Call 1

Assessment of species abundance, distribution and habitats in the Pelagos Sanctuary, with a priority given to the Cuvier's beaked whale, the fin whale, the sperm whale and the bottlenose dolphin

INTERMEDIATE TECHNICAL REPORT

Floriane Plard

Independant freelance researcher, Barraque de la Pinatelle, Trémoulet, 15500 Molompize, France, floriane.c.plard@gmail.com

17/03/2024

Introduction

Within the main objectives of the Pelagos Sanctuary to implement the action «Coexistence between marine mammals and use of the seas» of the management Plan and relevant Action Plan 2022-2027, the aim of the call 1 is to estimate cetacean abundances and distributions within the Pelagos Sanctuary. These estimates will provide information to help assess the status of marine mammal populations in the Pelagos Sanctuary, creating an initial status of the cetacean populations and identifying areas of high densities within the Pelagos Sanctuary. The priority has been given to four species, particularly sensitive to human activities: the bottlenose dolphin living close to coast and to human activities, the fin whale impacted by collision with boats, and the sperm whale and the Cuvier's beaked whale.

Abundance and density distribution are key population parameters. Abundance informs about the status of a species in a region, and maps of population densities allow targeting the preferential areas of presence (Wade 1998, Hammonds et al. 2021a, Waggitt et al. 2020). Using these population characteristics grants taking informed decisions about population management and conservation, accounting for threat and risk that might influence the distribution of populations (Freeman 2008, Pace et al. 2015, Jewell et al. 2012).

Population abundance and maps of densities are particularly challenging to estimate in cetaceans (Field et al. 2005, Hammond et al. 2021b, McPherson & Myers 2009). Among several reasons, the natural seasonal and annual extensive movements of these species make their distribution particularly variable among seasons and years. Moreover, the data required to estimate cetacean densities are particularly costly to collect, making them often spatially and temporally sparse and few in quantity. Usually, cetacean abundances in one region are derived from a distance sampling survey covering the region homogeneously along ship or aerial transects (Buckland et al. 1993). While often informative, a given survey reflects the abundance of the population at the time the survey was done. In addition, the low densities of some cetacean species and/or the low probability to detect animals, make the results of only one survey quite uncertain (Hammond et al. 2013; 2021a, Laran et al. 2017, Waggitt et al. 2020).

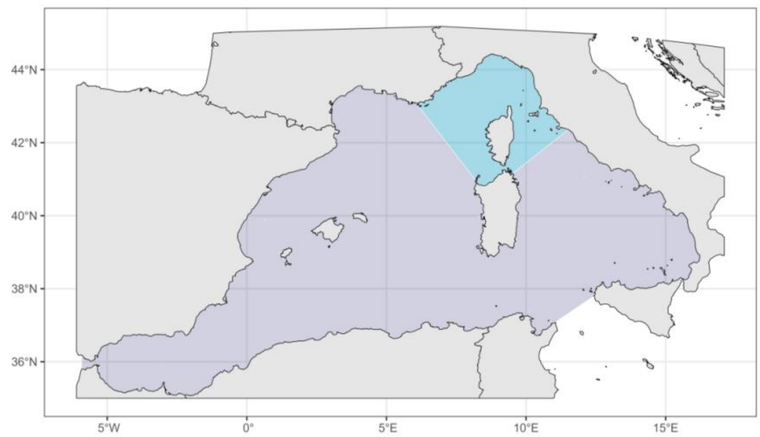


Figure 1: Study area including the MFSD Western Mediterranean region (dark blue) and the Pelagos Sanctuary (cyan).

The aim of this project is to gather multiple aerial and boat cetacean surveys from different areas and to combine them to get more robust estimates of cetacean abundance and densities within the Pelagos Sanctuary. Combining multiple surveys from different areas have several advantages. It increases the precision of the results by gathering more detection data. It also attenuates the temporal bias as different surveys are collected at different times of the year reflecting the variable distribution of species in different seasons. It makes the relationships between environmental variables and cetacean densities more robust. Thus, we chose to extend the studied region of the Pelagos Sanctuary to the Western Mediterranean region (Figure 1) of the The Marine Strategy Framework Directive (MFSD) to increase the number of data gathered and get more robust estimates of cetacean densities and abundances.

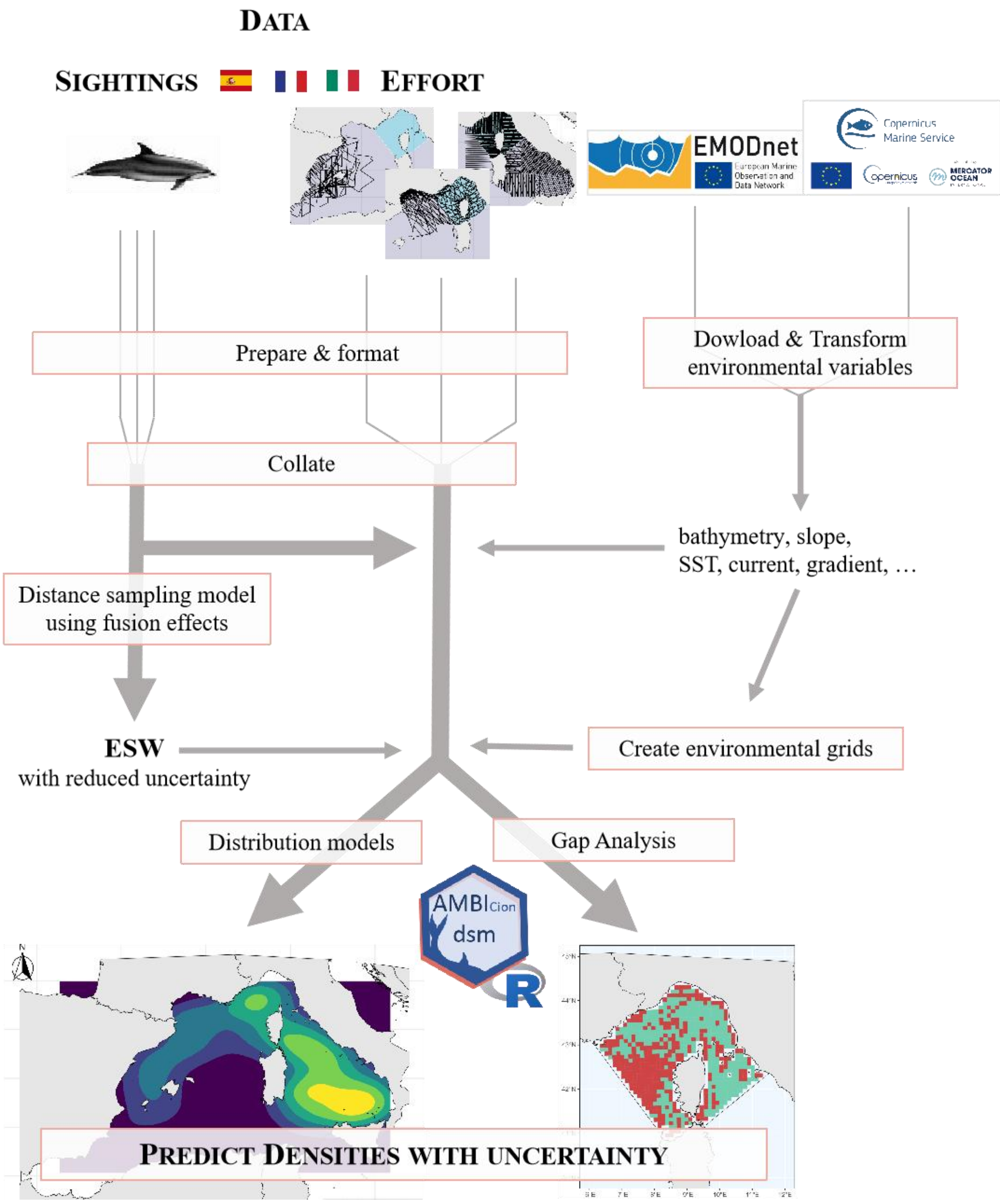


Figure 2: Graphical presentation of the method and tasks that will be realized in this call. Adapted from Plard et al. 2024

