, fitting and selection

A multi-stage modelling approach was implemented with the objective of reducing bias in the density estimates generated from the habitat models. Methods largely followed those described in Gilles et al. (2016), Becker et al. (2016; 2019) and Lacey et al. (2022), with the addition of a summary of pertinent and novel aspects.

The general structure of the Generalised Additive Models (GAM), using a logarithmic link function, was:

$$\begin{cases} n_i \sim \text{NegBin}(\omega, \eta_i) \\ \log(\eta_i) = \log(a_i) + \vartheta_0 + \sum_{k=1}^p f_k(z_{ki}) \end{cases} \quad (\text{Eq. 1})$$

where the response variable n_i is the number of individuals detected in the i th effort segment and η_i is the linear predictor. The offset a_i is the effective area searched for the i th segment, ϑ_0 is the intercept, $f_k(\cdot)$ are smoothed functions (splines) of the explanatory environmental covariates, and z_{ki} is the value of the k th explanatory covariate in the i th segment. A negative binomial likelihood (NegBin) was assumed for the response variable n_i to account for over-dispersion in the count data, with the over-dispersion parameter (ω) estimated during model fitting.

Smooth functions were fitted using restricted maximum likelihood (REML) with automatic term selection (Marra & Wood, 2011). Cubic regression splines were used for all covariates, with a maximum number of knots set to 10. For the covariate ‘aspect’, a cyclic cubic regression spline was used.

This model-fitting method helps to avoid overfitting of the smooth functions by including a penalization (Marra & Wood, 2011). The method can reduce the estimated degrees of freedom of a covariate term towards zero if it does not contribute sufficiently to account for the variability in the data. Consequently, covariates with estimated degrees of freedom close to 0 were not removed from a single model and all covariates, including those with estimates penalized down to 0, were retained. For each species, models were fitted for each possible combination of two to five uncorrelated covariates (that is covariates with a Pearson’s pairwise correlation coefficient < 0.50). The five models with the best goodness of fit, based on leave-one-out cross-validation, were selected and their respective predictions were stacked (Yao et al. 2019) for further investigation, with their respective contribution to the final prediction estimated with the *loo* R-package (Vehtari et al., 2017). QQ plots, rootograms (Kleiber & Zeilis, 2016), degrees of freedom, fitted relationships, predicted species distributions and abundances were inspected for this selection of models. Goodness-of-fit and model performance diagnostics were overall consistent among selected models, and the model contributing the most to the final prediction was finally selected for each species or species group.

This two-tiered model selection process enabled us to assess and account for model uncertainty in predicting cetacean density surfaces with stacking predictions from five different models in the first step; and to ease interpretation of important environmental covariates in the second step by focusing on the model with the highest contribution to the stacked predictions.

For dolphin species (common dolphin, striped dolphin, white-beaked dolphin and bottlenose dolphin), the single response model using the number of individuals per transect segment did not fit the data well because the greater range of group sizes recorded for these species, compared with harbour porpoise and whales, caused severe over-dispersion in the distribution of counts of individuals. For these species, a two-step modelling process was used. This process first modelled the group density (per squared km), and then secondly, conditional on sightings, modelled the average group size per sighting (i.e. number of individuals). Thus, in the first step, the number of sightings of groups in each effort segment was modelled as the response variable, rather than number of individuals:

$$\begin{cases} g_i \sim \text{NegBin}(\omega, \eta_i) \\ \log(\eta_i) = \log(a_i) + \vartheta_0 + \sum_{k=1}^p f_k(z_{ki}) \end{cases} \quad (\text{Eq. 2.1})$$

where the response variable g_i is the number of groups detected in the i th effort segment. Separate group size models were developed in the second step, as a function of locational coordinates X and Y only, applying the soap-film smoother. These group size models were fitted with only those effort segments that included sightings, using a negative binomial (Eq. 2.2). The general structure of the GAM, using a logarithmic link function, was:

$$\begin{cases} (n_i - 1) \sim \text{NegBin}(\omega_g, \eta_{i,g}) \\ \log(\eta_{i,g}) = \log(g_i) + \vartheta_{0,g} + \text{soap}_{\text{smooth}}(X_i, Y_i) \end{cases} \quad (\text{Eq. 2.2})$$

where the response variable is the number of individuals detected in the i th effort segment, n_i , minus 1 (to set the lowest possible value to 0, as only segments with sightings were used in this model) and the linear predictor is $\eta_{i,g}$. The offset g_i is the number of groups observed in the i th segment, $\vartheta_{0,g}$ is the intercept, and $\text{soap}_{\text{smooth}}(\cdot)$ is the smoothed function used for the locational coordinates X and Y segment in the soap-film smoother. A negative binomial likelihood (NegBin) was assumed for the response variable to account for over-dispersion in the number of individuals, as dolphin species may aggregate in large groups, and the over-dispersion parameter (ω_g) was estimated during model fitting. The number of groups (Eq. 2.1) and the group sizes (Eq. 2.2) were predicted for each cell in the study area, and their product generated estimates of individual density, equivalent to estimates from models of individuals for the other four species (Eq. 1).

For accurate uncertainty quantification, a pseudo-posterior approach was taken (e.g. King et al., 2000). The pseudo-posterior approach allows for seamless quantification of uncertainty for any derived quantities (e.g. abundance) from model parameters (Eq. 1). Maximum likelihood estimates of parameters and their associated covariance matrix were extracted from fitted models (using the function *rmvnorm* from package *mvtnorm*; Genz & Brentz, 2009) and used to generate a sample of 1,000 values from a pseudo-posterior, assuming a multivariate normal distribution for the parameters (King et al., 2000). This sample was used to carry out predictions at a daily level over the survey period. As the variability of some smooth functions was high for species with limited number of sightings or different ecotypes (e.g. bottlenose dolphins) and the upper values could reach extreme densities that are ecologically unrealistic due to the over-dispersion parameter, a threshold in densities was defined per species or species group. This limit was set to the 99.9% quantile of all the initial predicted species densities (i.e. by sample, cell and day). Densities above this limit were removed from the samples as they could not be considered as ecologically realistic. Finally, the predicted densities were averaged over the survey period for each cell and sample, providing a pseudo-posterior distribution per cell.

The pseudo-posterior approach that is used in this data analysis was recently shown to introduce a bias in the point estimate of any derived quantity from model parameters (Rainey, 2017; 2024). This bias can be separated into two components with the same magnitude and direction: “transformation bias” and “simulation bias”. Note that the bias-correction does not affect estimates of standard errors. The total bias was estimated per cell by predicting the cell densities with the *mgcv* R-package (Wood 2017) and averaging them over the survey period. The difference between the cell’s mean density produced by *mgcv* and the cell’s mean density from the pseudo posterior distribution is the “simulation bias”. This bias was multiplied by 2 for each cell to account for the “transformation bias” and corrected in the pseudo-posterior densities previously obtained. From this pseudo-posterior distribution, the cell’s mean density was calculated and its uncertainty estimated with the 2.5 % and 97.5 % quantiles. They represent the final results of the species distribution modelling undertaken in this study, and spatial grids containing these estimates were produced to plot the predicted species distribution over the SCANS-IV study area.

The model-based abundance estimates were checked for the selected models and used for model validation. To produce these model-based estimates, predictions were (i) summed per sample and day over the whole surveyed area or species-specific assessment units (e.g. for harbour porpoise), and (ii) averaged over the survey period to obtain pseudo-posterior distribution of model-based estimates of abundance. The “simulation bias” was calculated by 1) predicting the daily species cell abundance with *mgcv*, 2) summing them over the study area or species-specific assessment units, 3) averaging the total abundances over the survey period, 4) subtracting these averages to the pseudo posterior mean total abundance obtained in (ii). This bias was multiplied by 2 to include the “transformation bias” (Rainey 2017; 2024) and subtracted from the pseudo-posterior distribution model-based abundances. The mean abundance and coefficient of variations from this pseudo-posterior distribution were compared to the design-based abundances (Gilles et al., 2023).

Model application

Models were fitted to SCANS-IV data, collected in summer 2022, for harbour porpoise, bottlenose dolphin, white-beaked dolphin, common dolphin, striped dolphin, unidentified common or striped dolphin, common & striped dolphin (including unidentified common or striped dolphin), all beaked whale species combined, minke whale and fin whale.

For harbour porpoise modelling, data were restricted to survey blocks in shelf waters. Sightings of this species in deeper waters off the shelf are rare in the European Atlantic, as also shown during SCANS-III (Hammond et al., 2021). In SCANS-IV, no harbour porpoises were sighted in blocks BB1-BB3, IC-F, IC-E, IC-D and IC-B (**Figure 1**) (Gilles et al., 2023). For all other species, data collected in the Belt Seas and western Baltic Sea (i.e., blocks BS-A to BS-F) were excluded as these species are uncommon in this region and were not sighted during SCANS-IV.

In the case of striped dolphin, the initial models were fitted with all available data. However, subsequent analysis revealed that the group size model was irrelevant for the northern region of the study area, as no sightings were recorded. In addition, the group size model demonstrated a tendency to overcomplicate the soap-film smoother in this region since no data on group size were available, and the predicted group sizes were found to be highly extrapolated for some cells. The calibration data, soap-film smoother and predictions were therefore restricted to blocks in the southern study area: the Bay of Biscay (including offshore ship survey blocks), the Iberian Peninsula and the southern blocks of the Celtic Sea (i.e., CS-A and CS-B).

The initial models of white-beaked dolphin and bottlenose dolphin showed high variability and model-based abundance estimates not well aligned with the design-based abundance estimates (Gilles et al., 2023). Further investigation showed that the prediction of white-beaked dolphin distribution in September and October drove down the model-based abundance estimates. Since surveys covered mostly the Spanish coastal blocks over the course of these two months, and white-beaked dolphins were observed exclusively in the northern portion of the study area, the calibration data, soap-film smoother and predictions were constrained to the blocks in the Celtic Sea, Irish Sea, western Scotland and the North Sea for this species. Consequently, only predictions for the period 28 June to 31 August were used in the analysis. For bottlenose dolphins, the abundance was also highly variable over the time, and it was concluded that the presence of different ecotypes in the analysis induced variability in the smooth functions. Explanatory variables were therefore restricted to static variables, which limited the variability of the species distribution models. The group size model for bottlenose dolphins fitted complex smooth functions around the Hebrides Islands in the soap-film smoother, which predicted large group sizes. As these predicted group sizes exceeded twice the highest group size

observed, it was decided to simplify the soap-film smoother by removing the contours of the Hebrides Islands, which provided more realistic group sizes. For all other species, the soap-film smoother specification used for the models of number of groups was identical to that used for models of group sizes.

Given that no fin whales and 'unidentified common or striped dolphins' were observed around the Hebrides, using the contours of these islands in the soap-film smoother to model the distribution of these species groups has resulted in spatial smoothing functions that were unnecessarily complex and time-consuming to fit. The contours of the Hebrides were therefore removed from the soap-film smoother for these two groups.

For each species, spatially-explicit density values for the SCANS-IV study area were derived from model predictions using the selected best model on daily environmental conditions at a 10x10 km grid resolution.

Model evaluation

The performance of the model was evaluated using several established metrics. These included the percentage of explained deviance, deviance residuals, information criteria and visual inspection of predicted and observed distributions during the SCANS-IV surveys.

Rootograms were implemented as a model diagnostic tool to assess the goodness of fit of the model (Kleiber & Zeileis, 2016). This tool is particularly useful for diagnosing and treating issues such as overdispersion and/or excess zeros in count data models such as the ones presented here. The observed values of the response are compared with those expected from the fitted model, by grouping the data into bins. Rootograms can be estimated at the scale of the whole survey area, or at the scale of survey-block. Both were used to assess model goodness-of-fit, but the latter mostly to investigate the choice of knot number and locations in the soap-film smoother (not shown).

The potential for bias introduced by the habitat-based model was assessed by comparing the models' study area abundance estimates to conventional line-transect estimates derived from the same dataset used for modelling (and described in Gilles et al., 2023).

RESULTS

Searching effort and sightings

The final number of effort segments, groups and individuals sighted per segment are reported in **Table 2**.

A maximum of 7,152 effort segments were used to develop the habitat-based density models.

Except for harbour porpoise, common dolphin and the combined category common or striped dolphin, the percentage of segments with sightings of groups was small: around 2-3% for bottlenose dolphin, fin whale and white-beaked dolphin, and around 1% or fewer for minke whale and beaked whales, illustrating the extent of over-dispersion in the data (arising from a large fraction of segments with 0-sighting and a handful with many sightings).

Table 2. Number of effort segments and number of groups and individuals sighted of each species used in analysis of the SCANS-IV data in 2022. The lower number of effort segments for harbour porpoise, white-beaked dolphin and striped dolphin is due to exclusion of certain blocks with no sightings that are out of known species range (see explanation above).

Species	Total no. of effort segments	No. of effort segments with groups	% effort segments with groups	Number of groups	Number of individuals
Harbour porpoise	6,941	1,165	16.8	1,974	2,717
Bottlenose dolphin	7,152	218	3.0	361	1,549
White-beaked dolphin	4,190	66	1.6	97	511
Common dolphin	7,152	534	7.5	1,051	8,996
Striped dolphin	3,562	79	2.2	103	1,475
Unidentified common or striped dolphin	7,152	247	3.5	353	2,818
Common, striped and unid. common or striped dolphins	7,152	742	10.4	1,507	13,289
Beaked whales	7,152	36	0.5	43	75
Fin whale	7,152	191	2.7	297	346
Minke whale	7,152	62	0.9	70	73

Model results and predicted density surfaces

The soap smooth was retained in all models and accounted for the most estimated degrees of freedom in all models. The other selected covariates varied from species to species. In most cases, these covariates had much less estimated degrees of freedom attributed to them than to the soap smooth.

The final models explained 40% or more of the deviance for the majority of species. Striped dolphins and minke whales were the species with the lowest explained deviance, under 30%.

Model results are summarized in **Table 3** to **Table 12** below. Figures in these tables are rounded to one decimal place for greater legibility (hence 0.0 means in effect a number less than 0.1).

Table 3. Model description and diagnostics for the final selected models of **harbour porpoise** individuals in SCANS-IV (2022). Covariates in the models are described in **Table 1**.

Distribution	Model covariates	Estimated degrees of freedom	Model degrees of freedom	% Deviance explained
Negative Binomial (0.397)	Soap(X,Y)	55.1	71.8	42.6
	NPPV	1.7		
	mSST	5.0		
	Depth	7.3		
	Aspect	0.0		
	Distance to coast	2.7		

Table 4. Model description and diagnostics for the final selected models of **bottlenose dolphin** groups in SCANS-IV (2022). Covariates in the models are described in **Table 1**.

Distribution	Model covariates	Estimated degrees of freedom	Model degrees of freedom	% Deviance explained
Negative Binomial (0.157)	Soap(X,Y)	36.7	46.1	41.2
	Depth	1.4		
	Seabed Slope	2.0		
	Aspect	1.7		
	Distance to 200m	4.3		

Table 5. Model description and diagnostics for the final selected models of **white-beaked dolphin** groups in SCANS-IV (2022). Covariates in the models are described in **Table 1**.

Distribution	Model covariates	Estimated degrees of freedom	Model degrees of freedom	% Deviance explained
Negative Binomial (0.143)	Soap(X,Y)	14.9	22.1	51.2
	mSST	1.7		
	EKE	0.2		
	gradSST	0.9		
	Aspect	0.0		
	Distance to coast	4.4		

Table 6. Model description and diagnostics for the final selected models of **common dolphin** groups in SCANS-IV (2022). Covariates in the models are described in **Table 1**.

Distribution	Model covariates	Estimated degrees of freedom	Model degrees of freedom	% Deviance explained
Negative Binomial (0.276)	Soap(X,Y)	43.7	53.8	45.9
	NPPV	0.7		
	mSST	5.0		
	EKE	0.1		
	Seabed Slope	0.7		
	Distance to 200m	3.6		

Table 7. Model description and diagnostics for the final selected models of **striped dolphin** groups in SCANS-IV (2022). Covariates in the models are described in **Table 1**.

Distribution	Model covariates	Estimated degrees of freedom	Model degrees of freedom	% Deviance explained
Negative Binomial (0.153)	Soap(X,Y)	6.9	11.2	28.1
	gradSST	0.4		
	Depth	2.8		
	Aspect	0.0		
	Distance to 200m	1.1		

Table 8. Model description and diagnostics for the final selected models of **unidentified common or striped dolphin** groups in SCANS-IV (2022). Covariates in the models are described in **Table 1**.

Distribution	Model covariates	Estimated degrees of freedom	Model degrees of freedom	% Deviance explained
Negative Binomial (0.236)	Soap(X,Y)	22.0	32.9	39.9
	NPPV	2.6		
	mSST	3.8		
	Aspect	1.7		
	Distance to 50m	1.4		
	Distance to 200m	1.4		

Table 9. Model description and diagnostics for the final selected models of **common and striped dolphin, including unidentified common or striped dolphin** groups in SCANS-IV (2022). Covariates in the models are described in **Table 1**.

Distribution	Model covariates	Estimated degrees of freedom	Model degrees of freedom	% Deviance explained
Negative Binomial (0.333)	Soap(X,Y)	44.7	52.6	43.4
	mSST	2.9		
	Seabed Slope	0.9		
	Aspect	1.3		
	Distance to 50m	0.0		
	Distance to 200m	2.8		

Table 10. Model description and diagnostics for the final selected models of **beaked whales** (all species combined) individuals in SCANS-IV (2022). Covariates in the models are described in **Table 1**.

Distribution	Model covariates	Estimated degrees of freedom	Model degrees of freedom	% Deviance explained
Negative Binomial (0.014)	Soap(X,Y)	5.0	11.7	46.5
	gradSST	3.6		
	Depth	2.0		
	Aspect	1.1		

Table 11. Model description and diagnostics for the final selected models of **fin whale** individuals in SCANS-IV (2022). Covariates in the models are described in **Table 1**.

Distribution	Model covariates	Estimated degrees of freedom	Model degrees of freedom	% Deviance explained
Negative Binomial (0.265)	Soap(X,Y)	32.0	46.0	57.5
	NPPV	0.9		
	mSST	2.9		
	Seabed Slope	1.1		
	Aspect	3.1		
	Distance to coast	6.0		

Table 12. Model description and diagnostics for the final selected models of **minke whale** individuals in SCANS-IV (2022). Covariates in the models are described in **Table 1**.

Distribution	Model covariates	Estimated degrees of freedom	Model degrees of freedom	% Deviance explained
Negative Binomial (0.064)	Soap(X,Y)	12.9	19.2	28.8
	NPPV	4.3		
	mSST	1.4		
	Distance to coast	0.0		
	Distance to 200m	0.6		

The maps showing surfaces of predicted density and associated estimated coefficient of variation (CV) are shown for each species in **Figure 2-Figure 11**. The patterns of predicted density are influenced by the covariates retained in the models (see **Table 3-Table 12**), the fitted smooth functions (see Appendix **Figure A. 11 - Figure A. 20**), and spatial variation in the values of the covariates in the prediction grid. Overall, model-predicted density surface plots captured observed distribution patterns for all species (see Gilles et al., 2023), revealing high-density areas within the study area and demonstrating the models' capacity to predict distribution patterns in complex environmental contexts.

The maps of CVs provide a measure of the confidence in predicted density across the survey area. Lower CVs are generally associated with areas of higher density when predictions are interpolations *sensu* Pigeault et al. (2024b). On the other hand, high CV values may betray either extrapolations, a high between-day variance in predictions, or be associated with areas of very low density.

Confidence in the predictions in areas of low density is generally much poorer. This is expected; in areas with no sightings or few sightings the model has little information to learn on a lower bound to the parameters, which generates a large uncertainty relative to the mean (the uncertainty remains small in absolute value however). The magnitude of the CVs is influenced by the number of sightings as well as by how well the models fit the data. Thus, the CVs for predicted harbour porpoise density (**Figure 2**) are relatively low across most of the survey area because of the much larger number of sightings, despite the fact that the models accounted for less deviance for porpoises than for some other species (**Table 3**). The low density of harbour porpoises in offshore areas of the Bay of Biscay and Iberian Coasts sub-region are associated with high CVs: no sightings were made in these offshore areas, illustrating the point above on low confidence associated with some low-density areas.

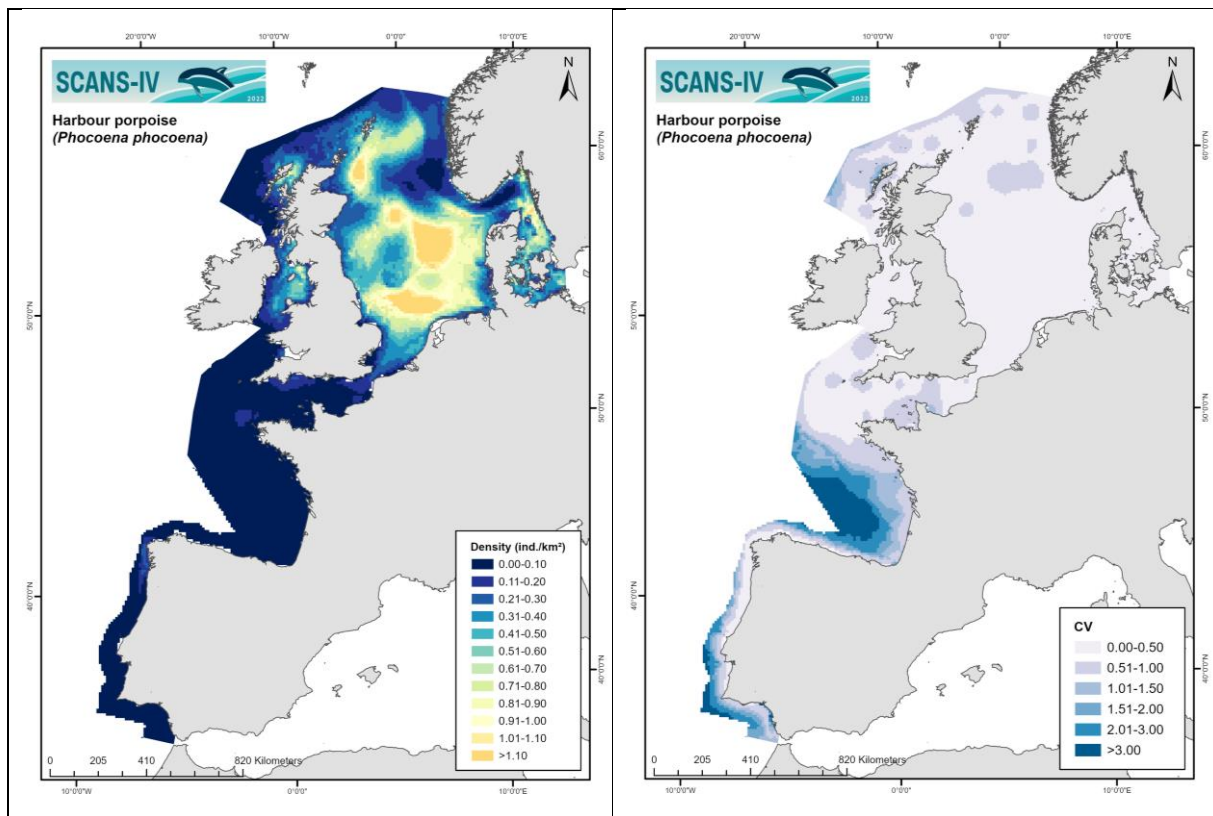


Figure 2. Predicted surfaces of estimated density [left] and associated coefficient of variation (CV) [right] for **harbour porpoise** in SCANS-IV (2022).

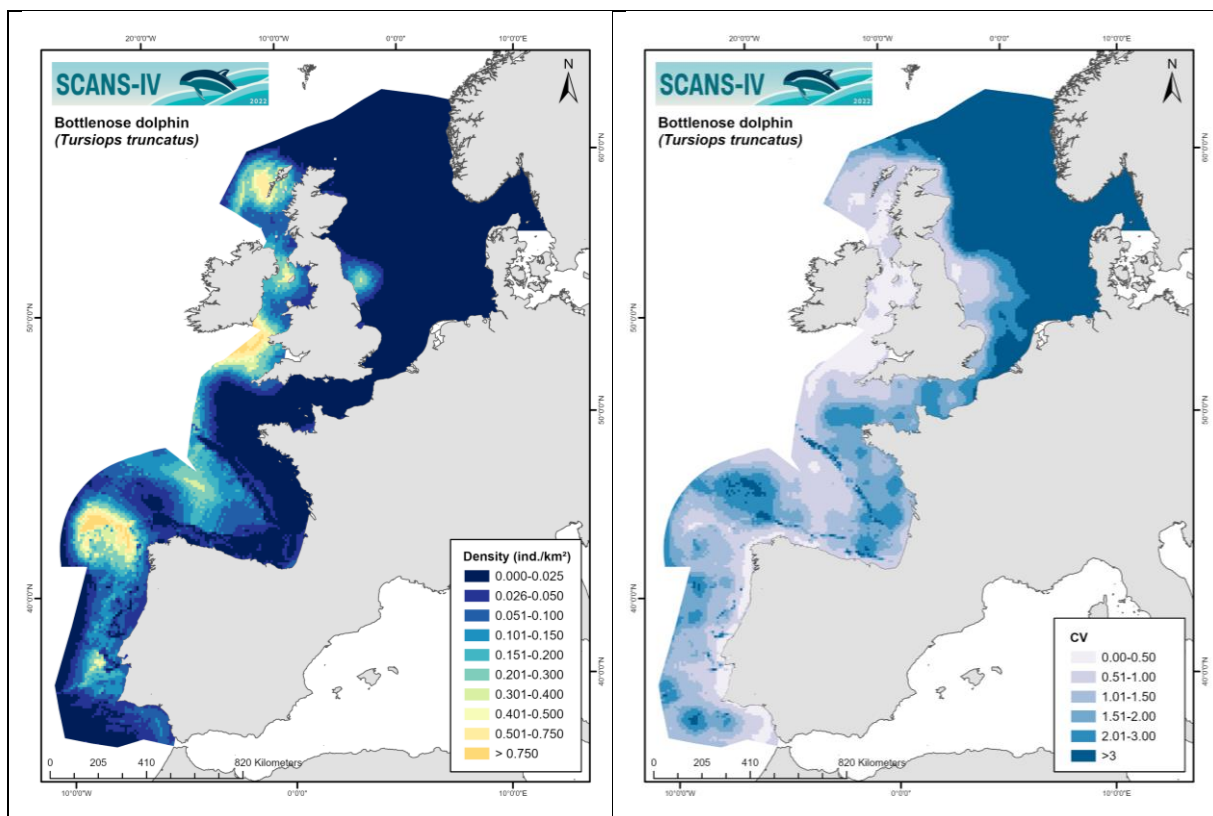


Figure 3. Predicted surfaces of estimated density [left] and associated coefficient of variation (CV) [right] for **bottlenose dolphin** in SCANS-IV (2022).

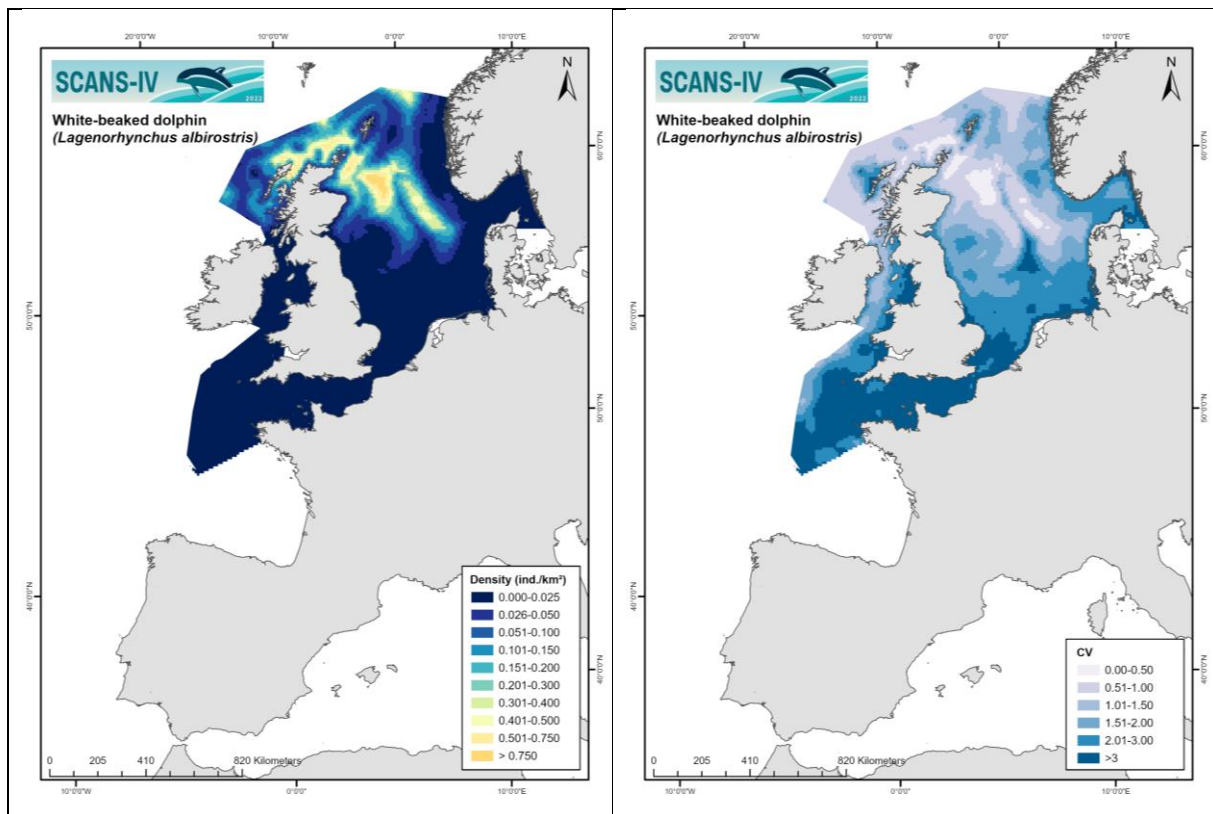


Figure 4. Predicted surfaces of estimated density [left] and associated coefficient of variation (CV) [right] for **white-beaked dolphin** in SCANS-IV (2022).

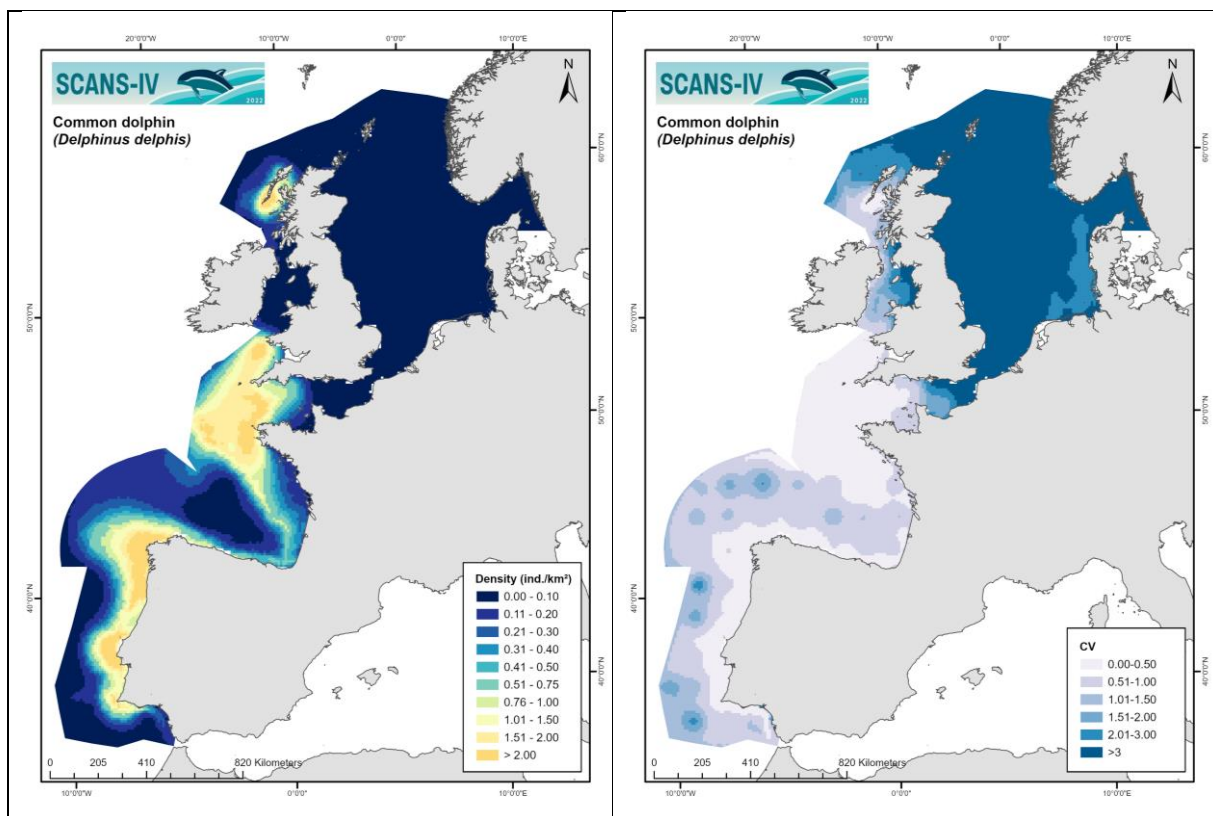


Figure 5. Predicted surfaces of estimated density [left] and associated coefficient of variation (CV) [right] for **common dolphin** in SCANS-IV (2022).

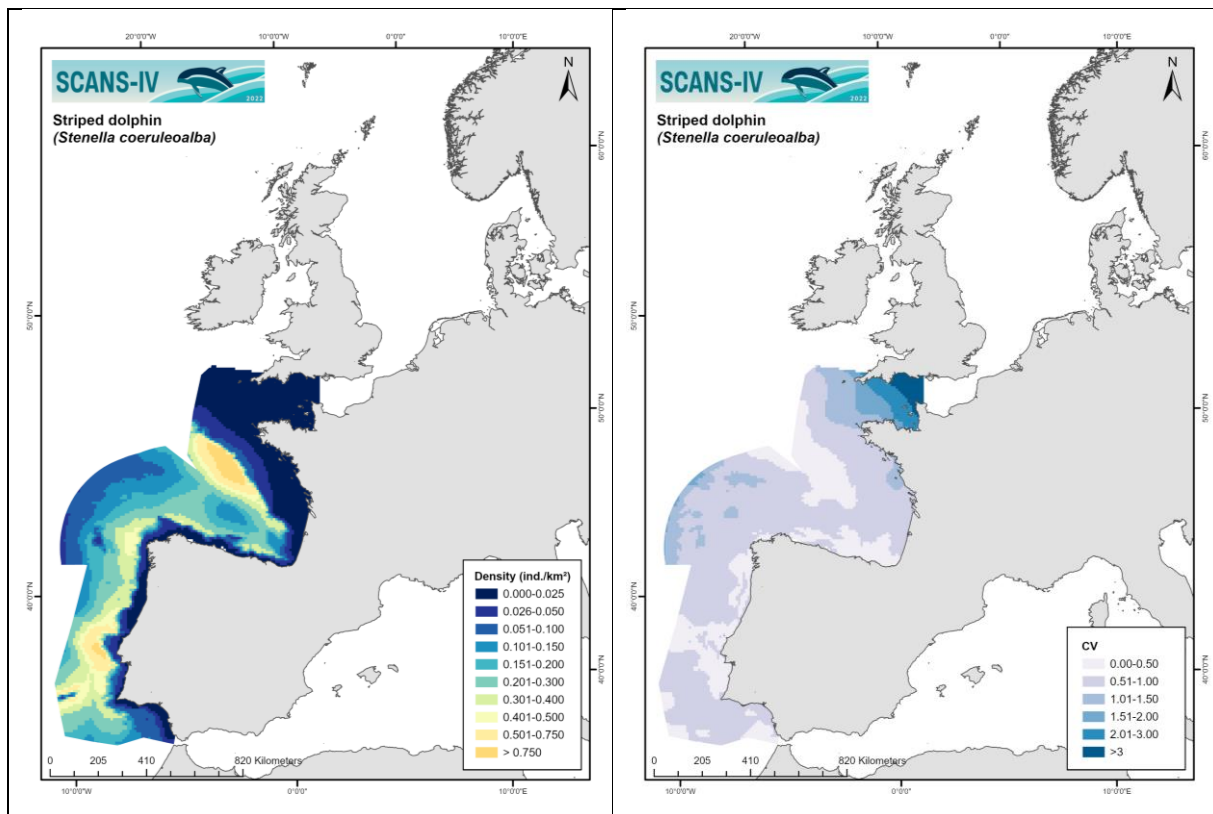


Figure 6. Predicted surfaces of estimated density [left] and associated coefficient of variation (CV) [right] for **striped dolphin** in SCANS-IV (2022).

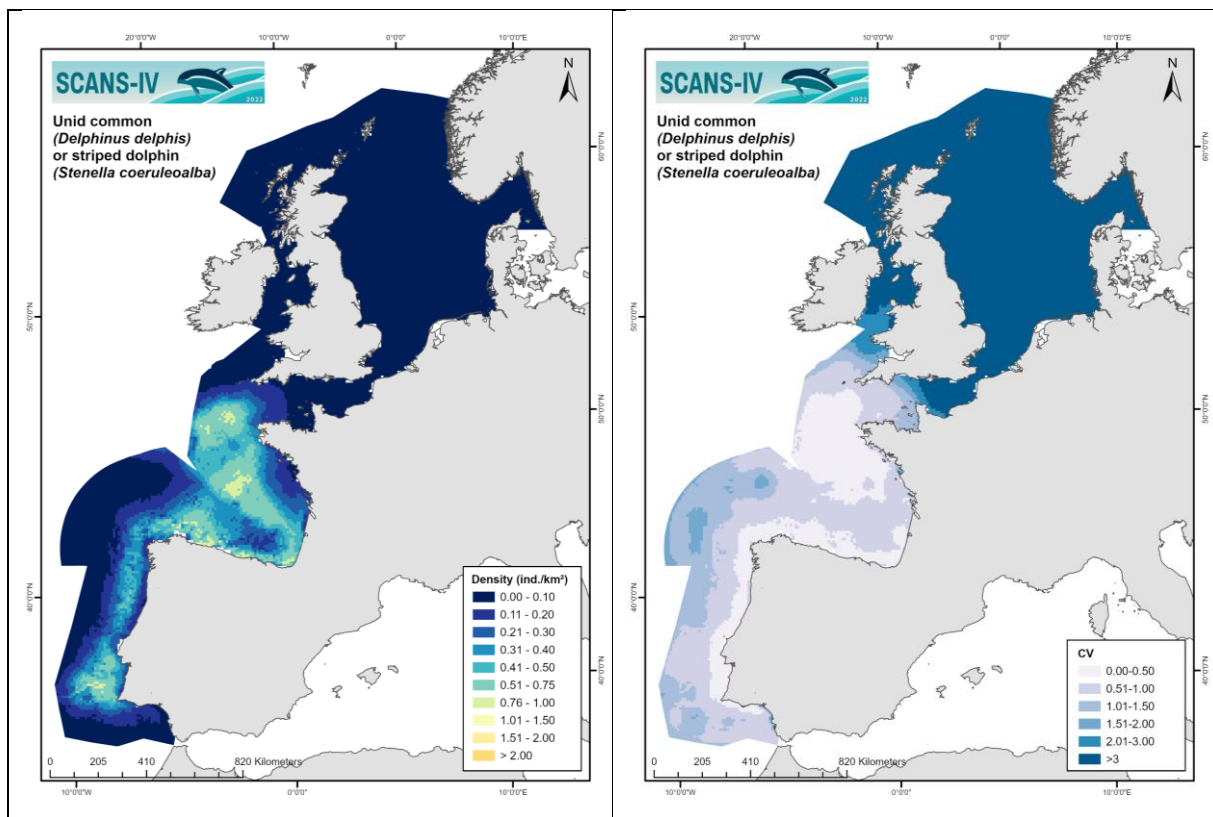


Figure 7. Predicted surfaces of estimated density [left] and associated coefficient of variation (CV) [right] for **unidentified common or striped dolphin** in SCANS-IV (2022).