The different steps and tasks of this project to estimate cetacean abundance and densities from the gathered data are outlined in Figure 2. After formatting and collating the different dataset, this project will estimate detection probabilities for each species and survey. Then, based on the relationship between marine environmental variables and cetacean densities along survey transects, I will predict model based cetacean densities within the Pelagos Sanctuary. The uncertainty associated with each prediction will be reported in the form of maps of coefficients of variation of densities. Moreover, either the density predictions provided are well informed by data or are fully extrapolated by the model will be analysed using a gap analysis.

1/ Data Call

The Pelagos Secretariat has sent a formal data request to the coordinators of the MFSD reports in Italy, Spain and France. The three countries have kindly agreed to collaborate. Data collected with a distance sampling protocol only have been requested. The three countries have provided the data on effort and sightings from distance sampling surveys they have collected within the Western Mediterranean region (Table 1). Thus, all effort data collected share a common protocol for recording sightings which is the distance sampling protocol (Buckland et al. 1993). At each observation, the distance from the animal to the observer as well as the angle between the animal and the transect line are collected. This information is used to derive the perpendicular distances from the sightings to the transect line and then, to be able to estimate detection probabilities. The detection probability, often decreasing with distance from the transect line, is estimated using a distance sampling model. This method allows estimating corrected animal densities, accounting for imperfect detection of sightings on the transect. Thus, this protocol guarantees robust estimates of densities along effort data.

Table 1: Summary of surveys collected by the Permanent Secretariat of the Pelagos Agreement.

	SurveyID	Platform	Season	Effort (km)	Years	In Sanctuary
FR	SAMM	Plane	fall	5993	2011	Yes
			spring	10908	2012	Yes
			summer	7561	2012	Yes
			winter	19309	2012, 2019	Yes
FR	MOOSE	Ship	spring	2443	2019, 2021	Yes
			summer	141	2021	No
FR	PELMED	Ship	spring	577	2018, 2019, 2021	No
			summer	5708	2017->2021	Yes
IT	Pelagos	Plane	winter	8542	2009	Yes
			summer	8849	2009	Yes
IT	PelaTir_2010	Plane	spring	11219	2010	Yes
			summer	3856	2010	Yes
IT	PelaTir_2020	Plane	fall	18170	2020	Yes

			summer	941	2020	No
IT	ISPRA23A	Plane	fall	6994	2023	Yes
			summer	2781	2023	No
IT	ISPRA23S	Plane	spring	3099	2023	No
			summer	8113	2023	Yes
SP	ICCAT	Plane	spring	12408	2015	No
SP	MEDIAS	Ship	summer	2613	2022, 2023	No
SP	DMESAL	Plane	summer	2096	2023	No
SP	DMLEBA	Plane	summer	5486	2023	No

A/ Italy

The surveys provided by Italy (Figure 3) covered the East part of the Western Mediterranean region, covering the Pelagos Sanctuary in recent and old years. They total 72564 km of effort. These aerial surveys included the ISPRA surveys conducted in summer and fall 2023, the PelaTir surveys conducted in 2010 and 2020, and the oldest survey collected so far; the Pelagos survey conducted in winter and summer 2009. However, for this last survey, I still did not receive the sightings data in an exploitable format. Nevertheless, as I expect to get these cetacean sightings soon, the effort of this survey is included in this report.

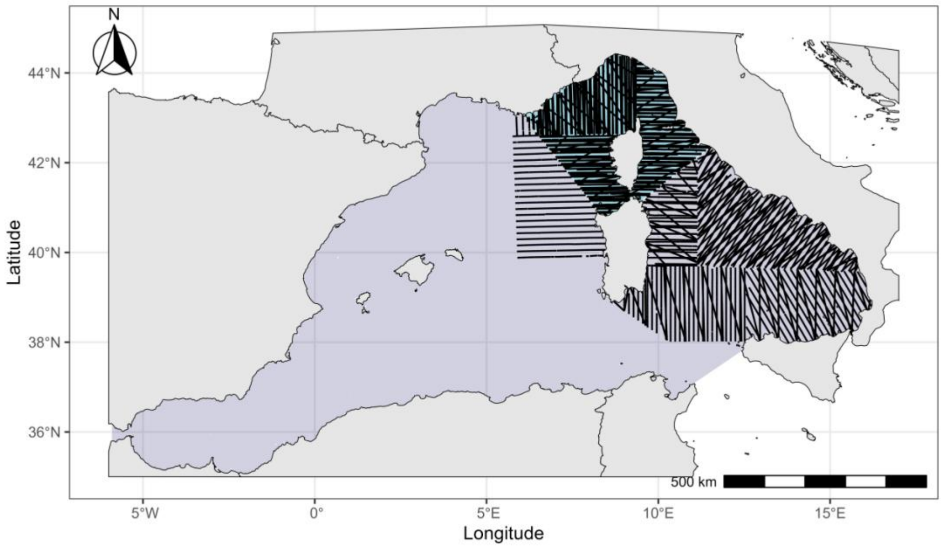


Figure 3: Spatial coverage of aerial surveys shared by Italy.

B/ France

The surveys provided by France (Figure 4) include two ship surveys MOOSE and PELMED conducted from 2017 to 2021 and the aerial large survey SAMM conducted in 2012 and 2019. They cover mainly the Gulf of Lion and the Pelagos Sanctuary. They total 52640 km of effort.