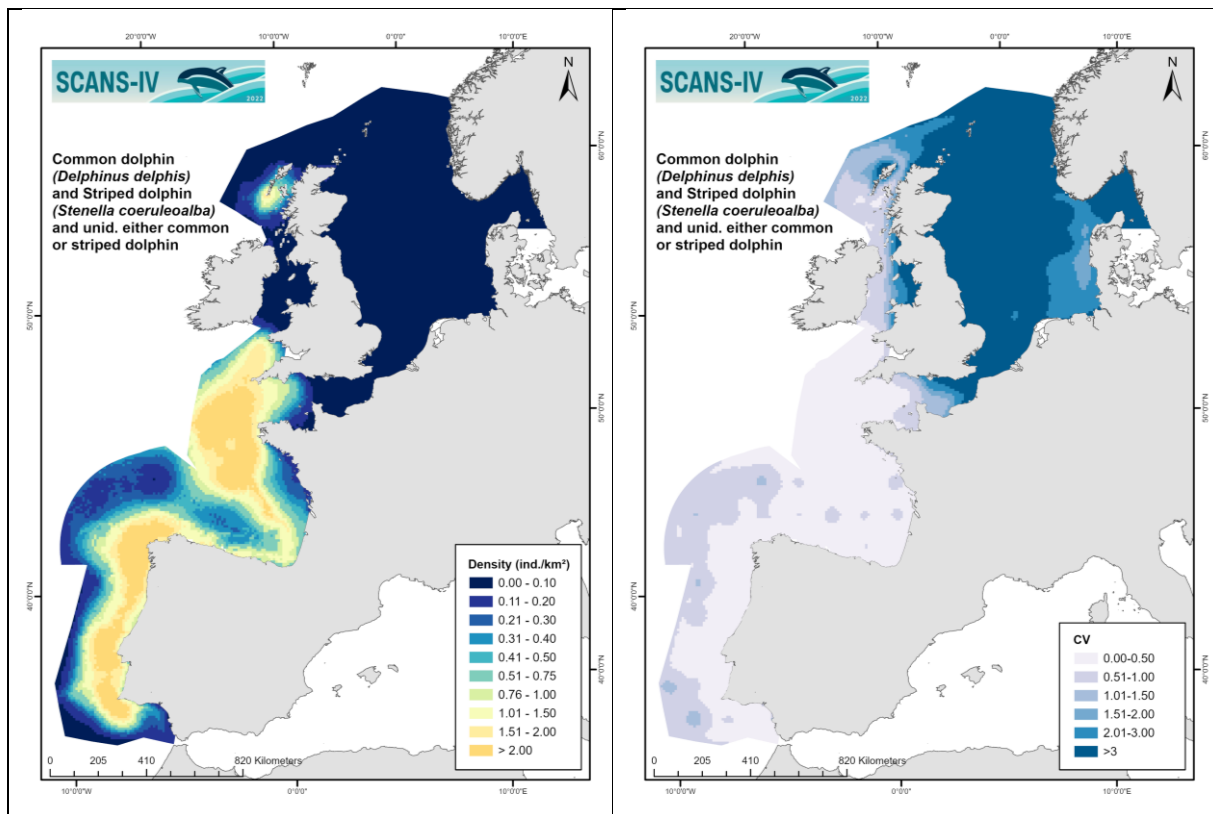


Figure 8. Predicted surfaces of estimated density [left] and associated coefficient of variation (CV) [right] for **common dolphin, striped dolphin and unidentified common or striped dolphin** in SCANS-IV (2022).

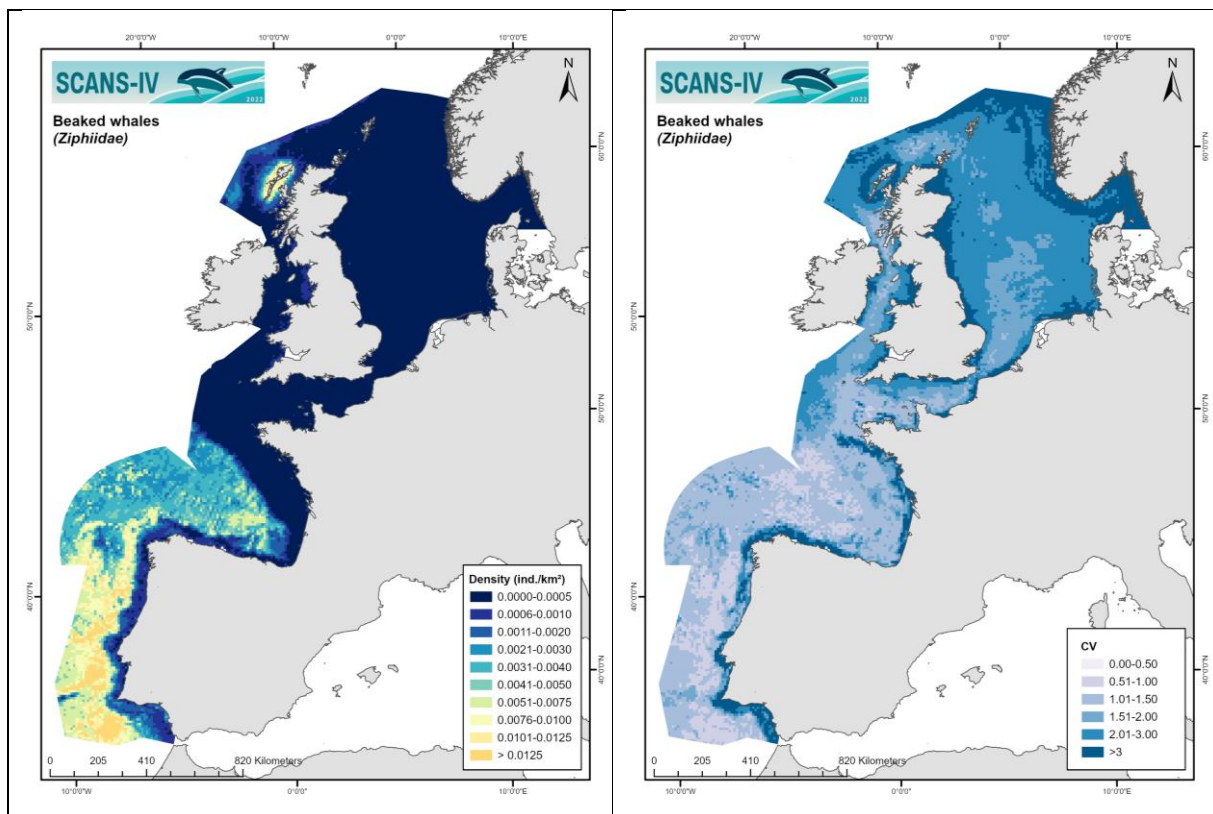


Figure 9. Predicted surfaces of estimated density [left] and associated coefficient of variation (CV) [right] for **beaked whales (all species combined)** in SCANS-IV (2022).

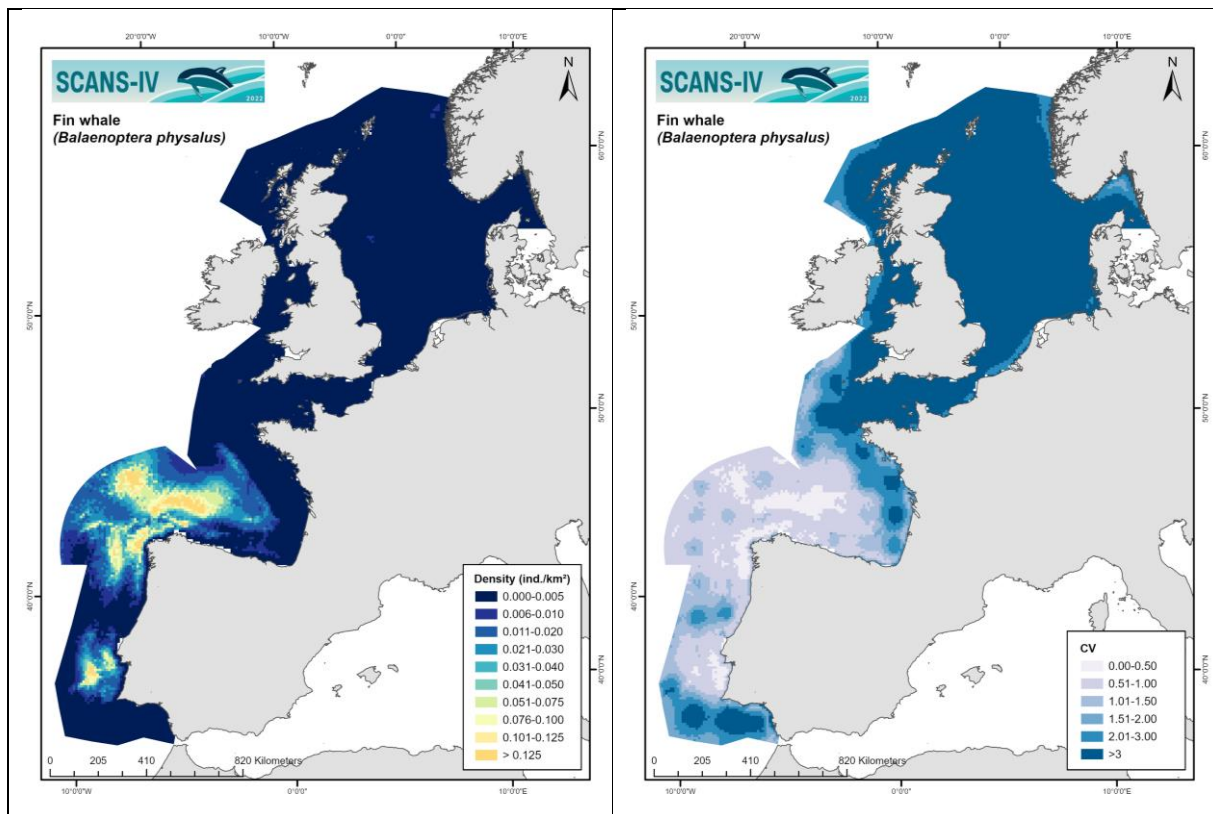


Figure 10. Predicted surfaces of estimated density [left] and associated coefficient of variation (CV) [right] for **fin whale** in SCANS-IV (2022).

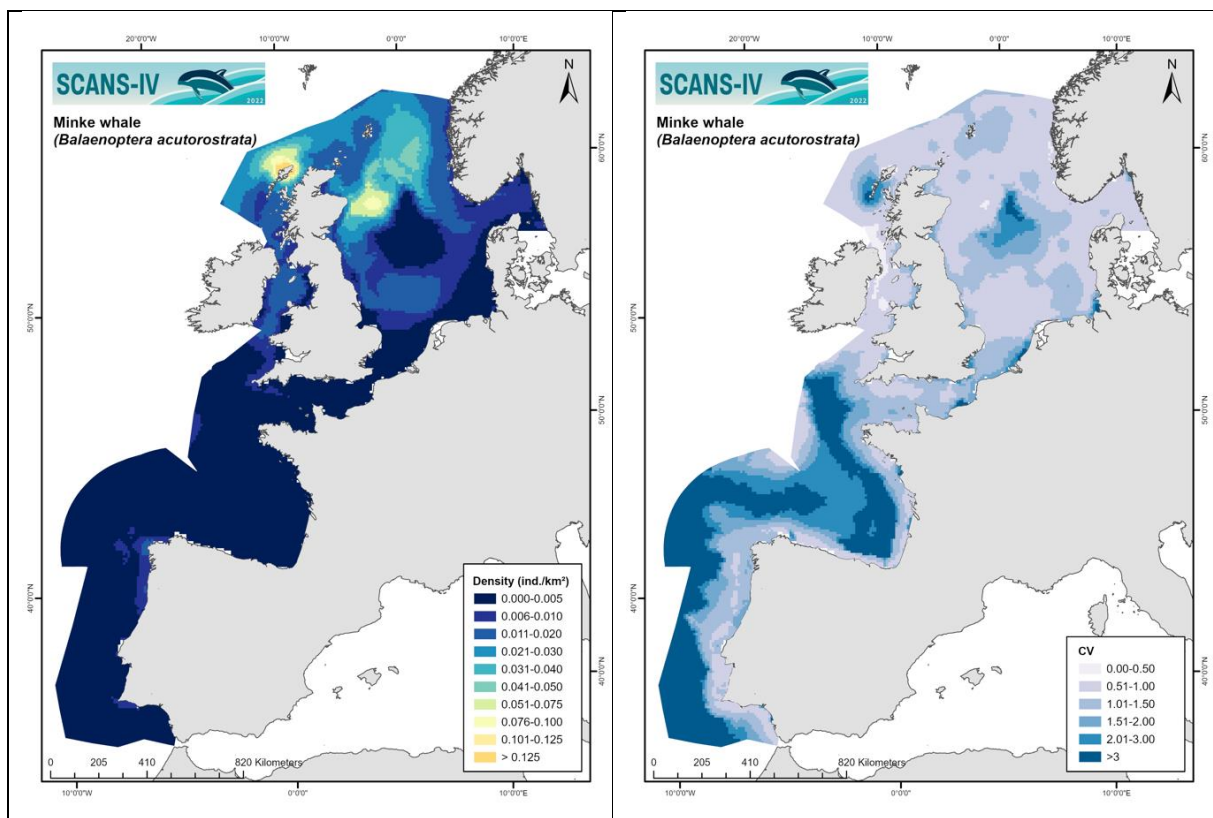


Figure 11. Predicted surfaces of estimated density [left] and associated coefficient of variation (CV) [right] for **minke whale** in SCANS-IV (2022).

DISCUSSION

The new models fitted using the recently collected SCANS-IV data will support risk-based assessments as well as multiple other evidence needs for cetacean conservation in European Atlantic waters. Environmental conditions are subject to inter-annual and decadal-scale variation, and to directional change as a result of ocean warming. Long-term data are required to keep track of how spatial distribution, and the environmental features that drive that distribution, may be changing.

The modelling workflow adopted for SCANS-IV built on the one used for SCANS-III (Lacey et al., 2022) but added some new developments while striving to maintain comparability of results. The main changes stemmed from using a soap-smooth instead of a bivariate smooth on locational covariates to account for any residual spatial autocorrelation not captured by environmental covariates. The soap-smooth prevents leakage across land barriers and, in theory, enables more accurate predictions. However, little guidance exists in the published literature with respect to knot placements and numbers of knots. These issues were dealt with by several trials and use of another development made during the modelling analyses of SCANS-IV: rootograms. These plots were used as a graphical goodness-of-fit diagnostics both at the level of the whole survey (see Appendix **Figure A. 22**), and at block-level (not shown) for an in-depth investigation of model fit, especially model overfitting. A third development was the adoption of a multi-model framework using stacking. Stacking is an ensemble method that takes the outputs of many models and weighs these outputs to combine them in a single prediction that accounts for model-uncertainty. The full potential of stacking was not harnessed in the present modelling workflow to maintain interpretability of results, especially relationships with environmental covariates which are easier to interpret from a single model, and comparability with previous results. Although stacking weights for the five best-fitting models (with respect to AIC) were computed, the model with the highest weight was selected to predict distribution and density. A fourth development was the use of a pseudo-posterior approach for uncertainty quantification: this approach allowed for seamless (i) trickling down of estimation uncertainty and (ii) accounting for correlation in daily predictions before aggregation over the surveyed period. Use of the pseudo-posterior approach nevertheless required some care with respect to “transformation bias” and “simulation bias” (Rainey, 2017; 2024). These four developments allowed for a more robust modelling with respect to taking into land barriers in the large surveyed area (**Figure 1**) and uncertainty quantification.

Harbour porpoise

In summer 2022, the highest harbour porpoise densities were predicted in the central and southern North Sea, with several distinct hotspots, namely in offshore areas on the slopes of the Dogger Bank and around the Frisian Front (**Figure 2**). Higher densities were also predicted off north-eastern Scotland, south of Orkney and in the northern Kattegat.

The most noticeable difference between the modelled distributions from SCANS-II (in 2005), SCANS-III (in 2016) and now SCANS-IV, is the further reduction in density observed in the Celtic Sea (southwest of Britain), the Irish Sea and the Belt Sea. It should be noted that part of the Celtic Sea was surveyed by the ObSERVE-2 survey concurrently with SCANS-IV in 2022 (see **Figure 1**), but this sister survey also reported a marked decline in the abundances estimated for harbour porpoise compared to surveys conducted in 2015-2017 (Giralt Paradell et al., 2024). Given the high reported bycatch of porpoises in these areas (ICES, 2024; Kindt-Larsen et al., 2023), with likely unsustainable levels in these harbour porpoise assessment units (Taylor et al., 2022; Owen et al., 2024), effective conservation measures are needed.

In comparison to SCANS-III in summer 2016, much lower densities were predicted in the Belt Sea porpoise population region in the waters shared between Denmark, Germany and Sweden, as well as off eastern England and southeastern Scotland in summer 2022, in a region from Flamborough Head in the south to Buchan Deep in the north and offshore to about 250 km.

The previously reported southward expansion of harbour porpoise distribution in the North Sea towards the Channel extended slightly further but is comparable to the modelled distribution from SCANS-III and did not extend further south.

In the south, density was predicted to be generally low, except for coastal waters west of Galicia/Cape Finisterre and in northern Portugal. It is important to point out that the IC blocks off Portugal and off Spain were surveyed in different months (Spanish blocks were surveyed later than the rest of the survey period), and due to the particular situation of the Iberian harbour porpoise, future simultaneous and seasonal surveys are needed. The relative high density area predicted in Galician waters is coincident with the proposed critical area included in the preliminary Recovery Plan that the government of Spain has been developing since November 2020. At this time, the status of harbour porpoise in the Spanish Catalogue of Threatened Species was updated from “vulnerable” to “in danger of extinction” (Order TED/1126/2020).

Bottlenose dolphin

The modelled distribution of bottlenose dolphins in summer 2022 (**Figure 3**) highlights a number of higher density areas in waters along the western edge of the survey area. In the south, density is relatively high in offshore waters to the west of Portugal and particularly high west of Galicia. Densities are moderate in offshore waters of the Bay of Biscay. Another high density area is apparent to the west of southwest Wales and the Bristol Channel. Moderate density is predicted in the rest of the Irish Sea and off western Scotland, where density is particularly high to the west of the southern Hebrides Islands. The modelled distribution of bottlenose dolphins from the ObSERVE-2 survey in summer 2022 shows the highest densities along the southeastern edge of the survey area (Giralt Paradell et al., 2024), which matches well with the high predicted density to the west of southwest Wales and the Bristol Channel from SCANS-IV (**Figure 3**). However, overall abundance estimates for bottlenose dolphins were lower during ObSERVE-2 than in phase I (Giralt Paradell et al., 2024).

Bottlenose dolphins in European Atlantic waters comprise two ecotypes: those in small resident coastal populations, and offshore animals (Geelhoed et al., 2022). The resident coastal populations in these waters are included in the SCANS-IV survey area but such large-scale line transect surveys are not appropriate either to estimate their abundance or to map their distribution at a fine spatial scale. Instead, photo-identification surveys provide the data for these populations (e.g. Arso Civil et al., 2019; Cheney et al., 2024). It is clear from **Figure 3** that the large majority of the animals represented by the areas of higher predicted density are offshore animals. However, **Figure A. 2** shows that there were sightings close to the coast throughout much of the survey area. These patterns reflect that the resident coastal populations are very much smaller than the offshore population.

One such coastal population resides along the east coast of Britain in the North Sea. Since studies of this population began in 1989, its range has expanded south from the Moray Firth in northeast Scotland to the Tay estuary and adjacent waters in central Scotland and, most recently, further south to waters off northeast England (Arso Civil et al., 2019, in prep; Cheney et al., 2024; Ellis et al., in press). **Figure 3** shows a relatively high density area off northeastern England, highlighting the extent to which this population has spread southwards. The handful of sightings in the wider North Sea (**Figure A. 2**)

also reflect that the ranges of some individuals from this population are not restricted to coastal waters of eastern Britain (Cheney et al., 2024).

Distribution modelling of the SCANS-III data from 2016 (Lacey et al., 2022) resulted in a strongly smoothed predicted distribution with higher density areas limited to west of southwestern England and Brittany, and off Galicia, both areas broadly reflected in 2022. Modelling SCANS-II and CODA data from 2005/07 resulted in the highest densities predicted along the coast of Portugal and Galicia, with relatively high density also predicted west of Galicia and on the outer Celtic Shelf (similar to 2016 and 2022) but also along the western edge of the survey area west of Ireland and Scotland (Lacey et al., 2022).

The variation in predicted density from survey to survey likely reflects that, although there are two bottlenose dolphin ecotypes in the area, the majority of animals are from offshore waters and part of a poorly described population(s), the distribution of which may vary annually in response to environmental factors, especially the availability of prey.

White-beaked dolphin

The modelled distribution of white-beaked dolphins in summer 2022 (**Figure 4**) shows areas of high density in the northern North Sea between Scotland and Norway and also to the north and northwest of Scotland around the Orkney Islands and northern Hebrides. This pattern is very similar to the areas of high density predicted from the SCANS-II and SCANS-III surveys in 2005 and 2016 (Lacey et al., 2022) and is also similar to the pattern of sightings from the SCANS survey in 1994 (Hammond et al., 2002). There were too few sightings from the ObSERVE-2 survey in summer 2022 to support distribution modelling (Giralt-Paradell et al., 2024). There were more sightings from ObSERVE-1 in summer 2015 and 2016, a few off the west coast of Ireland but mostly further offshore along the shelf edge or on the Porcupine Bank (Rogan et al., 2018).

Overall, the series of SCANS surveys show a remarkably consistent summer distribution of white-beaked dolphins in European Atlantic waters over almost three decades. Although other authors have suggested a northward shift in distribution based on analysis of strandings and incidental sightings data (IJseldijk et al., 2018; Lambert et al., 2014; Williamson et al., 2021), we have found no evidence of this to date.

Surveys in the wider North Atlantic show that the sightings of white-beaked dolphins around northern Scotland and the northern North Sea from SCANS, and those off Ireland from ObSERVE, are disjunct from those found off east Greenland and around Iceland and off northern Norway (NAMMCO, 2023). DNA genotyping of tissue from animals stranded around Britain, Ireland and the North Sea shows that these animals are strongly differentiated from the rest of the North Atlantic. Furthermore, admixture analysis indicates that there are likely two populations in this region, one in the North Sea and another off western Scotland and Ireland, with an area of overlap off eastern Scotland (Gose et al., 2024).

Thus, unlike other dolphin species in the European Atlantic, the white-beaked dolphin likely has two overlapping populations that have been shown to be distinct from conspecifics in the rest of the North Atlantic. This provides a strong motivation for continuing to monitor the distribution and abundance of this species in this region.

Common and striped dolphin

The modelled distribution of **common dolphins** in summer 2022 showed high predicted density around the Iberian coasts (Spain and Portugal shelf waters, excluding the Bay of Cadiz) and along the shelf edge in the northern Bay of Biscay and Celtic Sea, up to St Georges' Channel, between Ireland and Wales (**Figure 5**). High density areas were also predicted to the west of Galicia and around the Hebrides, west of Scotland. The ObSERVE-2 surveys in Irish waters in 2021-22 recorded many common dolphins mostly in summer 2021; the species was seen in all ObSERVE strata but infrequently recorded in the Irish Sea. During ObSERVE-2, common dolphin sightings were numerous in continental shelf waters, where animals were sighted both in coastal and offshore areas (Giralt Paradell et al., 2024). The modelled distribution of common dolphins in summer 2022 from ObSERVE-2 in waters within the Irish EEZ predicted high densities primarily in continental shelf waters along the west coast of Ireland, across the Celtic Basin and west of Scotland, with a low density area in the western Irish Sea (Giralt Paradell et al., 2024). This pattern is congruent with the modelled distribution from SCANS-IV (**Figure 6**). Model results supported the hypothesis of a northward extension in the distribution of common dolphins (Gilles et al., 2023). Except for the high-density areas west of Scotland and at the mouth of St George's Channel (**Figure 5**), the modelled distribution of common dolphins in 2022 was similar to the distribution modelled from 2016 (Lacey et al., 2022).

The modelled distribution of **striped dolphins** in summer 2022 predicted high densities in waters off the shelf throughout the Bay of Biscay, west of Galicia and in offshore waters in front of Cape da Roca (centre-west Portugal) (**Figure 6**). This distribution is thus largely disjunct from that of common dolphins, except that there is overlap in predicted high density of the two species along the shelf edge of the Bay of Biscay and west of Galicia. There were eleven sightings of striped dolphins during the ObSERVE-2 survey, with the majority during summer 2021 (Giralt Paradell et al., 2024), supporting the modelled distribution mostly in waters south to the Celtic Seas (**Figure 6**).

The modelled distribution of either common or striped dolphins (individuals that could not be identified to species level) reflected the overlap between the two species along the shelf break of the Bay of Biscay and Iberian Coasts, and in offshore waters in the southern Bay of Biscay (**Figure 7**). The modelled distribution should be interpreted as areas where species identification from an aerial platform is the most challenging as both common or striped dolphins are co-occurring in these areas. Results from species identification using digital photos are reported in Gilles et al. (2023). In block CS-B, which covers the continental shelf in the southern Celtic Seas down to the northern Bay of Biscay, the analysis of digital photos of sightings of the category 'either common or striped dolphin' showed that 100% of these sightings corresponded to common dolphins. In block BB-B, which covers the continental shelf of the Bay of Biscay, 93% of such unidentified sightings were labelled as common dolphins from digital photos, with the remainder (7%) labelled as striped dolphins. In block BB-A which covers the shelf break and offshore waters in the inner Bay of Biscay, the proportion of unidentified sightings labelled from digital photos as common and striped dolphins were respectively 19% and 81%.

The modelled distribution of common and striped dolphins combined, including those unidentified to species, in 2022 largely reflects those of common dolphins, driven by the order of magnitude greater number of sightings of that species (**Figure 8**). Except for a high-density area west of Scotland, the modelled distribution in 2022 was similar to that in 2016 (Lacey et al., 2022).

Beaked whales

The modelled distribution of beaked whales (all species) in 2022 shows the highest predicted density along the shelf edge and in the deep waters of the Bay of Biscay, off the north coast of Spain and Portugal (**Figure 9**). A relatively large number of beaked whale sightings had already been registered

during the LIFE MarPro offshore ship survey performed in 2011 in Portuguese waters (Vingada & Eira 2018). These results for beaked whales emphasise the importance of further surveying offshore waters in the Iberian Coast. The observed distribution is similar to previous SCANS surveys; however, no conclusion can be drawn about the deep waters west of Scotland, where high densities were predicted in SCANS-III and particular in CODA 2007, since SCANS-IV did not cover this area.

Lacey & Hammond (2023) modelled available data for beaked whales from systematic surveys conducted in 2015 and 2016, including SCANS-III, ObSERVE and the Faroes surveys of NASS-2015. Models were fitted for individual species where possible, which broadly showed patterns of highest densities of Cuvier's beaked whales in southern areas, northern bottlenose whales in northern areas west of Scotland, around the Faroe Islands and off Iceland, and Sowerby's beaked whales off the shelf edge. But sample sizes were very small and predicted densities therefore had high uncertainty, so these results should be considered cautiously. The model of all beaked whale species combined predicted the highest densities northwest of the SCANS survey area in deep waters off western Scotland and around the Faroe Islands and Iceland, and in the Bay of Biscay. These results reflected those obtained from models of data from SCANS-II and SCANS-III (Lacey et al., 2022). These previous model predictions lacked data from offshore waters of Portugal, which were not initially surveyed as part of the SCANS series but were included in SCANS-IV. It is evident that these offshore Portuguese waters are of particular importance to beaked whales.

Giralt Paradell et al. (2024) combined predictions for beaked whales in summer from both ObSERVE phases. High densities were predicted in deep waters off the continental shelf, along the continental slope throughout the survey area, which aligns closely with the modelled distribution from SCANS-IV.

Fin whale

The modelled distribution of fin whales in summer 2022 was similar to those from 2016 and 2005/2007, showing high predicted densities in offshore waters of the Bay of Biscay (**Figure 10**). However, the modelled distribution in 2022 suggested a more coastal distribution of fin whales on the north-western part of the Iberian Coasts (esp. around Galicia). Yet, it is imperative to acknowledge that the aerial survey in the coastal waters of Spain was carried out later than the main survey period, i.e. in September and October 2022 (Gilles et al., 2023). Further (seasonal) surveys need to be conducted to improve understanding of the coastal presence of fin whales in this region.

Also, SCANS-IV being the first survey in the SCANS series to cover Portuguese offshore waters, revealed a high-density hotspot of fin whales in the southwestern offshore waters of Portugal. During the ObSERVE-2 survey, fin whales were mainly sighted west of Ireland, beyond the continental slope in waters >500m, particularly in the Porcupine Basin: of the 15 sightings, only two were made in summer 2022 (Giralt Paradell et al., 2024).

Minke whale

The modelled distribution of minke whales in summer 2022 showed the highest densities in the north-western North Sea, and in shelf waters west of Scotland (**Figure 11**). Compared to the modelled distribution in summer 2016, low densities were predicted in the central North Sea and in the Moray Firth. Low densities were also predicted in the Celtic and Irish Seas compared to modelled distributions from 2016 and 2005/2007 (Lacey et al., 2022). In contrast to both 2005/07 and 2016, the modelled

distribution in summer 2022 was not suggestive of areas of high predicted density towards the western edge of the survey area west of Scotland (**Figure 11**).

There were several sightings of minke whales in Galician and Portuguese coastal waters (**Figure A. 10**), leading to an area of slightly elevated predicted density off the west coast of Galician and northern Portugal (**Figure 11**). This is in contrast to previous SCANS surveys in 2005 and 2016, in which there had been no sightings in these waters (Hammond et al., 2013; Hammond et al., 2021). The two sightings in Portuguese waters were made in July but the cluster of sightings off Galicia were made in September/October. Although these new results could suggest a disjunct Iberian segment of the North Atlantic population, the majority of the sightings being made in autumn could reflect animals on migration.

During ObSERVE-2, minke whales were sighted all around Ireland except on the north coast, with most of the sightings in continental shelf waters <200m and highest densities in coastal waters off the south coast of Ireland (Giralt Paradell et al., 2024).

ACKNOWLEDGMENTS

This study has been conducted using E.U. Copernicus Marine Service Information; <https://doi.org/10.48670/moi-00019>, <https://doi.org/10.48670/moi-00021>. We thank Leonie Mahlke for support with ArcGIS, Mathieu Genu for computational support and scripting regarding the implementation of the soap smooth and Paula Gutierrez Muñoz for reviewing species distribution maps.

The SCANS-IV project was supported by funding from: Miljøministeriet (Denmark); Office Français de la Biodiversité (France); Bundesamt für Naturschutz and the Bundesministerium für Umwelt, Naturschutz, nukleare Sicherheit und Verbraucherschutz (Germany); Ministerie van Landbouw, Natuur en Voedselkwaliteit (Netherlands); Fundo Ambiental and Instituto da Conservação da Natureza e das Florestas (Portugal); El Ministerio para la Transición Ecológica y Reto Demográfico and the Ministerio De Agricultura, Pesca Y Alimentación (Spain); Havs- och vattenmyndigheten (Sweden); Joint Nature Conservation Committee, Department for Environment, Food and Rural Affairs and Department for Business, Energy & Industrial Strategy (UK).

REFERENCES

- Arso Civil, M, Quick, N J, Cheney, B, Pirotta, E, Thompson, PM, & Hammond, PS (2019). Changing distribution of the east coast of Scotland bottlenose dolphin population and the challenges of area-based management. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 29(S1), 178-196.
- Arso Civil, M, Ellis, GK, Ivanchikova, J, Coxon, JC, Montabaranom, J, Cheney, BJ, Thompson, PM & Hammond, PS (in prep). Monitoring the Scottish east coast bottlenose dolphin population during a southward range expansion. Final report to Forth and Tay windfarm developers (Seagreen, Inch Cape, NnG).
- Becker, EA, Forney, KA, Fiedler, PC, Barlow, J, Chivers, SJ, Edwards, CA, Moore, AM, Redfern, JV (2016). Moving Towards Dynamic Ocean Management: How Well Do Modeled Ocean Products Predict Species Distributions? *Remote Sensing*, 8(2), 149. <https://doi.org/10.3390/rs8020149>
- Becker, EA, Forney, KA, Redfern, JV, Barlow, J, Jacox, MG, Roberts, JJ, Palacios, DM (2019). Predicting cetacean abundance and distribution in a changing climate. *Diversity and Distributions*, 25, 626–643. doi:10.1111/ddi.12867
- Becker, EA, Forney, KA, Miller, DL, Fiedler, PC, Barlow, J, Moore, JE (2020). Habitat-based density estimates for cetaceans in the California Current Ecosystem based on 1991-2018 survey data, U.S. Department of Commerce, NOAA Technical Memorandum NMFS-SWFSC-638.
- Cheney, BJ, Arso Civil, M, Hammond, PS, Thompson, PM (2024). Site Condition Monitoring of bottlenose dolphins within the Moray Firth Special Area of Conservation: 2017-2022. NatureScot Research Report 1360.
- CODA (2009). Cetacean Offshore Distribution and Abundance in the European Atlantic (CODA). Final Report. University of St Andrews, UK.
- Ellis, GK, Arso Civil, M, Cheney, BJ, Tapp, S & Sparling, CE (in press). Bottlenose dolphin ecology in coastal northeast England: Baseline abundance and connectivity of the East Coast Scotland population using photo identification in an area of range expansion. Natural England Report
- EMODnet Digital Bathymetry (DTM 2020). EMODnet Bathymetry Consortium <https://doi.org/10.12770/bb6a87dd-e579-4036-abe1-e649cea9881a>
- EMODnet Bathymetry World Coastline. EMODnet Bathymetry Consortium <https://doi.org/10.12770/36821cff-56db-4e96-8426-ddfe68240c4b>
- Geelhoed, SCV, Authier, M, Pigeault, R, Gilles, A (2022). Abundance and Distribution of Cetaceans. In: OSPAR, 2023: The 2023 Quality Status Report for the Northeast Atlantic. OSPAR Commission, London. Available at: <https://oap.ospar.org/en/ospar-assessments/quality-status-reports/qsr-2023/indicator-assessments/abundance-distribution-cetaceans/>.
- Genz, A, Bretz, F (2009). Computation of Multivariate Normal and t Probabilities series. Lecture Notes in Statistics. Springer-Verlag, Heidelberg. ISBN 978-3-642-01688-2.
- Gilles, A, Viquerat, S, Becker, EA, Forney, KA, Geelhoed, SCV, Haelters, J, Nabe-Nielsen, J, Scheidat, M, Siebert, U, Sveegaard, S, van Beest, FM, van Bemmelen, R, Aarts, G (2016). Seasonal habitat-based density models for a marine top predator, the harbor porpoise, in a dynamic environment. *Ecosphere* 7(6), e01367. doi: 10.1002/ecs2.1367
- Gilles, A, Authier, M, Ramirez-Martinez, NC, Araújo, H, Blanchard, A, Carlström, J, Eira, C, Dorémus, G, Fernández-Maldonado, C, Geelhoed, SCV, Kyhn, L, Laran, S, Nachtsheim, D, Panigada, S, Pigeault, R,

- Sequeira, M, Sveegaard, S, Taylor, NL, Owen, K, Saavedra, C, Vázquez-Bonales, JA, Unger, B, Hammond, PS (2023). Estimates of cetacean abundance in European Atlantic waters in summer 2022 from the SCANS-IV aerial and shipboard surveys. Final report published 29 September 2023. 64 pp. <https://www.tiho-hannover.de/itaw/scans-iv-survey>
- Giralt Paradell, O, Cañadas, A, Bennison, A, Todd, N, Jessopp, M, Rogan, E (2024). Aerial surveys of cetaceans and seabirds in Irish waters: Occurrence, distribution and abundance in 2021-2023. Department of the Environment, Climate & Communications and Department of Housing, Local Government & Heritage, Ireland. 260 pp. <https://assets.gov.ie/static/documents/aerial-surveys-of-cetaceans-and-seabirds-in-irish-waters-occurrence-distribution-and-a.pdf>
- Global Ocean Biogeochemistry Hindcast. E.U. Copernicus Marine Service Information (CMEMS). Marine Data Store (MDS). DOI: 10.48670/moi-00019 (Accessed on 19-01-2024)
- Global Ocean Physics Reanalysis. E.U. Copernicus Marine Service Information (CMEMS). Marine Data Store (MDS). DOI: 10.48670/moi-00021 (Accessed on 18-01-2024)
- Gose, MA, Humble, E, Brownlow, A, Wall, D, Rogan, E, Sigurðsson, GM, Kiszka, JJ, Thøstesen, CB, IJsseldijk, LL, ten Doeschate, M, Davison, NJ, Øien, N, Deaville, R, Siebert, U, Ogden, R (2024). Population genomics of the white-beaked dolphin (*Lagenorhynchus albirostris*): Implications for conservation amid climate-driven range shifts. *Heredity*, 132, 192-201. <https://doi.org/10.1038/s41437-024-00672-7>.
- Hammond, PS, Berggren, P, Benke, H, Borchers, DL, Collet, A, Heide-Jørgensen, MP, Heimlich, S, Hiby, AR, Leopold, MF, Øien, N (2002). Abundance of harbour porpoises and other cetaceans in the North Sea and adjacent waters. *Journal of Applied Ecology*, 39, 361-376.
- Hammond, PS, Macleod, K, Berggren, P, Borchers, DL, Burt, ML, Cañadas, A, Desportes, G, Donovan, GP, Gilles, A, Gillespie, D, Gordon, J, Hedley, S, Hiby, L, Kuklik, I, Leaper, R, Lehnert, K, Leopold, M, Lovell, P, Øien, N, Paxton, C, Ridoux, V, Rogan, E, Samarra, F, Scheidat, M, Sequeira, M, Siebert, U, Skov, H, Swift, R, Tasker, ML, Teilmann, J, Van Canneyt, O, Vázquez, JA. (2013). Cetacean abundance and distribution in European Atlantic shelf waters to inform conservation and management. *Biological Conservation*, 164, 107-122.
- Hammond, PS, Lacey, C, Gilles, A, Viquerat, S, Börjesson, P, Herr, H, Macleod, K, Ridoux, V, Santos, MB, Scheidat, M, Teilmann, J, Vingada, J, Øien, N (2021). Estimates of cetacean abundance in European Atlantic waters in summer 2016 from the SCANS-III aerial and shipboard surveys. SCANS-III project report 1, 39 pp. https://scans3.wp.st-andrews.ac.uk/files/2021/06/SCANS-III_design-based_estimates_final_report_revised_June_2021.pdf
- Hijmans R (2023). raster: Geographic Data Analysis and Modeling. R package version 3.6-26, <https://CRAN.R-project.org/package=raster>.
- ICES (2024). Bycatch of endangered, threatened and protected species of marine mammals, seabirds and marine turtles, and selected fish species of bycatch relevance. In: Report of the ICES Advisory Committee, 2024. ICES Advice 2024, byc.eu. <https://doi.org/10.17895/ices.advice.27999401>
- IJsseldijk, LL, Brownlow, A, Davison, NJ, Deaville, R, Haelters, J, Keijl, GO, Siebert, U, ten Doeschate, M (2018). Spatiotemporal analysis in white-beaked dolphin strandings along the North Sea coast from 1991-2017. *Lutra*, 61, 153-163.
- Kindt-Larsen L, Glemarec G, Berg CW, Königson S, Kroner A-M, Sjøgaard M, Lusseau D (2023). Knowing the fishery to know the bycatch: bias-corrected estimates of harbour porpoise bycatch in gillnet fisheries. *Proceedings of the Royal Society B*, 290, 20222570. <https://doi.org/10.1098/rspb.2022.2570>