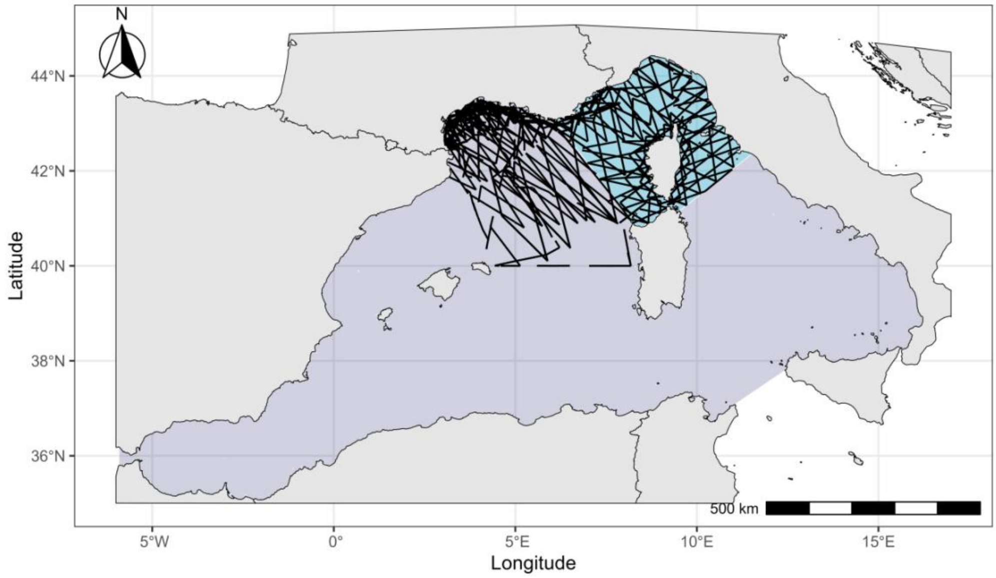


Figure 4: Spatial coverage of surveys shared by France.

C/ Spain

The surveys provided by Spain (Figure 5) cover the West part of the Western Mediterranean region covering the coast and waters of Spain. The survey data gathered so far total 22603 km of effort. Spain surveys include a recent ship survey MEDIA conducted in 2022 and 2023 along the coasts of Spain. Two aerial surveys DMESAL and DMLEBA have also been collected in 2023 in the Spain waters. A third aerial survey ICCAT is available. This survey flying over the Balearic islands originally aimed to collect information about tuna species, but also collected data on marine megafauna. I have received the data of 2015 for this survey. Data from other years are also available. As I received them mid-March; they have not been included in this report. In this report ICCAT data from 2015 only have been included.

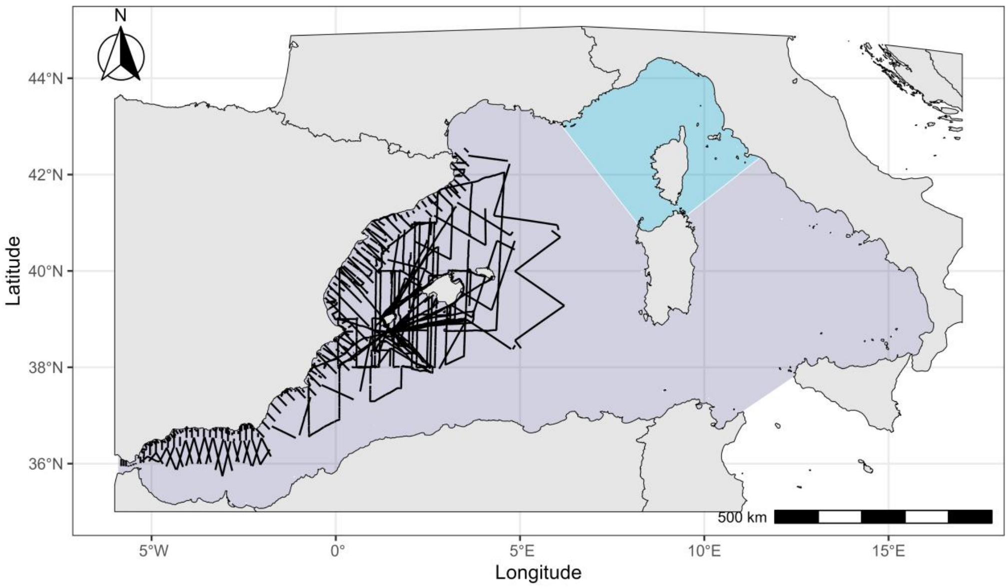


Figure 5: Spatial coverage of surveys shared by Spain.

2/ Data Preparation

A first important step to collate data from different surveys is to homogenize and format all data in the same way so they can be used together. In this step, effort data of each survey have been checked to include effort collected when at least one observer was present. Effort over land, during circle backs and unstandardised effort were excluded. Boat and plane speed, and plane altitude greatly influence detection probability. To start analyses, effort data were kept if the speed was between 8 and 16 knots for boats and between 50 and 150 knots for planes. Plane altitude was limited from 150 to 400 meters. These values might be restricted if high heterogeneity in detection probability is revealed in the distance sampling model.

In a second step, efforts were linearised to avoid increasing the actual covered and sampled area. Effort data were then cut into segments of about 10 km, the conditions of observation including the number of observers, the altitude of observation, the sea state and the subjective conditions remaining similar along each segment. After this preparation, the effort data were collated.

The sightings data were also checked and prepared. Data were included if information about species name, group size, spatial location, date and time of the sighting, perpendicular distance, and observation side were available. Using this information, each sighting was linked to the effort segment it was observed on, in order to gather the condition of observation of each sighting that influenced detection probability.

A first filter was applied on species name and the following species were retained: fin whale *Balaenoptera physalus*, common dolphin *delphinus delphis*, risso's dolphin *Grampus griseus*, pilot whale *Globicephala melas*, sperm whale *Physeter macrocephalus*, striped dolphin *Stenella coeruleoalba*, bottlenose dolphin *Tursiops truncatus,* Cuvier's beaked whale *Ziphius cavirostris*, and sightings of unidentified ziphius species, large whales, small and medium cetaceans.

A/ Bottlenose Dolphin

361 sightings of individuals or groups of bottlenose dolphins (Figure 6) are available in the collated and prepared dataset. Many of these sightings have occurred in the Gulf of Lion and in the Pelagos Sanctuary.

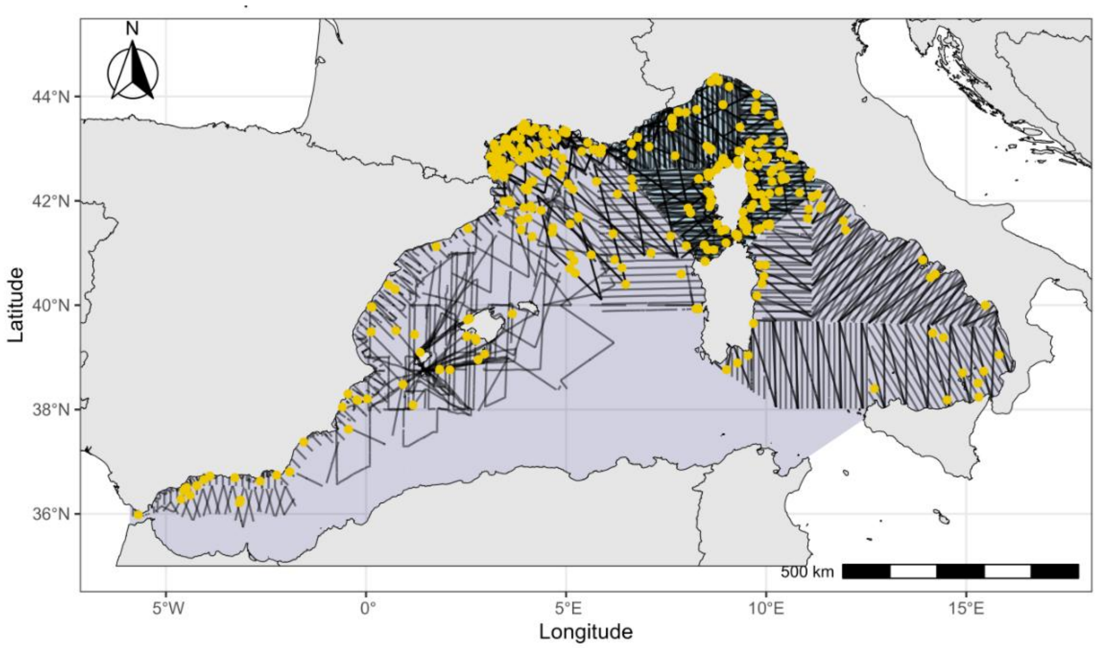


Figure 6: Spatial distribution of the sightings of Bottlenose dolphins on gathered effort.