- King, G, Tomz, M, Wittenberg, J (2000). Making the Most of Statistical Analyses: Improving Interpretation and Presentation. *American Journal of Political Science*, 44(2), 347–361. <https://doi.org/10.2307/2669316>
- Kleiber, C, Zeileis, A (2016). Visualizing Count Data Regressions Using Rootograms. *The American Statistician*, 70(3), 296–303. <https://doi.org/10.1080/00031305.2016.1173590>
- Lacey, C, Gilles, A, Börjesson, P, Herr, H, Macleod, K, Ridoux, V, Santos, MB, Scheidat, M, Teilmann, J, Vingada, J, Viquerat, S, Øien, N, Hammond, P (2022). Modelled density surfaces of cetaceans in European Atlantic waters in summer 2016 from the SCANS-III aerial and shipboard surveys. Final report (March 2022). SCANS-III project report 2. https://scans3.wp.st-andrews.ac.uk/files/2022/08/SCANS-III_density_surface_modelling_report_final_20220815.pdf
- Lacey, C, Hammond, PS (2023). Distribution and abundance of beaked whales and other deep diving cetaceans in European Atlantic waters. Report to BEIS. 59 pp. <https://www.gov.uk/government/publications/uk-offshore-energy-strategic-environmental-assessment-research-projects>
- Lambert, E, Pierce, GJ, Hall, K, Brereton, T, Dunn, TE, Wall, D, Jepson, PD, Deaville, R, MacLeod, CD (2014). Cetacean range and climate in the eastern North Atlantic: future predictions and implications for conservation. *Global Change Biology*, 20, 1782-1793.
- Lau-Medrano, W (2024). grec: Gradient-Based Recognition of Spatial Patterns in Environmental Data. R package version 1.6.0, <https://CRAN.R-project.org/package=grec>.
- Marra, G, Wood, SN (2011). Practical variable selection for generalized additive models. *Computational Statistics & Data Analysis*, 55, 2372-2387. <https://doi.org/10.1016/j.csda.2011.02.004>.
- NAMMCO (2023). Report of the Scientific Committee Working Group on Dolphins, October 2023, Copenhagen, Denmark. Available at https://nammco.no/dwg_reports/.
- Orden TED/1126/2020, de 20 de noviembre, por la que se modifica el Anexo del Real Decreto 139/2011, de 4 de febrero, para el desarrollo del Listado de Especies Silvestres en Régimen de Protección Especial y del Catálogo Español de Especies Amenazadas, y el Anexo del Real Decreto 630/2013, de 2 de agosto, por el que se regula el Catálogo Español de Especies Exóticas Invasoras. Boletín Oficial del Estado, 314, de 1 de diciembre de 2020. <https://www.boe.es/eli/es/o/2020/11/20/ted1126>
- Owen K, Gilles A, Authier M, Carlström J, Genu M, Kyhn LA, Nachtsheim DA, Ramírez-Martínez NC, Siebert U, Sköld M, Teilmann J, Unger B, Sveegaard S (2024). A negative trend in abundance and an exceeded mortality limit call for conservation action for the Vulnerable Belt Sea harbour porpoise population. *Frontiers in Marine Science*, 11, 1289808. doi:10.3389/fmars.2024.1289808
- Pigeault, R, Ruser, A, Ramírez-Martínez, NC, Geelhoed, SCV, Haelters, J, Nachtsheim, DA, Schaffeld, T, Sveegaard, S, Siebert, U, Gilles, A (2024a). Maritime traffic alters distribution of the harbour porpoise in the North Sea. *Marine Pollution Bulletin*, 208, 116925. doi:10.1016/j.marpolbul.2024.116925
- Pigeault, R, Authier, M, Ramírez-Martínez, NC, Virgili, A, Geelhoed, SCV, Haelters, J, Louzao, M, Saavedra, C, Gilles, A (2024b). Assessing the reliability of species distribution models under changing environments: A case study on cetaceans in the North-East Atlantic. *Global Ecology and Conservation*, 56, e03299. doi:10.1016/j.gecco.2024.e03299
- R Core Team (2024). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.

- Rainey, C (2017). Transformation-Induced Bias: Unbiased Coefficients Do Not Imply Unbiased Quantities of Interest. *Political Analysis*, 25(3), 402-409. doi:10.1017/pan.2017.11
- Rainey, C (2024). A careful consideration of CLARIFY: simulation-induced bias in point estimates of quantities of interest. *Political Science Research and Methods*, 12(3), 614-623. doi:10.1017/psrm.2023.8
- Rogan, E, Breen, P, Mackey, M, Cañadas, A, Scheidat, M, Geelhoed, S, Jessopp, M (2018). Aerial surveys of cetaceans and seabirds in Irish waters: Occurrence, distribution and abundance in 2015-2017. Department of Communications, Climate Action & Environment and National Parks and Wildlife Service (NPWS), Department of Culture, Heritage and the Gaeltacht, Dublin, Ireland. 297pp. https://secure.dccae.gov.ie/downloads/SDCU_DOWNLOAD/ObSERVE_Aerial_Report.pdf
- Taylor, N, Authier, M, Banga, R, Genu, M, Macleod, K, Gilles, A (2022). Marine Mammal By-catch. In: OSPAR, 2023: The 2023 Quality Status Report for the Northeast Atlantic. OSPAR Commission, London. Available at: <https://oap.ospar.org/en/ospar-assessments/quality-status-reports/qsr-2023/indicator-assessments/marine-mammal-bycatch>
- Venables, WN, Ripley, BD (2002). Modern Applied Statistics with S. Fourth edition. Springer, New York. ISBN 0-387-95457-0, <https://www.stats.ox.ac.uk/pub/MASS4/>.
- Vehtari, A, Gelman, A, Gabry, J (2017). Practical Bayesian model evaluation using leave-one-outcross-validation and WAIC. *Statistics and Computing*, 27, 1413-1432. <https://doi.org/10.1007/s11222-016-9696-4>
- Vingada, J, Eira, C (2018). Conservation of cetaceans and seabirds in Continental Portugal. In: The LIFE+ MarPro Project. Rainho & Neves, Lda. Aveiro, pp. 257.
- Virgili, A, Authier, M, Boisseau, O, Cañadas, A, Claridge, D, Cole, T, Corkeron, P, Dorémus, G, David, L, Di-Méglio, N, Dunn, C, Dunn, TE, García-Barón, I, Laran, S, Lauriano, G, Lewis, M, Louzao, M, Manocci, L, Martinez-Cedeira, J, Palka, D, Panigada, S, Pettex, E, Roberts, JJ, Ruiz, L, Saavedra, C, Santos, MB, Van Canneyt, O, Vázquez Bonales, JA, Monestiez, P, Ridoux, V (2019). Combining multiple visual surveys to model the habitat of deep-diving cetaceans at the basin scale. *Global Ecology and Biogeography*, 28, 300-314. <https://doi.org/10.1111/geb.12850>
- Wood, SN, Bravington, MV, Hedley, SL (2008). Soap Film Smoothing. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 70, 931-955. <https://doi.org/10.1111/j.1467-9868.2008.00665.x>
- Wood, S (2017). Generalized Additive Models: An Introduction with R, Second Edition. Chapman and Hall/CRC. ISBN 978-1498728331
- Yao, Y, Vehtari, A, Simpson, D, Gelman, A (2017). Using stacking to average Bayesian predictive distributions. *Bayesian Analysis*, 13(3), 917-1007. <https://doi.org/10.1214/17-BA1091>

APPENDIX

Sightings locations for each species in SCANS-IV (2022). For each species are shown the sightings locations (red dots) overlaid on the modelled density surfaces.

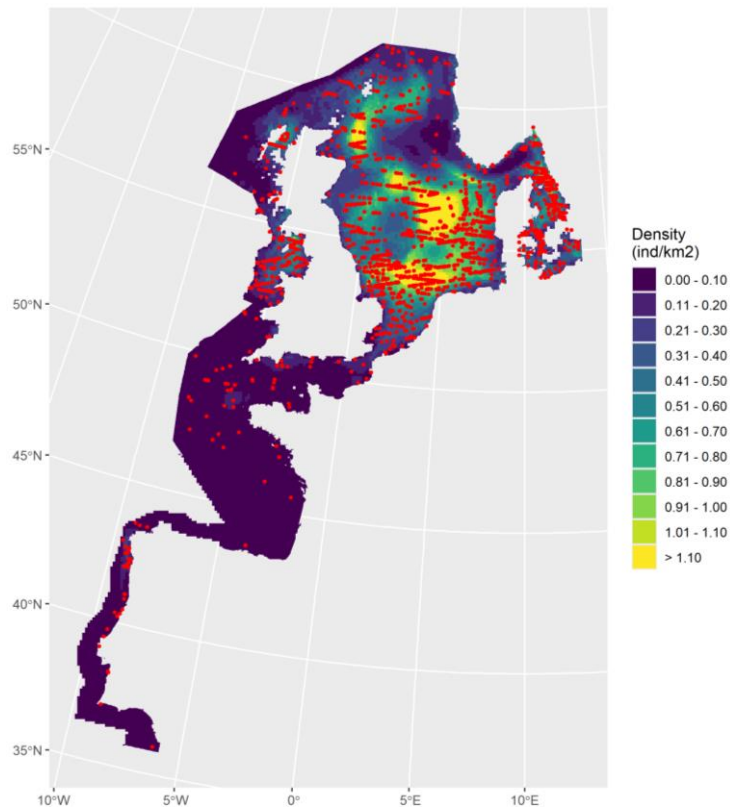


Figure A. 1. Harbour porpoise: sightings overlaid with modelled density surface.

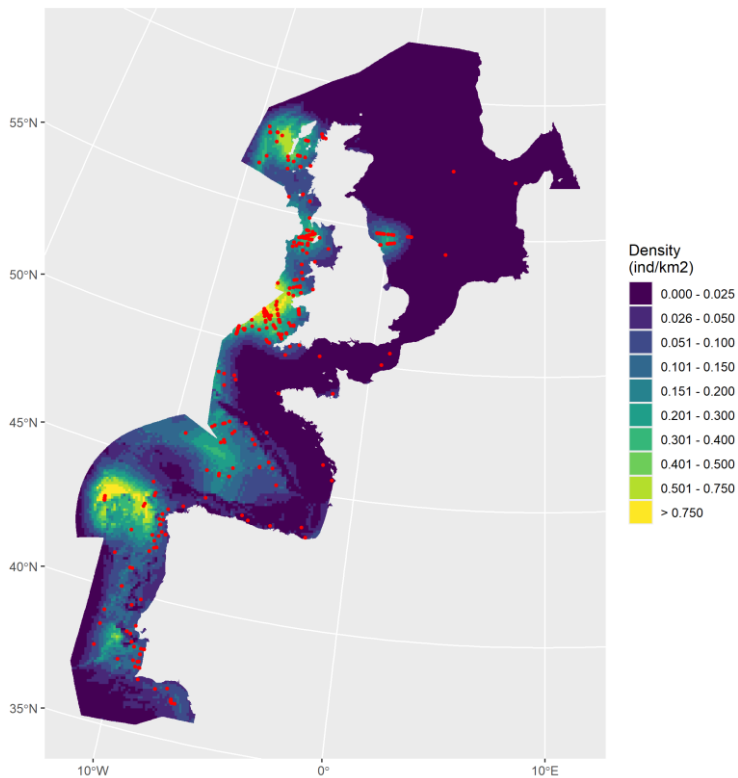


Figure A. 2. Bottlenose dolphin: sightings overlaid with modelled density surface.

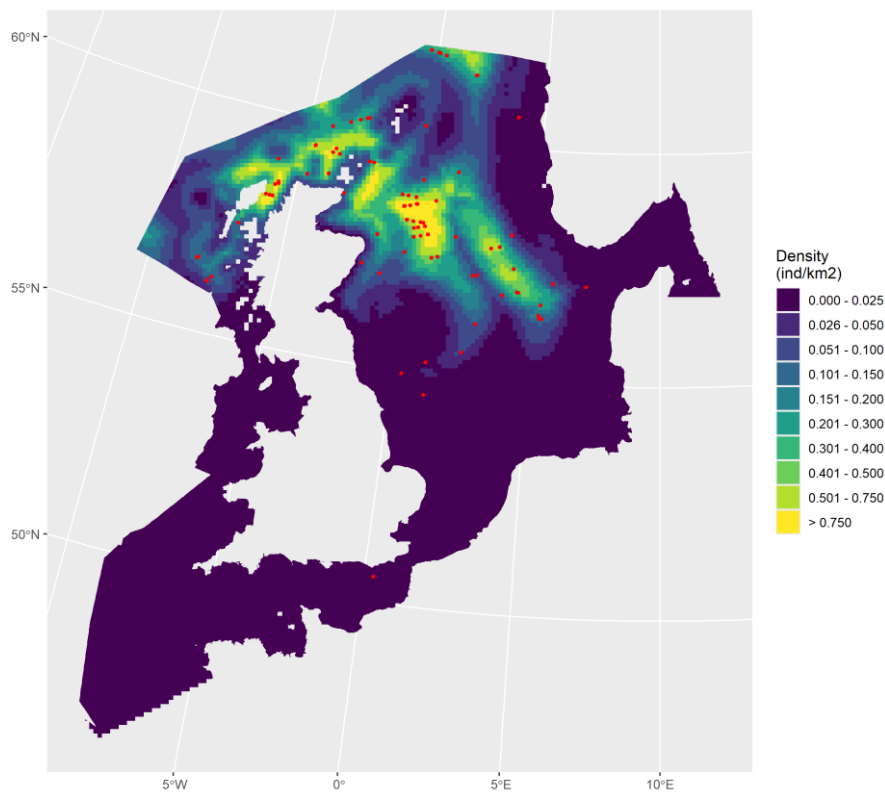


Figure A. 3. White-beaked dolphin: sightings overlaid with modelled density surface.

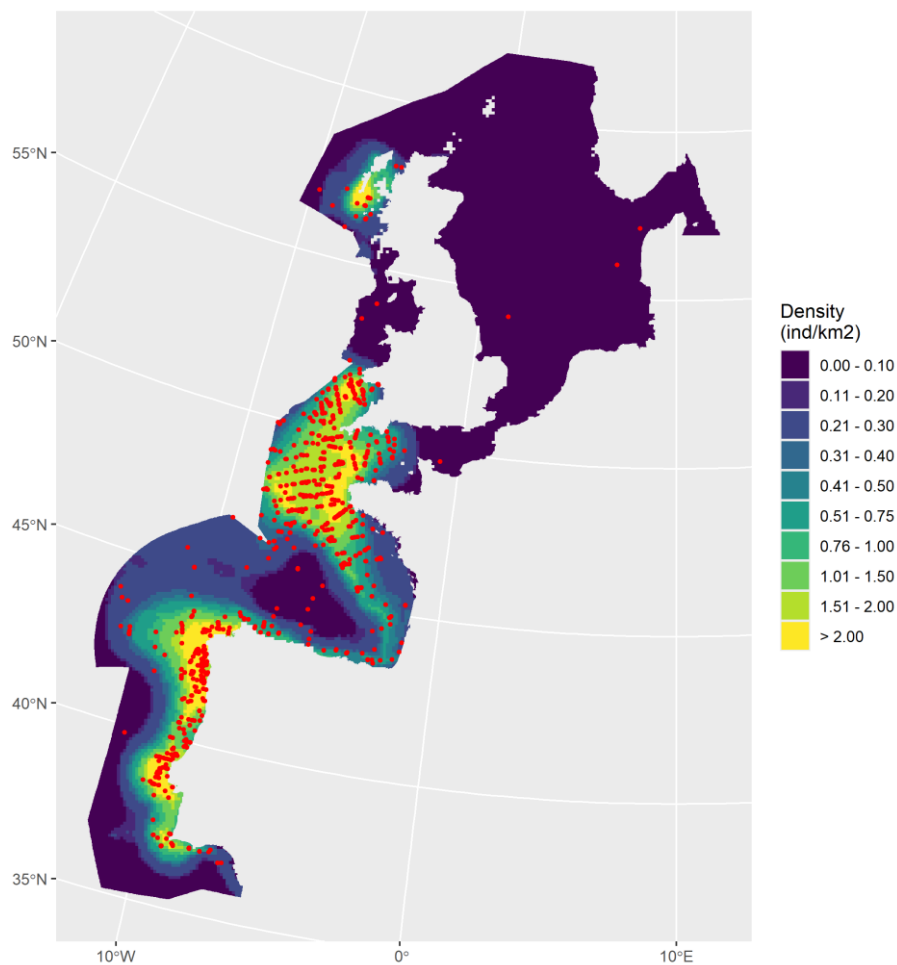


Figure A. 4. Common dolphin: sightings overlaid with modelled density surface.

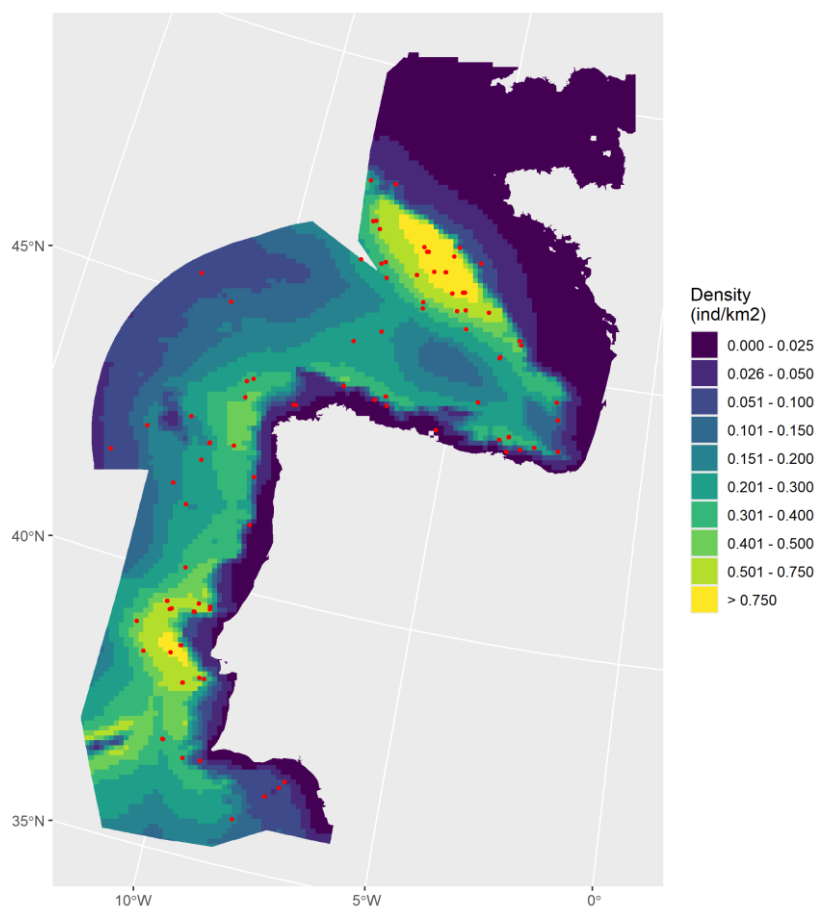


Figure A. 5. Striped dolphin: sightings overlaid with modelled density surface.

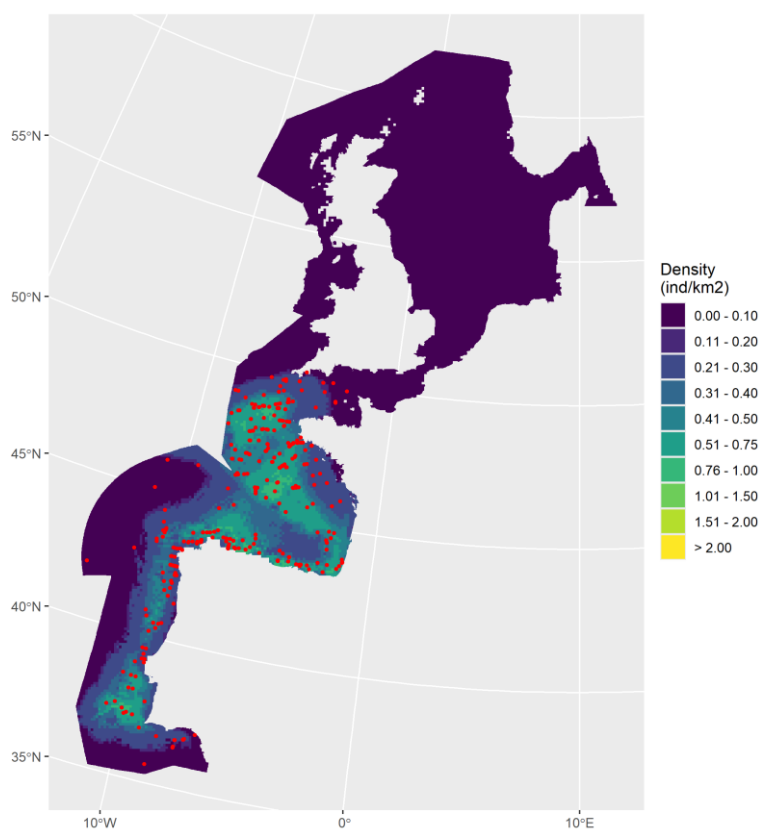


Figure A. 6. Unidentified common or striped dolphin: sightings overlaid with modelled density surface.