B/ Sperm whale

59 sightings of individuals or groups of sperm whales (Figure 7) are available in the current collated and prepared dataset. A lot of them occurred in the Pelagos Sanctuary.

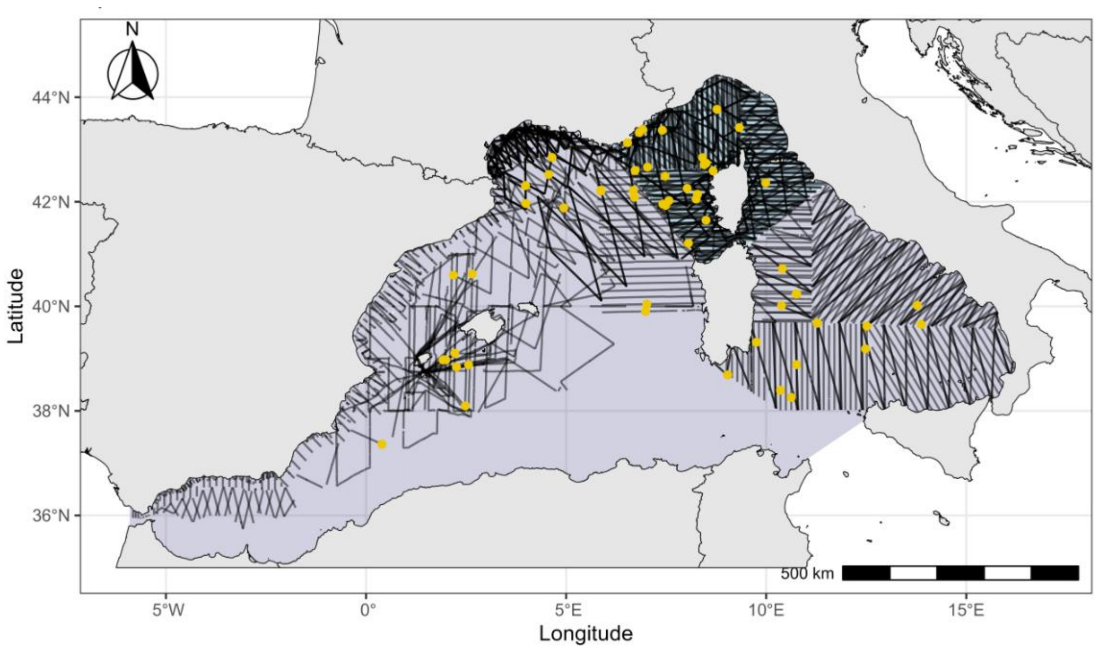


Figure 7: Spatial distribution of the sightings of Sperm whales on gathered effort.

C/ Fin whale

234 sightings of individuals or groups of fin whales (Figure 8) are available in the current collated and prepared dataset. Most of these sightings were observed in the Gulf of Lion and in the Pelagos Sanctuary.

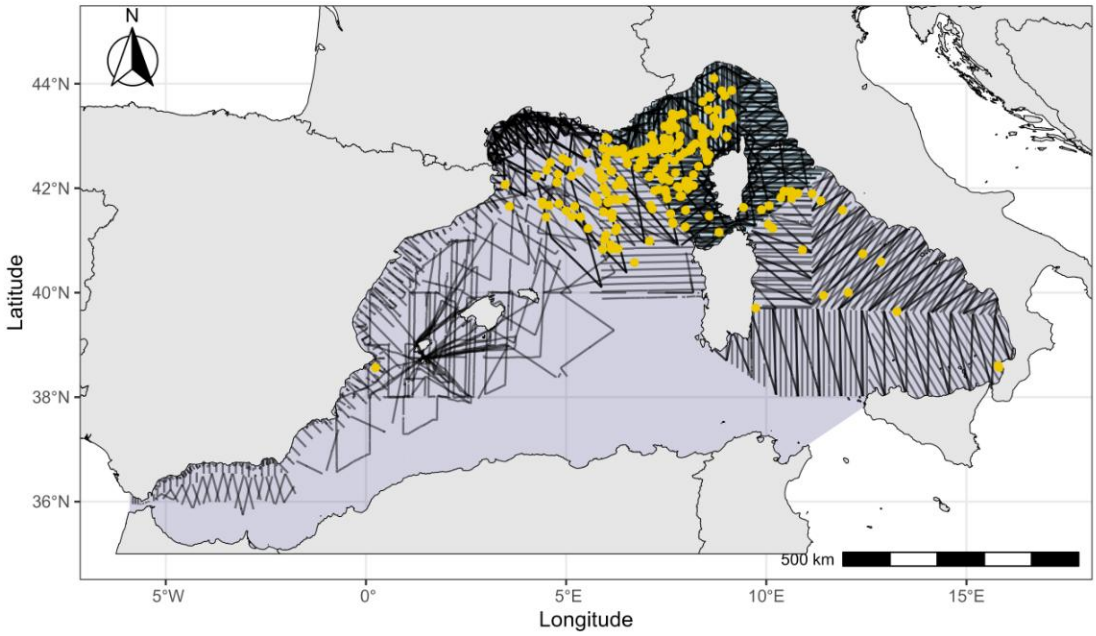


Figure 8: Spatial distribution of the sightings of Fin whales on gathered effort.

D/ Cuvier's beaked whale

29 sightings of individuals or groups of Cuvier's beaked whales (Figure 9) and 6 sightings of unidentified ziphius species are available in the current collated and prepared dataset. I intend to group both types of sightings to increase the number of data points to analyze the densities of Cuvier's beaked whale in the Western Mediterranean region as the sightings are widespread in the region.