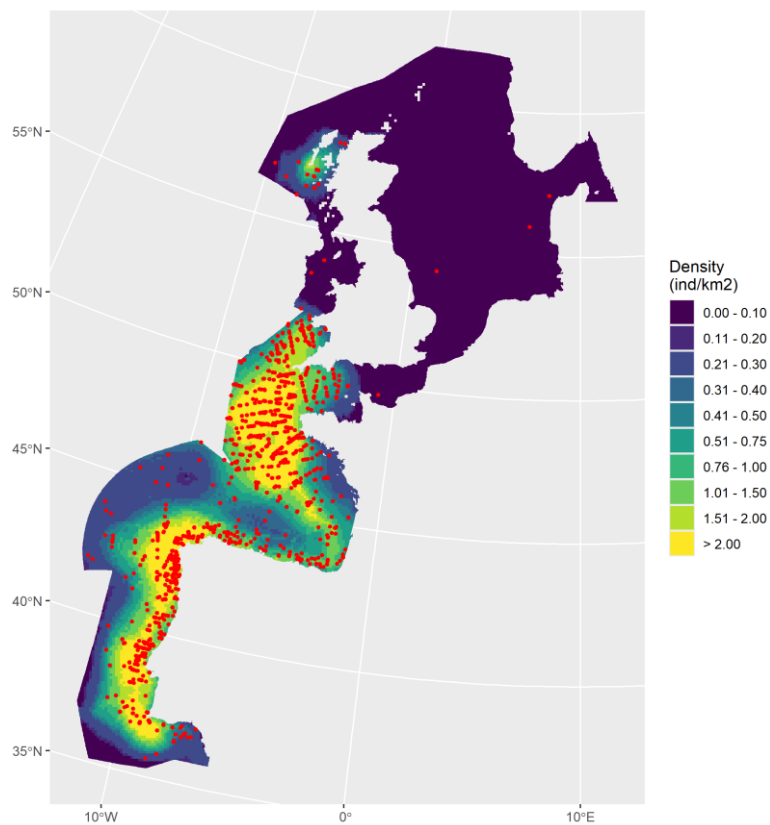


Figure A. 7. Common and striped dolphin, including unidentified common or striped dolphins: sightings overlaid with modelled density surface.

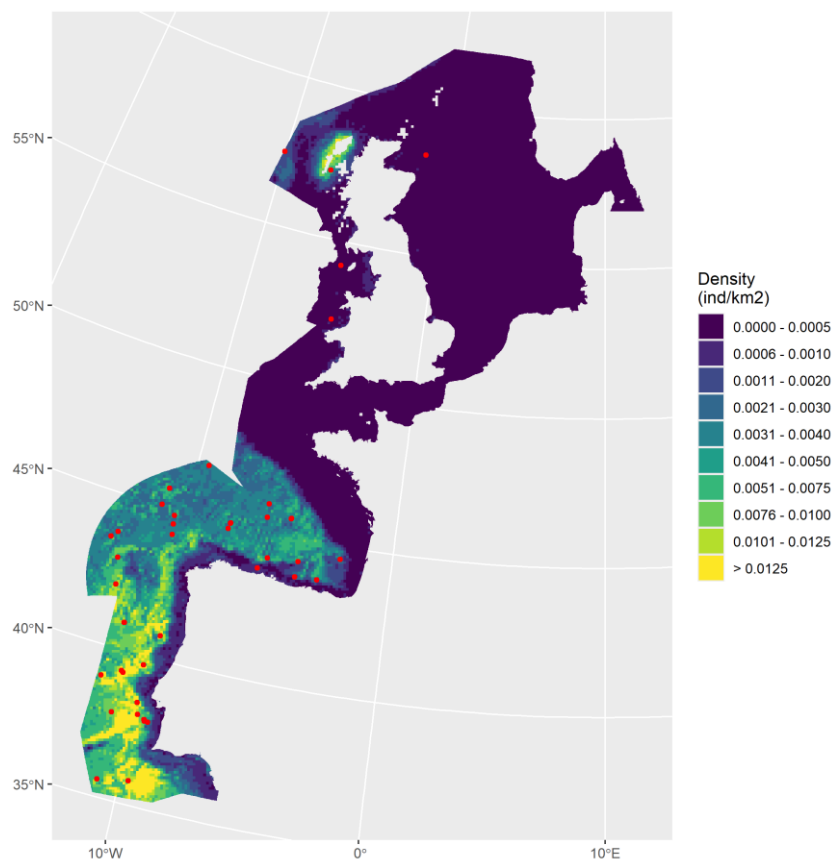


Figure A. 8. Beaked whales (all species): sightings overlaid with modelled density surface.

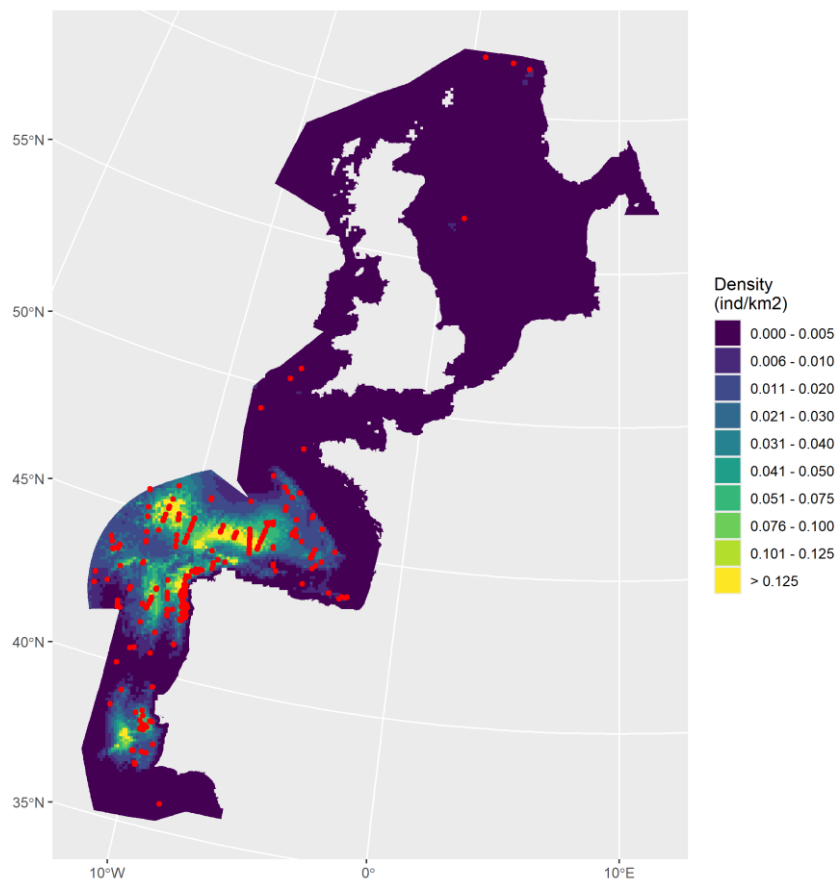


Figure A. 9. Fin whale: sightings overlaid with modelled density surface.

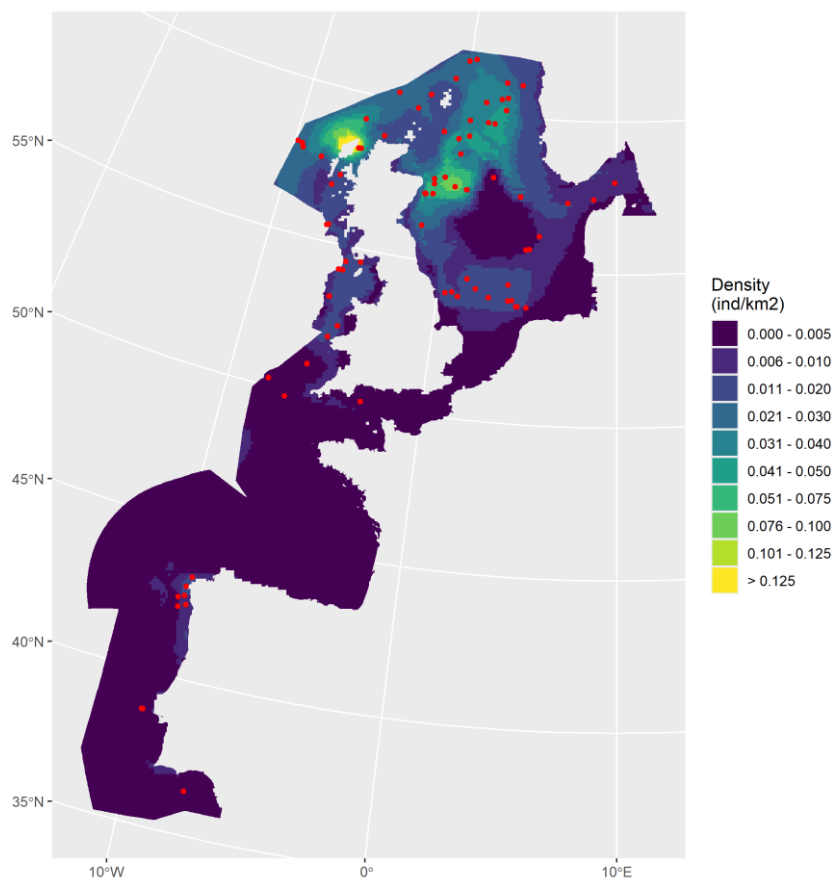


Figure A. 10. Minke whale: sightings overlaid with modelled density surface.

Diagnostic for final models for each species in SCANS-IV (2022). For each species are shown the fitted smooth relationships between relative density and the spatial soap smooth, as well as each covariate selected in the final best model. Covariates are described in **Table 1**.

Y-axes represent function of the term (linear or spline), with the degrees of freedom shown in brackets on the y-axis (linear terms are represented by a single degree of freedom). Zero on the y-axis corresponds to no effect of the predictor variable on the estimated response variable. Scaling of y-axis varies among predictor variables to emphasize model fit. The shading reflects 2× standard error bands (i.e., 95% confidence interval); tick marks (“rug plot”) above the X-axis show data values.

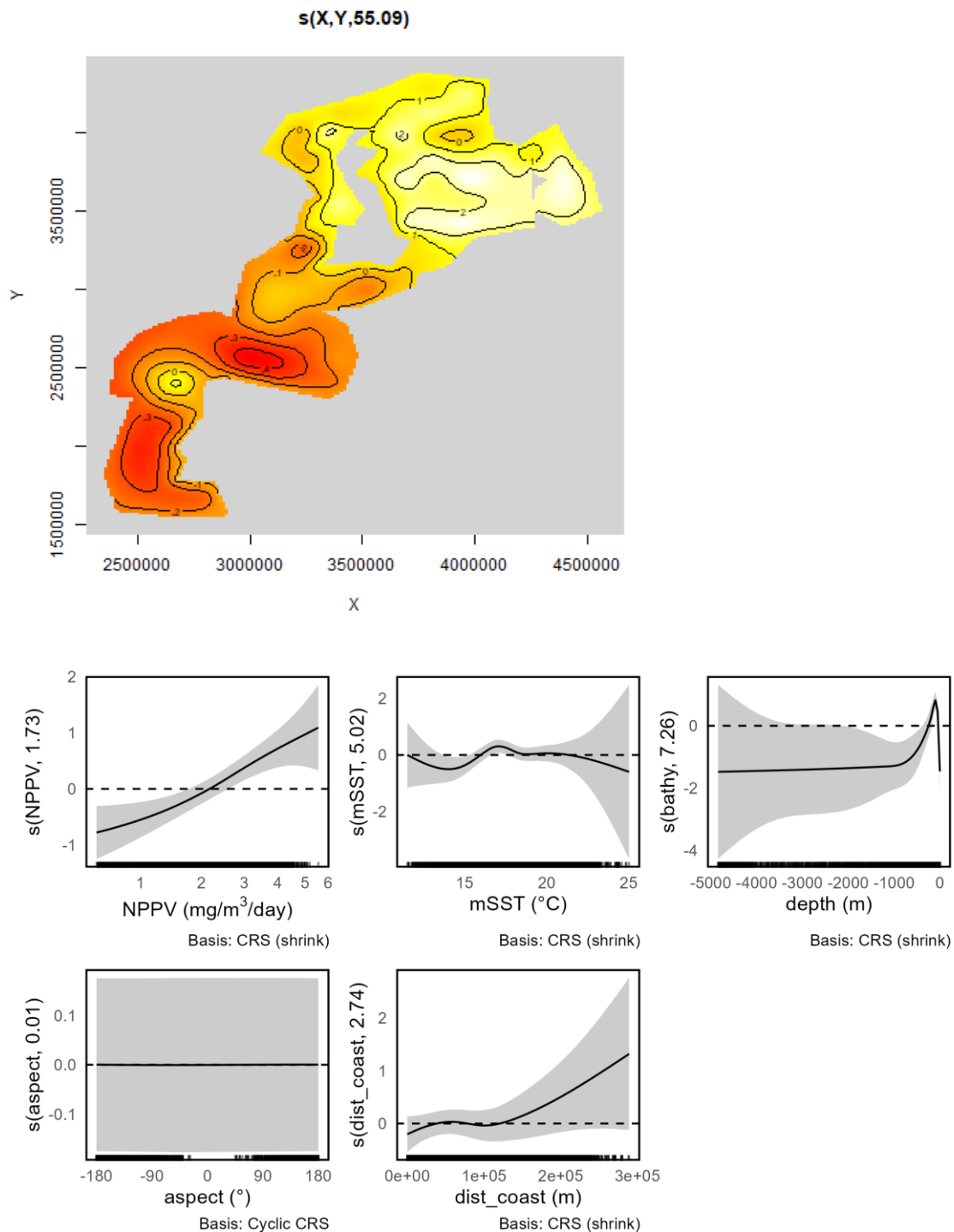


Figure A. 11. Diagnostic for harbour porpoise individuals – see **Table 3**.

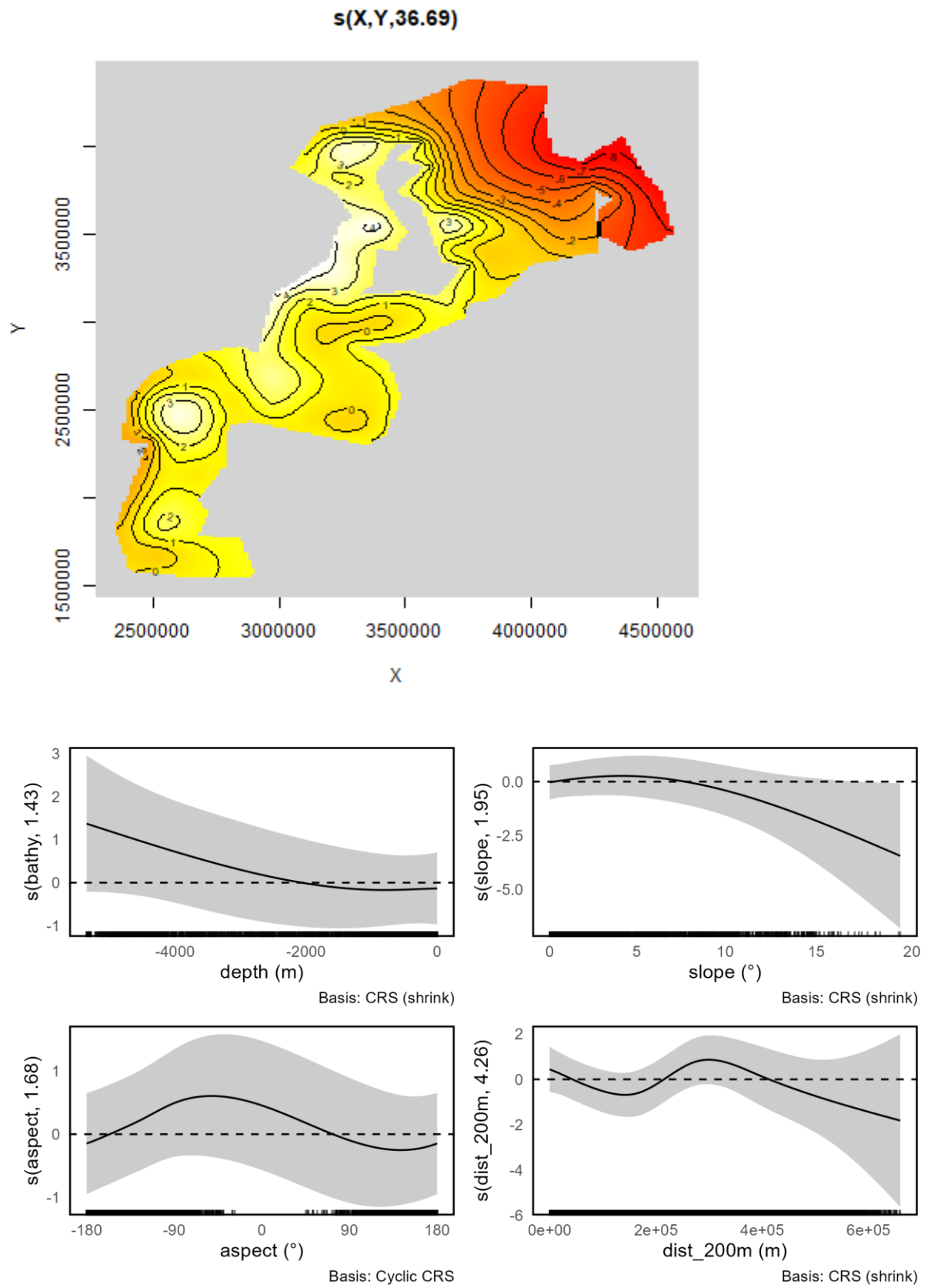


Figure A. 12. Diagnostic for bottlenose dolphin groups – see **Table 4**.

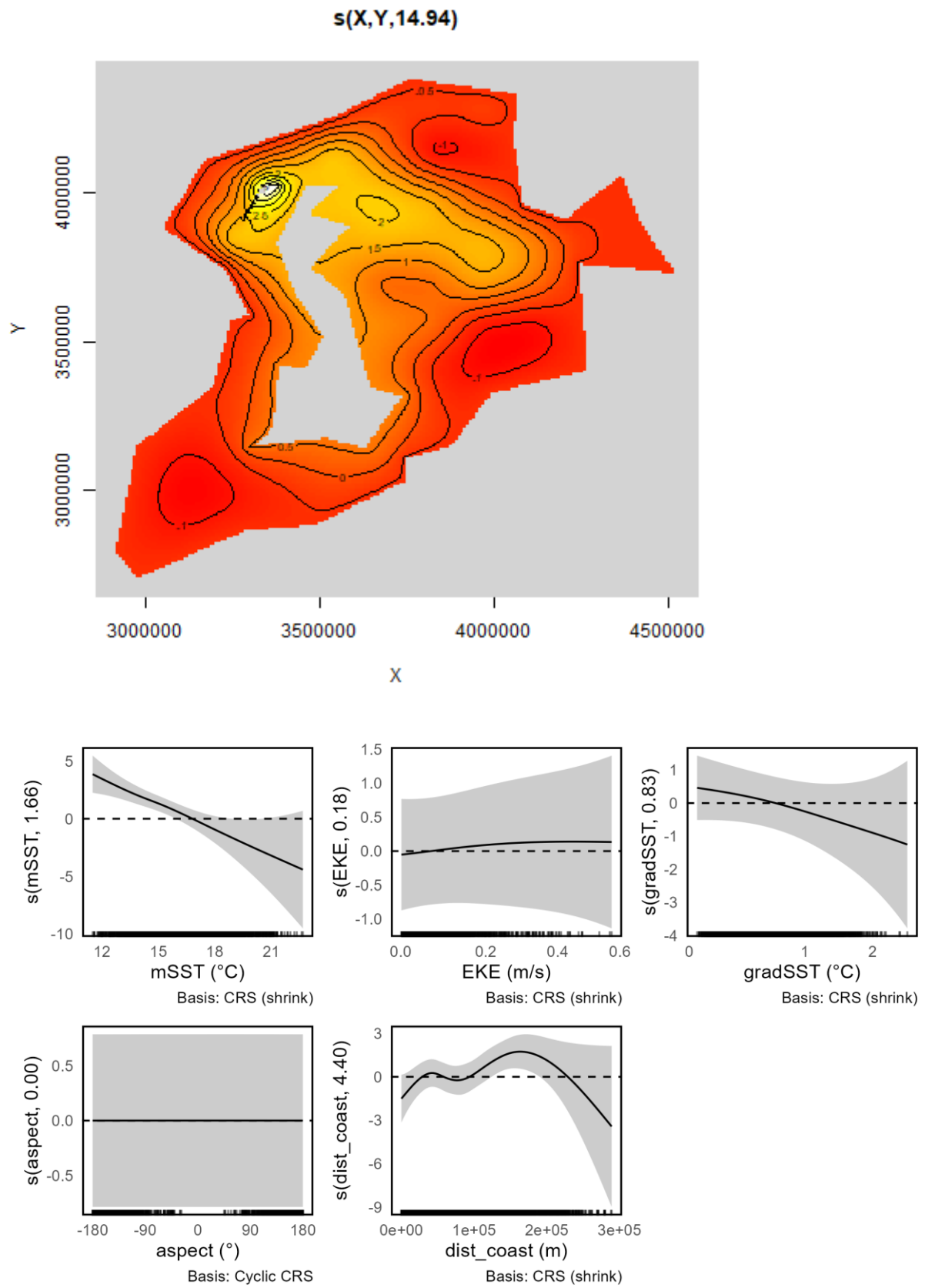


Figure A. 13. Diagnostic for white-beaked dolphin groups – see **Table 5**.