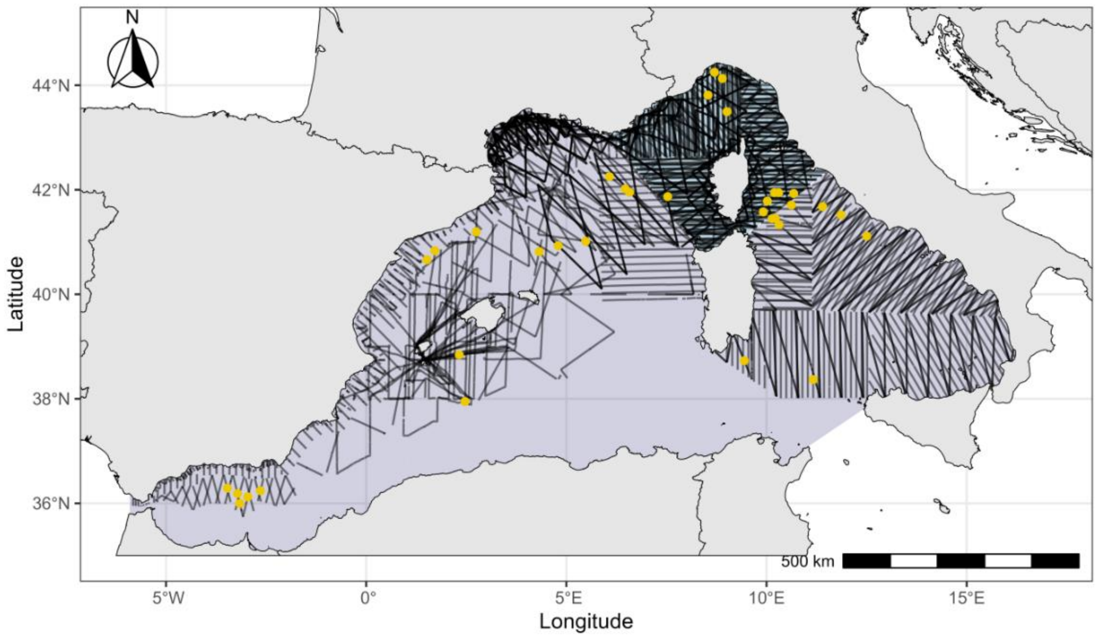


Figure 9: Spatial distribution of the sightings of Cuvier's beaked whales on gathered effort.

E/ Other species

Among other species available, one observation of rough-toothed dolphin (*Steno bredanensis*) has been removed. 40 sightings of long-finned pilot whales, 80 sightings of Risso dolphins, 35 sightings of common dolphin and 1792 sightings of striped dolphins (Figure 10 for this last species) were kept in the current collated and prepared dataset. They will be used to improve the estimates of detection probability in the distance sampling model (see below).

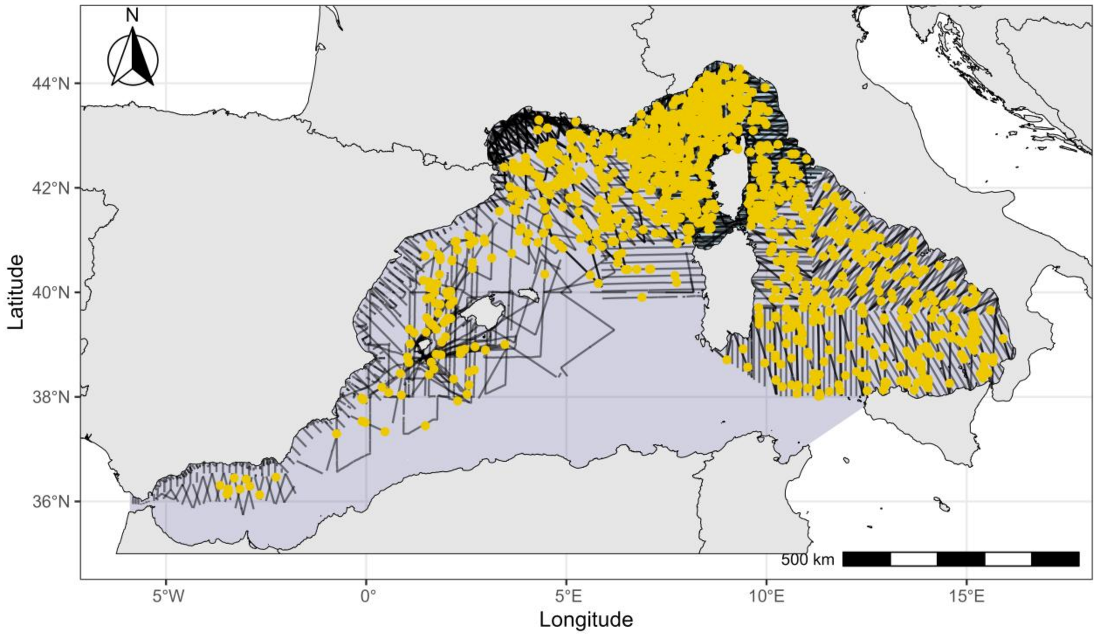


Figure 10: Spatial distribution of the sightings of Striped dolphin on gathered effort.

3/ Environmental Variables

The densities of cetacean in the Pelagos Sanctuary will be predicted from density surface models (Buckland et al. 2015, Miller et al. 2013) and marine environmental variables characterizing the area using relationships between environmental variables and cetacean densities estimated along the effort data. These models will not directly reflect the habitat of cetacean populations or their true distribution, but will reflect maps where cetaceans are expected to occur.

Marine environmental variables have been used in previous studies to predict the density and distribution of marine mammals (Astarloa et al. 2021, Virgili et al. 2019, Waggitt et al. 2020). The two main challenges we face with cetacean species are their wide distribution and large movements making their distribution highly dynamic among and within years. Their distributions result from complexe interactions of ecological processes including oceanographic and biological components (Croll et al. 1998, Barlow et al. 2020). Food availability plays a major role for their distribution (Benoit-Bird & Au, 2003; Hastie et al., 2004; Frederiksen et al., 2006). Unfortunately, robust data on the dynamics distribution of most prey of cetaceans are not available within the whole studied region, preventing us from directly using prey data to predict cetacean densities (Guisan & Zimmermann, 2000). Nevertheless, prey distribution can be correlated to oceanographic and physiographic environmental variables easier to collect at large spatial scale and with higher robustness (Forney 2020). These oceanographic variables are thus relevant to predict cetacean densities (Redfern et al., 2006, Forney 2020).

A candidate set of static and dynamic variables has been selected based on their use in previous models of cetacean densities and their accessibility at large spatial scale. They are summarized in table 2.

Table 2: Marine environmental covariables downloaded to predict densities of cetacean species.The static variable bathymetry was downloaded from the website of EMODnet (https://emodnet.ec. europa.eu/en/bathymetry). Other static variables: slope, aspect and roughness of the sea floor were derived from the variable bathymetry using the function terrain of the package terra (Hijmans et al. 2022) in the statistical software R. The dynamic variables: sea surface temperature, net primary productivity and marine currents were downloaded at a monthly temporal coverage from 2001 to 2023 from the website of Copernicus Marine Service (https://data.marine.copernicus. eu/products). Monthly temporal scale was chosen as the trade-off to minimize the amount of data download and to maximize potential predicted scale. Sea surface temperature gradient was derived from the sea surface temperature using the function DetecFronts of the package grec (Wencheng 2024) in the statistical software R.

Environmental variable		Original scale	Justification	Source
Static	Bathymetry (m)	1/16 arc minute	Deep and shallow column waters influence the presence of variable preys (e.g. squids or fish species)	EMODnet
	Slope (rad)		Associated with currents, high slopes induce enhanced primary production or prey aggregation	
	Aspect (rad)		Describe currents and prominent structures such as canyons, seamounts or mountain chains, used as proxies for predator hotspots and useful in locations where access to biological data is limited	
	Roughness (m)		Describe micro-irregularities in space, revealing heterogenous sea floor potentially enhancing diverse habitats and fish assemblages	
Dynamic	Sea surface temperature (SST) mean (°C)	0.083 degree	Variability of SST over time and space influences directly prey and cetacean distributions	Copernicus

Sea surface temperature gradient (°C/m)		Horizontal gradients of SST reveal front locations, mixing of water and is associated with enhanced primary
Eddy kinetic Energy (EKE ; m/s)		High EKE are linked to the development of eddies, upwelling of nutrients and enhanced primary production, which induce prey aggregation
Net primary productivity (NPPV ; mg.m-3.day-1)	0.25 degree	Net primary production is a proxy of zooplankton distribution, feeding cetacean preys

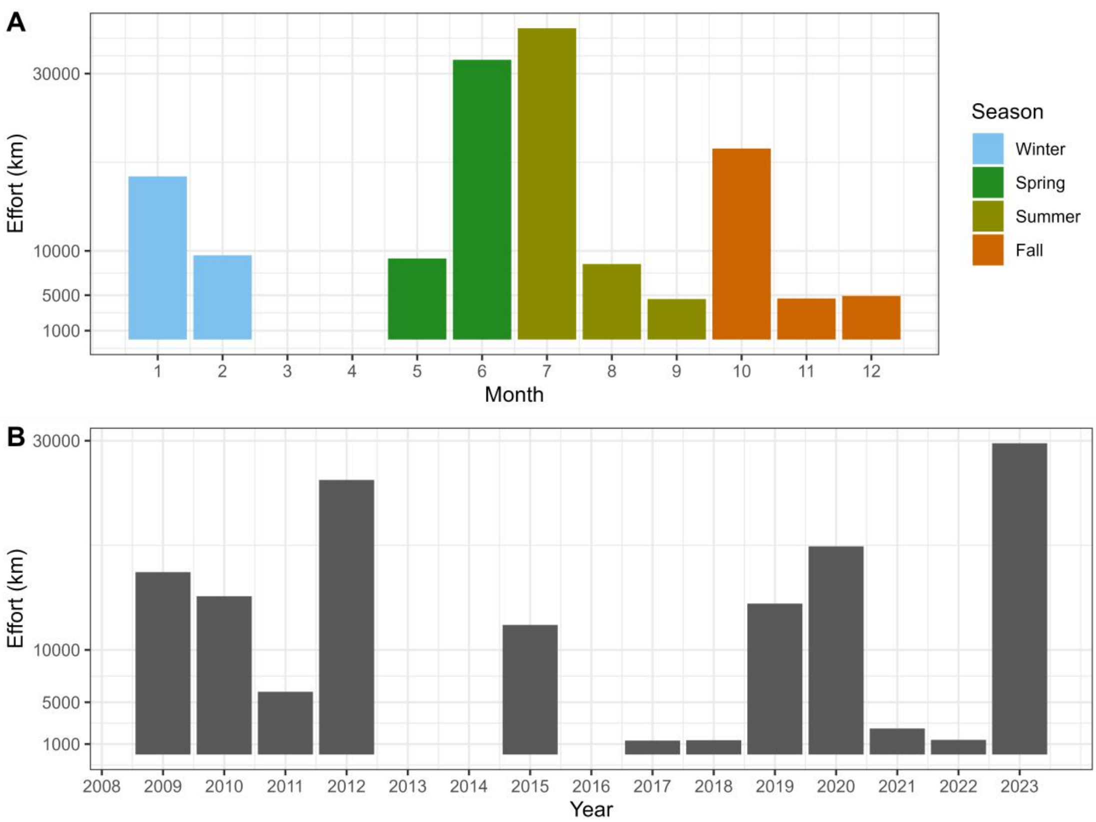
Environmental variables were used to prepare prediction grids covering the Pelagos Sanctuary area and the Western Mediterranean MFSD region. Grids of 10km² resolution were prepared to be consistent with prepared segments of effort that have been cut every 10km. Environmental variables were also associated with the centroid of each segment of effort to perform the gap analysis and the density surface model.

4/ Gap Analysis

Cetacean densities are going to be predicted using the density surface model from the marine environmental variables characterizing the Pelagos Sanctuary area and the western Mediterranean region. However, depending on the temporal resolution we chose to predict these densities, some of the predictions will occur in areas poorly informed by available data. To inform decision makers and stakeholders about the limits and uncertainty of the upcoming results, a gap analysis was performed. This gap analysis highlights areas where predictions of cetacean densities will be highly uncertain because informed by few data.