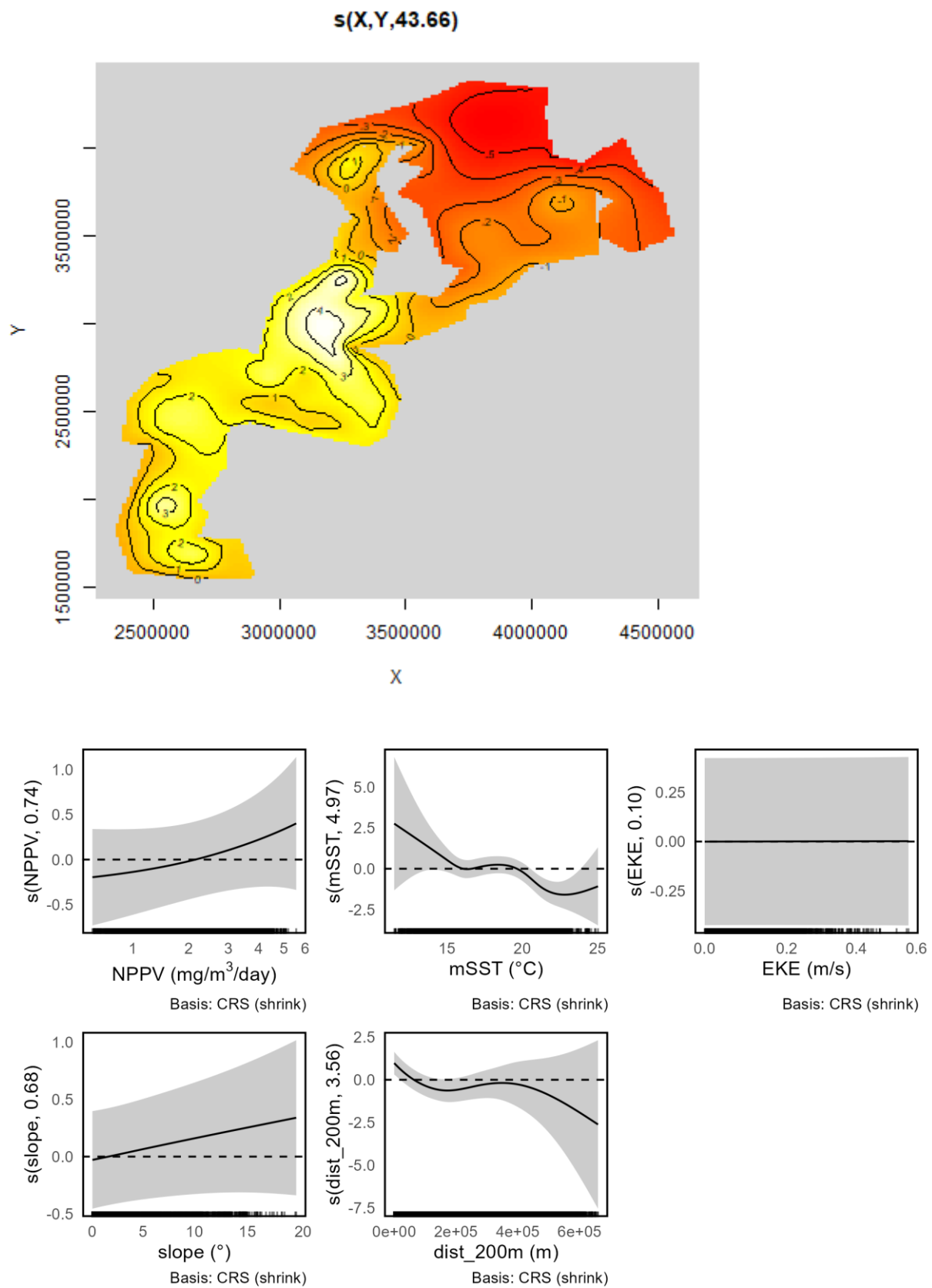


Figure A. 14. Diagnostic for common dolphin groups – see **Table 6**.

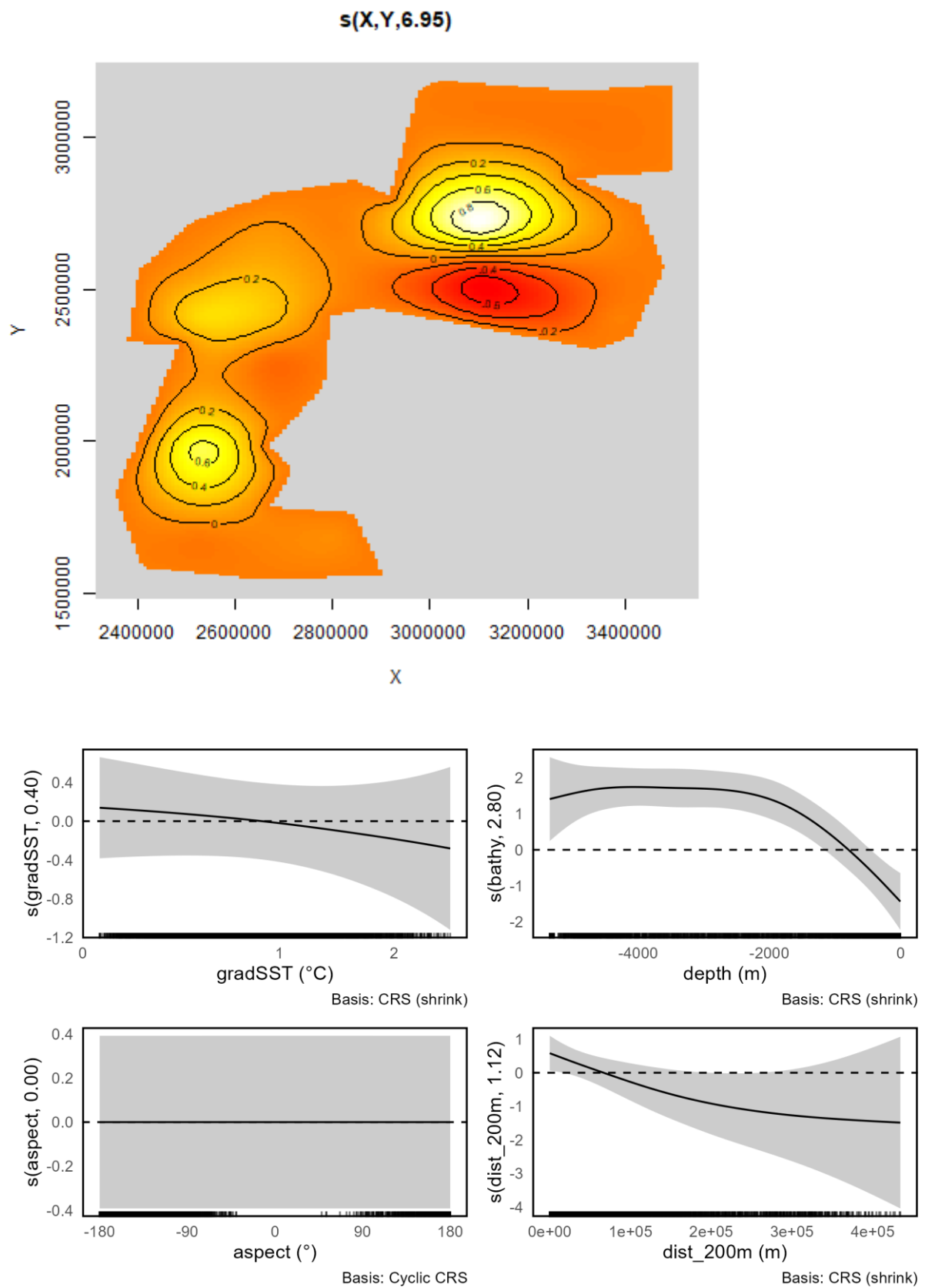


Figure A. 15. Diagnostic for striped dolphin groups – see **Table 7**.

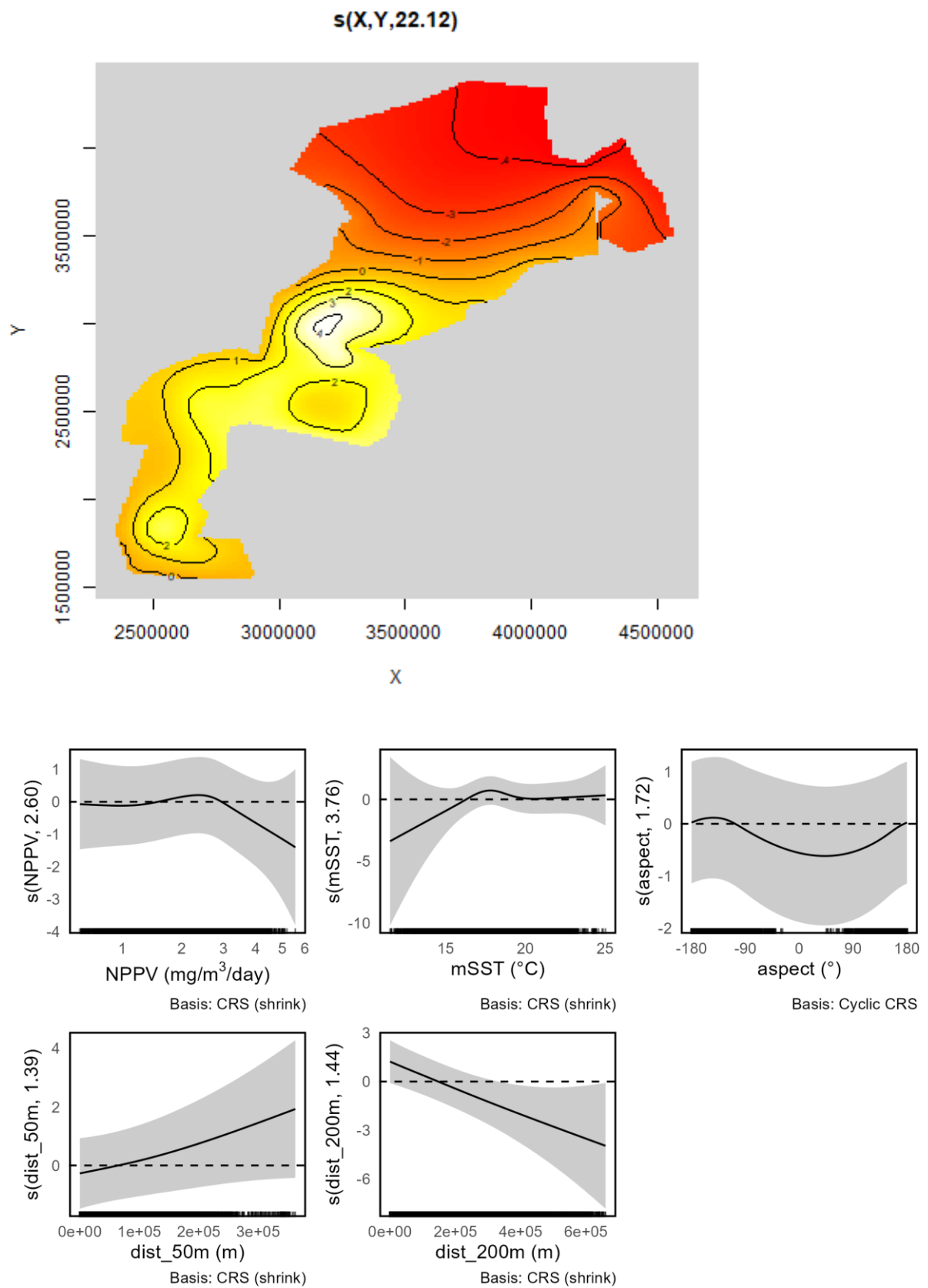


Figure A. 16. Diagnostic for unidentified common or striped dolphin groups – see **Table 8**.

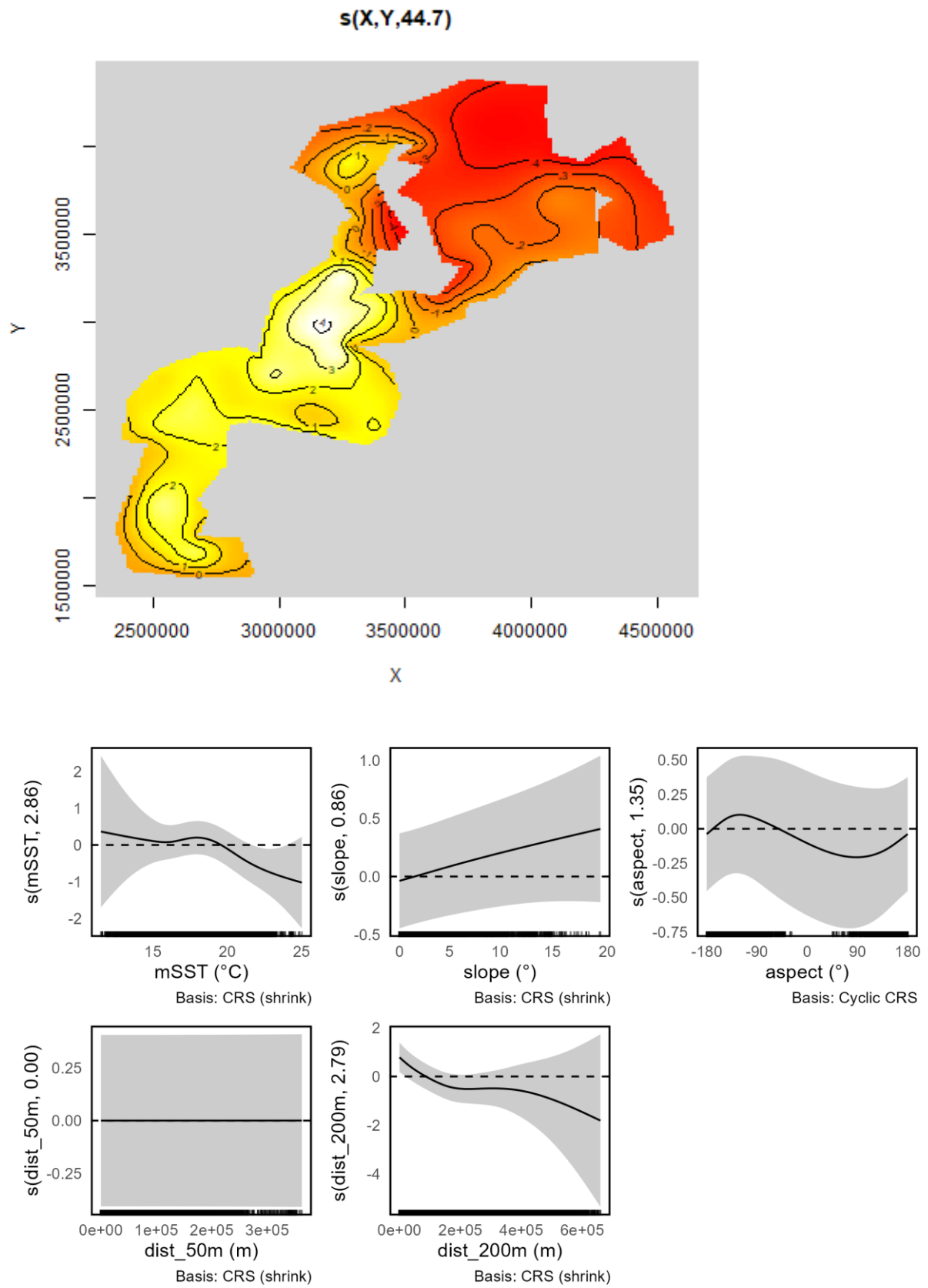


Figure A. 17. Diagnostic for common and striped dolphin, including unidentified common or striped dolphin groups – see Table 9.

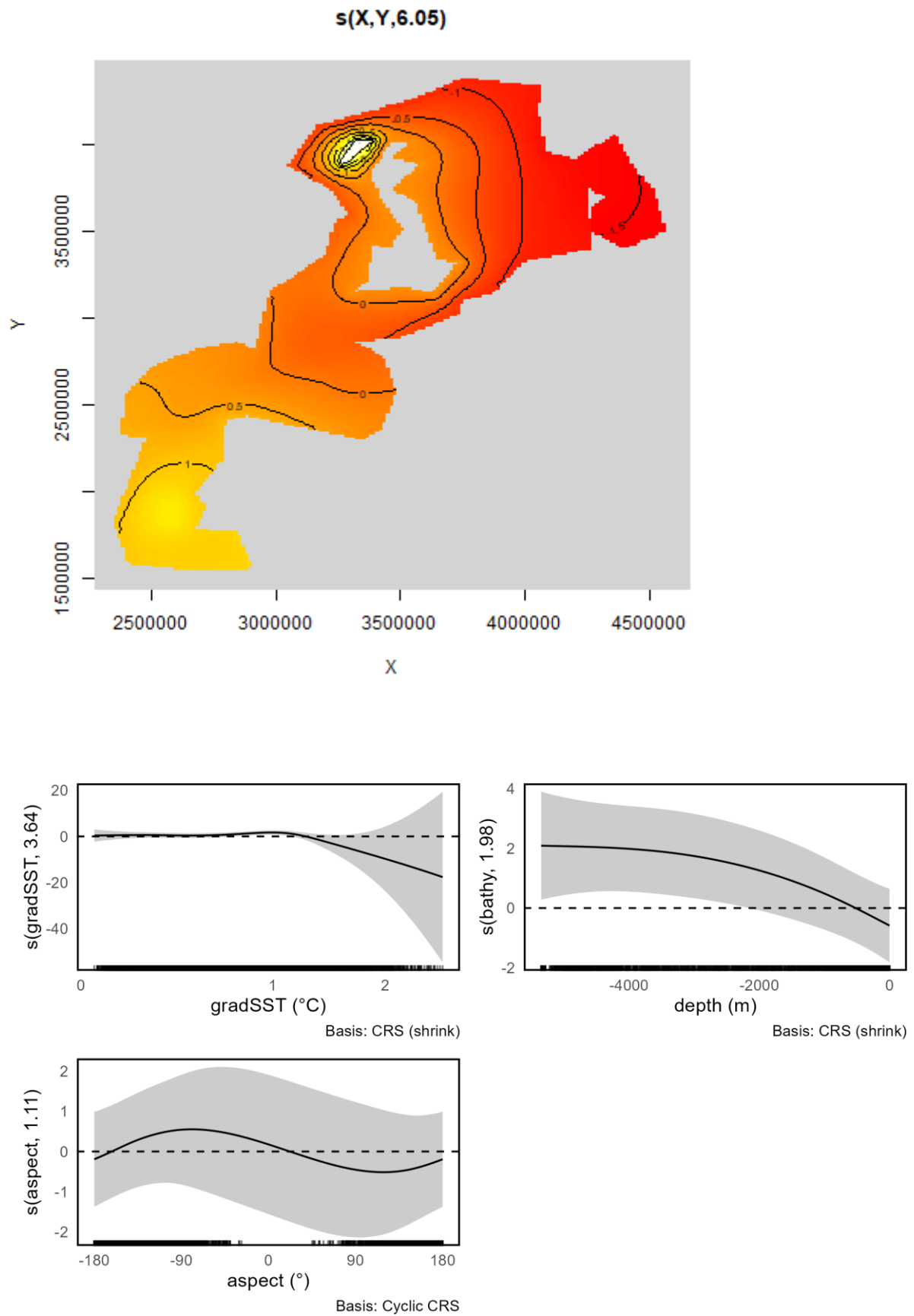


Figure A. 18. Diagnostic for beaked whales (all species combined) individuals – see **Table 10**.