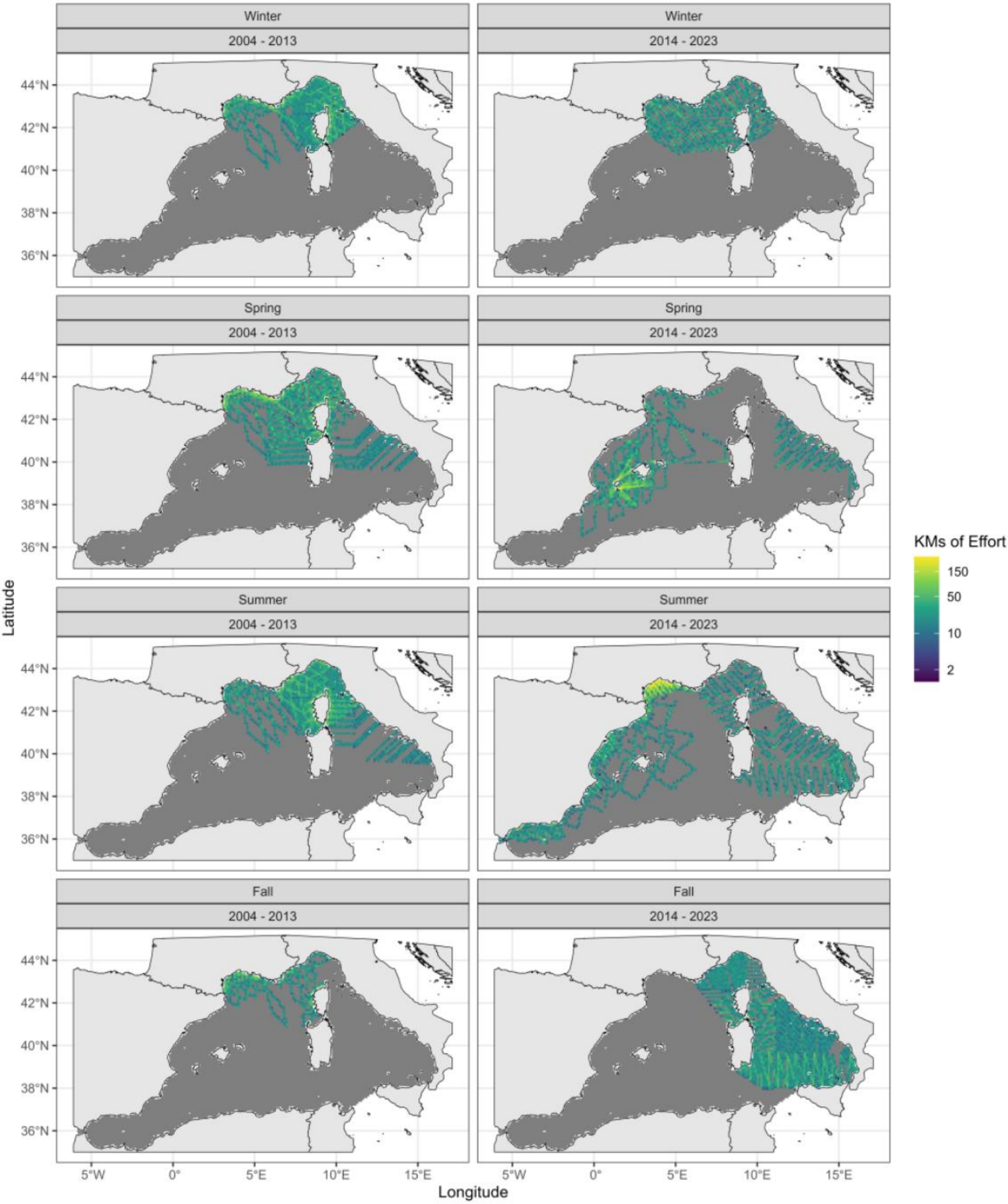
A/ Descriptive information

Originally in the consulting call, density maps were required for three periods: before 2004, from 2004 to 2013 and from 2014 to 2023. The data collected so far showed that we do not have any available data before 2009 (Figure 11). Thus, the available data make it impossible to get predictions of densities in the first period. A first summary of the collated effort data also revealed an unbalanced coverage of seasons and years. Most of the data were collected in June, July and October. The years 2023 and 2012 are more represented than other years. The years 2009, 2010, 2015, 2019 and 2021 total also a fair amount of effort data.



<u>Figure 11</u>: Number of kilometers of effort gathered per month (A) and per year (B), all surveys included.

In figure 12, the number of available kilometers of effort is shown according to the season and the period of data collection on a logarithmic scale. This figure reveals very little data for purple, blue to dark green colours. Grey areas were not covered by any survey. From this raw analysis, we can see that the Pelagos sanctuary was very well covered mainly in the winter and summer of the first period, namely by the PELAGOS survey. It also reveals that in fall of the first period and spring of the second period, the Pelagos sanctuary was poorly covered. However, conclusions are less obvious for other seasons.



<u>Figure 12</u> : Spatial coverage of the number of kilometers of effort gathered per period and season.

<u>B/ Analysis</u>

To be able to understand how all data from the Western Mediterranean region can inform prediction in the Pelagos Sanctuary, a gap analysis was performed. The gap analysis assumes that the Western Mediterranean region is a complex ecosystem that can be characterized by marine environmental variables (Tew Kai et al. 2020), the same ones as we described in the previous section. These environmental variables define an environmental space that shares similar or variables sea floor topography, current, surface temperature or primary production values in each season and period. This environmental space can thus be studied with geometric tools including distance and hulls to realize a gap analysis. It will reveal spatial areas where predictions will be extrapolated (i.e. not informed by data but extrapolated from the model only) and areas where predictions will be interpolated from data (i.e. well informed by data) (Authier et al. 2017; Bouchet et al. 2019).

The first step of this gap analysis is thus to study how the available data covered this environmental space. An environmental space must be seen as a space of N-dimension where each dimension is a marine environmental variable. If we consider only two environmental variables as an example, a simple polygon including all data points can be created from the observed values of both environmental variables at the centroid of each effort segment. The second step is to build the environmental space of the predicted area. Then, this analysis will reveal the gaps between the predicted and data-based environmental spaces and will allow us to map these gaps as extrapolations on the predicted maps.

A first result of this gap analysis informs about the extrapolated or interpolated value of each prediction point. The prediction points are the centroids of the cells including in the predicted maps (the different cells are visible on figure 13). I analysed if each prediction point was included or not within the data-based N-dimension environmental space. If a predicted point is included within this data-based environmental space, the associated prediction is considered as an interpolation while if the point is outside of the data-based environmental space, it is considered as an extrapolation (Figure 13).

In a second analysis, we can also map the different prediction points within the N-dimensional space and relatively to the data-based environmental space and analyse how each prediction point is well surrounded or not by data points. The second measure of this gap analysis derives the percentage of data points that are nearby each prediction point to inform it in the environmental space. This «nearby analysis» informs about the percentage of data that are used to make a prediction (Figure 14).

C/ Results

The extrapolation analysis (Figure 13) shows that separating the data in two periods will result in a high proportion of areas extrapolated. The predictions will be more robust if all years are pooled together. In this case only predictions in spring will mainly be extrapolations (i.e. fully extrapolated by the model and not directly informed by the data). Summer and fall predictions will mainly be interpolations (i.e. well informed by the data) while winter predictions will be a combination of both inter- and extra-polations. Note that the results of the gap analysis do not directly match raw Figure 12 because dynamic covariables were included at a monthly resolution on the effort data while they were included at a seasonal resolution in the prediction maps. Thus environmental spaces of dynamics variables between monthly and seasonal average may vary.