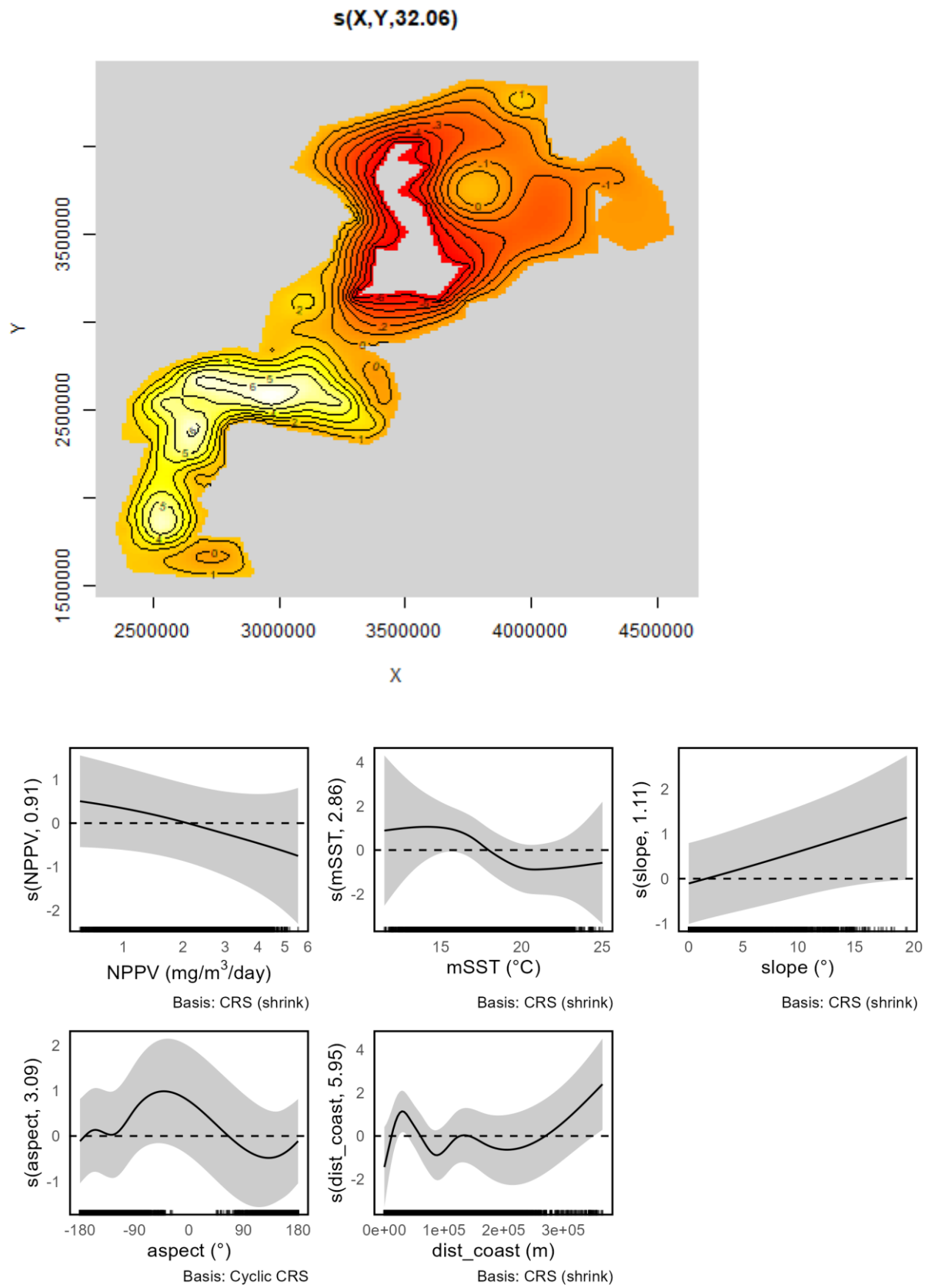


Figure A. 19. Diagnostic for fin whale individuals – see **Table 11**.

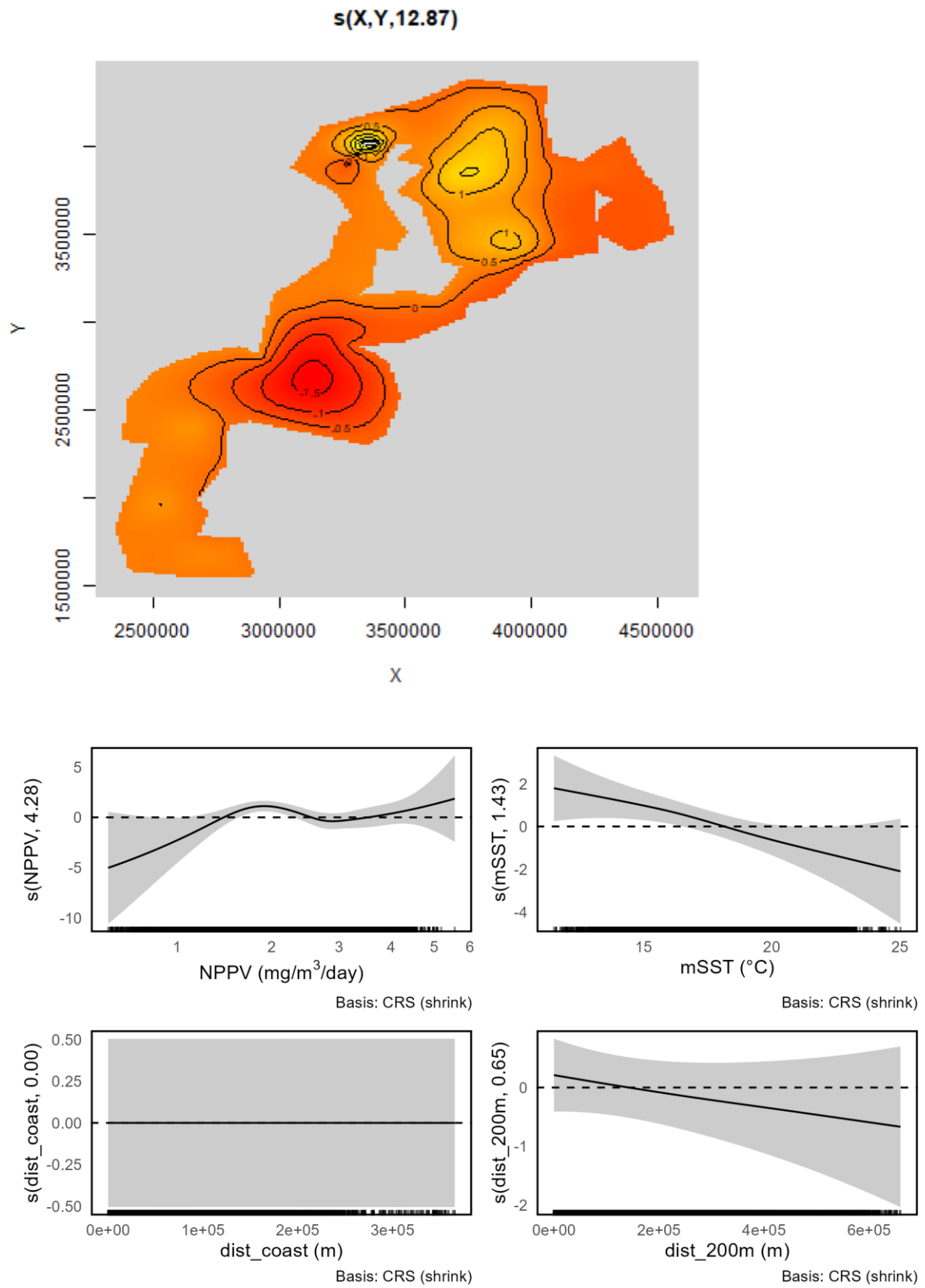


Figure A. 20. Diagnostic for minke whale individuals – see **Table 12.**

Deviance residuals

The Q-Q plot is the plot of the order statistics of the sample against the theoretical quantiles. For generalized linear models, deviance residuals are more appropriate to diagnose model misfit (Ben and Yohai, 2004). Plots display deviance residuals against approximate theoretical quantiles of the deviance residual distribution under the assumption that the model is correct. A well fitted model is indicated by a straight line with a slope of 1 and intercept of 0. The theoretical quantiles are estimated by repeatedly simulating new response data $\{n_i^{\text{new}}\}$ from the fitted model and computing the corresponding residuals (Augustin et al., 2012). Deviance residuals are computed:

$$d_i = \frac{\text{sign}(n_i - \eta_i)}{\sqrt{D_i}} \quad (\text{Eq. A1})$$

where D_i is the likelihood of datum n_i under the assumed model (Eq. 1).

For all taxa, save harbour porpoises, Q-Q plots suggested an adequate model fit. For harbour porpoises, deviance residuals between 0 and 1.5 fall below the confidence bands and those above 2 fall above the confidence bands. This pattern suggests some model misfit, especially for large quantiles as residuals are far from the confidence bands. In this regard, the rootogram offers a valuable insight into the underlying causes of the misfit (see the subsequent section for further details).

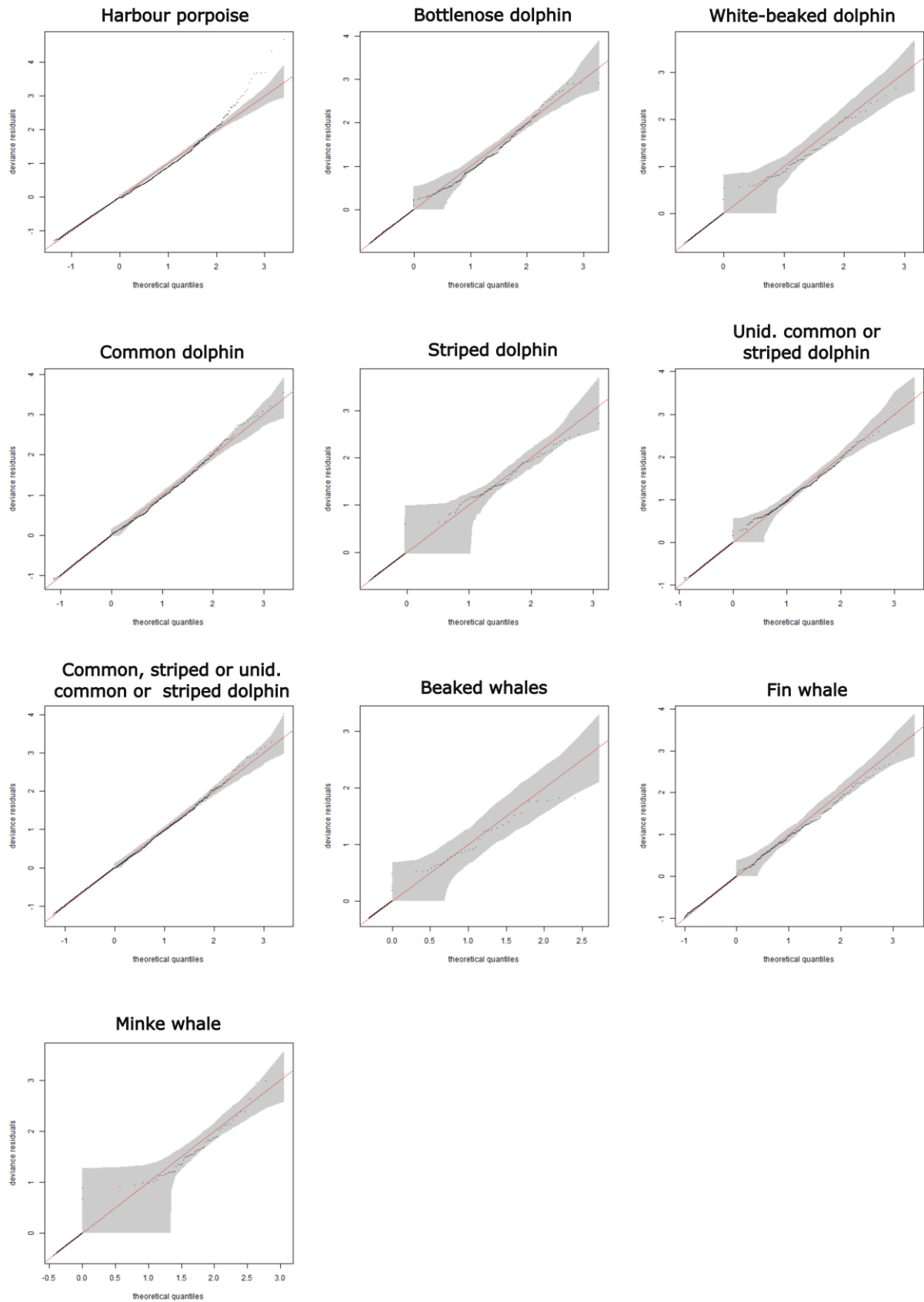


Figure A. 21. QQ plot for each species in SCANS-IV (2022). The QQ plot of deviance residuals plotted against theoretical quantiles is presented for each species.

Rootograms

Rootograms are histograms of count data, often on a square-root scale on the y-axis for better visualisation of small frequencies. Rootograms compare the empirical histogram of the data $\{n_i\}$ with histograms of data generated from a fitted model $\{n_i^{\text{new}}\}$. Once a model has been fitted to data, their parameters have been estimated and can be used to predict new data $\{n_i^{\text{new}}\}$. These new data are data that would be observed should the survey be re-run and under the assumption that the model is correct:

$$\begin{cases} n_i^{\text{new}} \sim \text{NegBin}(\hat{\omega}, \hat{\eta}_i) \\ \log(\hat{\eta}_i) = \log(a_i) + \hat{\vartheta}_0 + \sum_{k=1}^p \hat{f}_k(z_{ki}) \end{cases} \quad (\text{Eq. A2})$$

where the hat notation $\hat{\cdot}$ flags an estimated parameter. Rootograms thus compare the empirical dataset with hypothetical repetition of these data under a fitted model. The rationale for doing so is that a well-fitting model should be able to reproduce well the whole dataset. Badly fitting models are expected to generate new data that are different than the empirical data, whereas a well-fitting model should generate new data similar to the empirical data. Rootograms allow to assess whether a model is prone to over-prediction (e.g. the model generates too many data points with 0 count) or to under-prediction (e.g. the model is unable to generate a maximum count close to the empirical maximum in the data).

For harbour porpoise, and to a lesser extent common dolphin and common or striped dolphins, rootograms suggest some model misfit with an under-prediction of 0-count and an over-prediction of 1-count (**Figure A. 22**). The fitted models are also unable to generate large counts. This pattern in rootogram can explain the model misfit seen in the harbour porpoise Q-Q plot: the model is unable to generate large count, resulting in positive deviance residuals (Eq. A1). On the other hand, the excess 1-count under the model results in negative deviance residuals. This pattern, which is most acute in the harbour porpoise model, is suggestive of excess dispersion above that already taken into account by the choice of a negative binomial likelihood. Other choices of heavier-tailed distributions than the negative binomial may help (e.g. Gorgi, 2020) but are beyond the scope of this report.

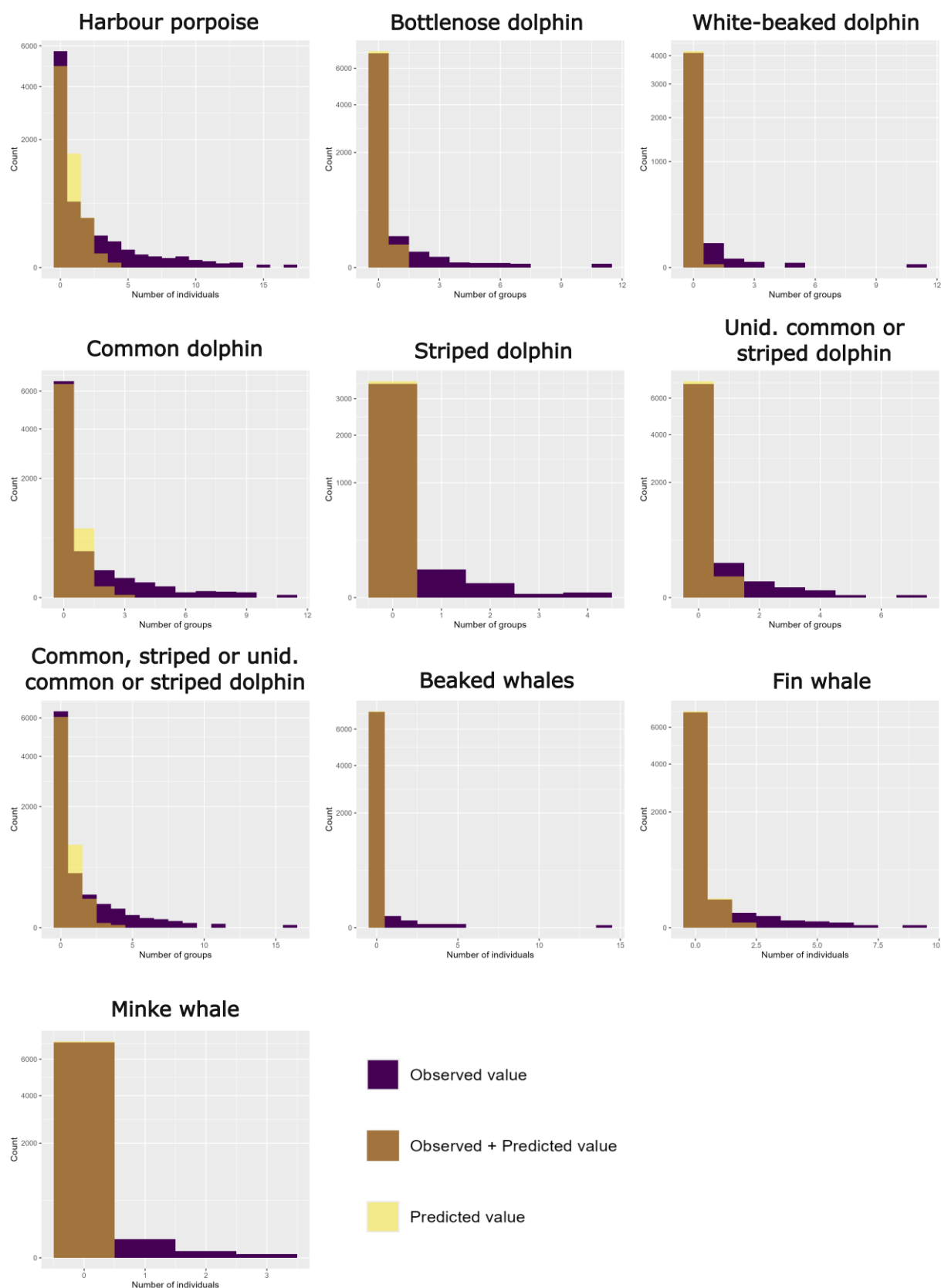


Figure A. 22. Rootogram plot for each species in SCANS-IV (2022). For each species, the rootogram plot is presented, which provides a comparison between the histograms of observed and predicted values for the number of individuals or groups per segment.

REFERENCES

Augustin, NH, Sauleaub, EA, Wood SN (2012) On quantile quantile plots for generalized linear models. *Computational Statistics & Data Analysis*, 56(8): 2404-3409.

Ben, MG, Yohai, VJ (2004) Quantile–Quantile Plot for Deviance Residuals in the Generalized Linear Model. *Journal of Computational and Graphical Statistics*, 13:1, 36-47, DOI: 10.1198/1061860042949_a

Gorgi, P (2020) Beta–negative binomial auto-regressions for modelling integer-valued time series with extreme observations. *Journal of the Royal Statistical Society*, 82 (5): 1325-1347