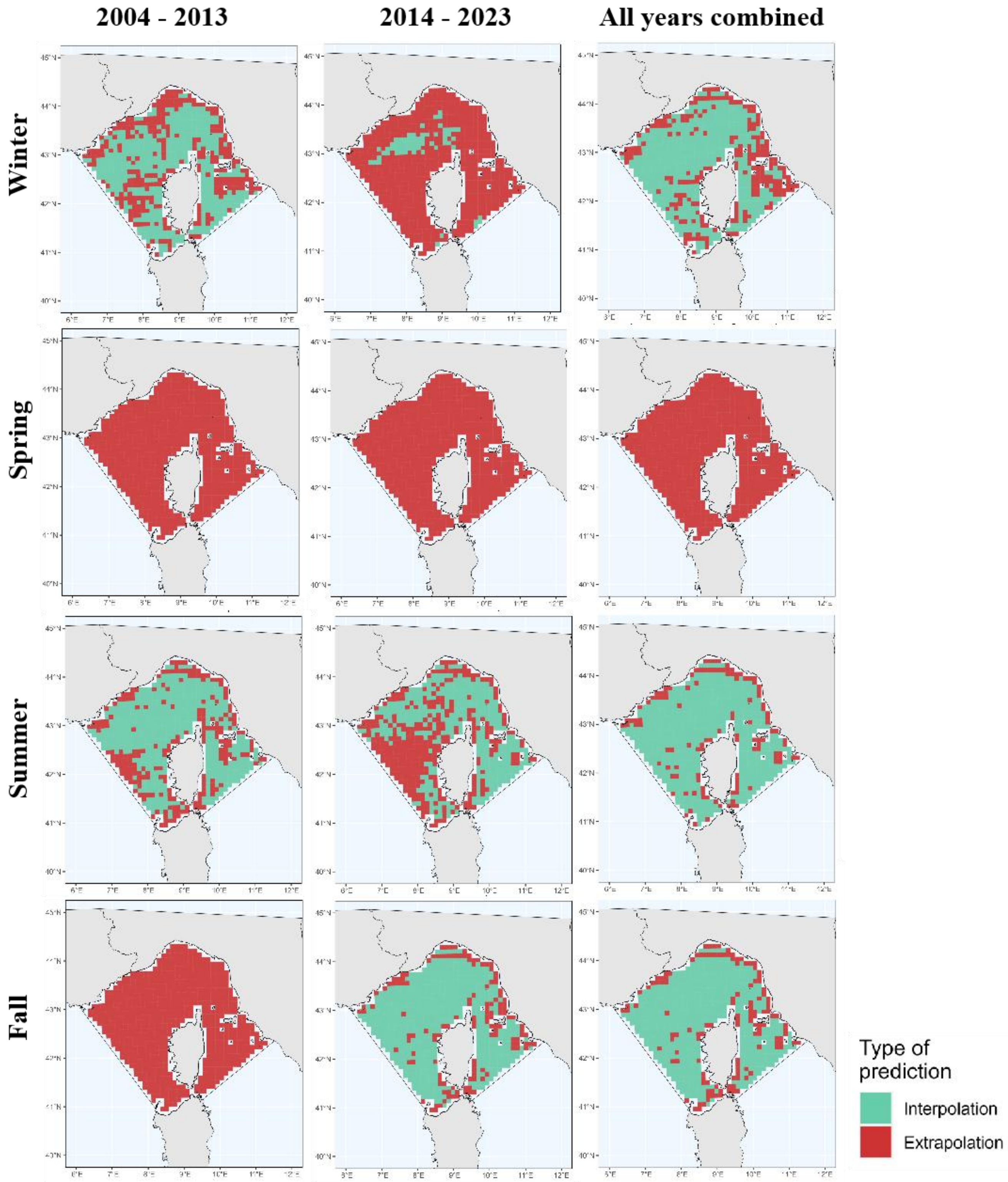


Figure 13: Gap Analysis: Interpolated and Extrapolated predictions per season and period. Systematic red points shared by all maps may reveal missing environmental variables that will be further checked in next analyses.

The nearby analysis (Figure 14) reveals the same main patterns as the extrapolation analysis with more details. It shows that if we combine all years, winter predictions in the Western part of the Pelagos Sanctuary would be particularly well informed by data. In fall, the western part of the Pelagos Sanctuary would also be better informed by data than the eastern part, while the opposite would be true for summer predictions.

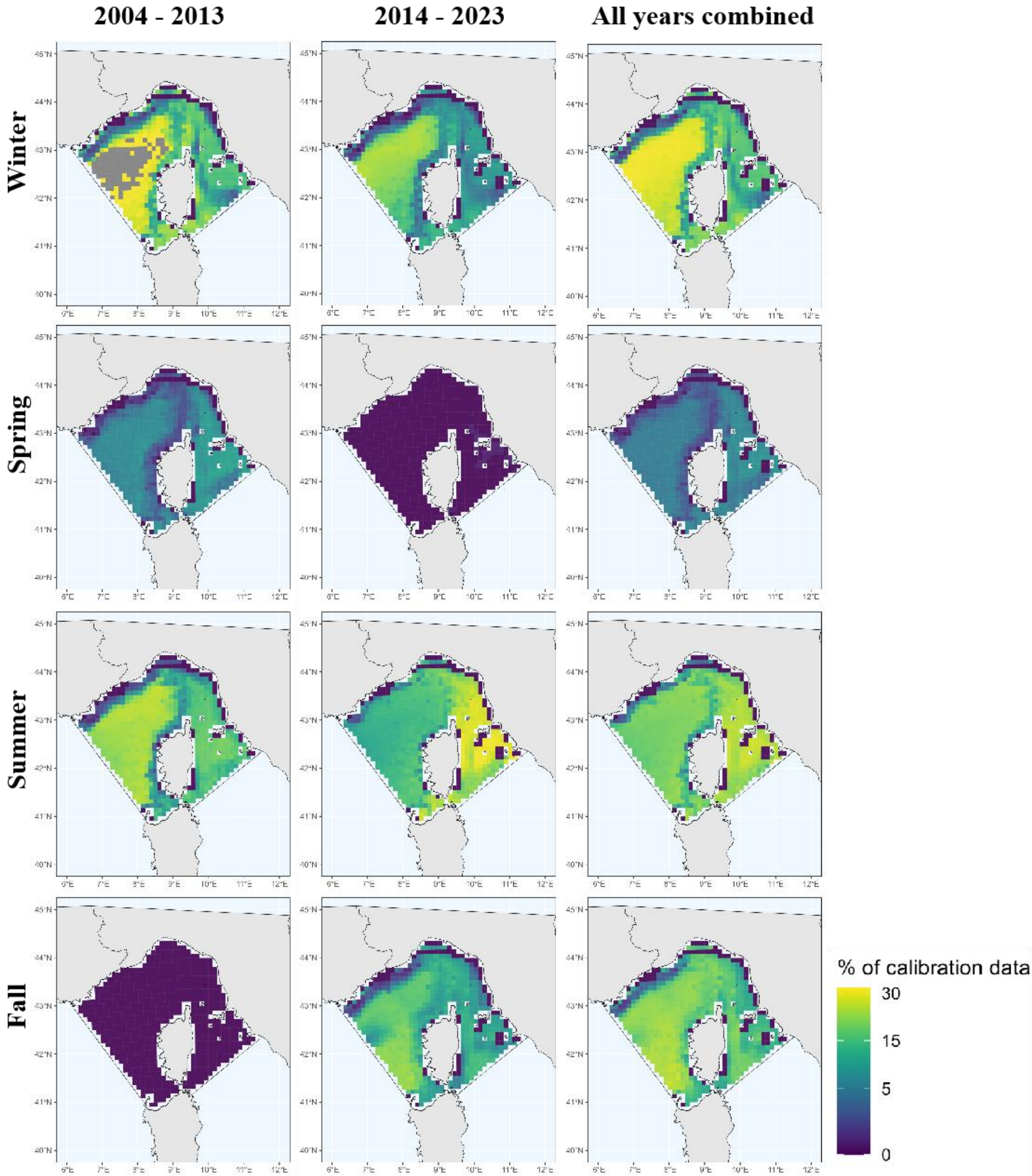


Figure 14: Gap Analysis: Percentage of data informing the prediction in each cell of the prediction grid per season and period. Systematic purple points shared by all maps reveal missing environmental variables that will be further checked in next analyses.

5/ Method to estimate cetacean densities along effort data

Some animals remain undetected along transects. The probability of detection often varies with species and environmental conditions including the survey platform of observation, the meteorological conditions and the marine conditions during the survey. The probability to detect an individual or group of individuals often decreases with increasing distance to the transect line (Miller et al. 2017). Distance sampling models are commonly used to estimate the probability of detection for a given species in a given survey (Buckland et al. 2001, Buckland et al. 2015). The declining detection probability is modelled by a decreasing function, often a half-normal or a hazard rate function (Buckland et al. 2015). This function is fitted to the observed number of detections according to observed perpendicular distances to the transect line (Figure 15). Distance sampling models aim to estimate the area actually covered by the survey. To do this, they estimate the effective strip half width (ESW), the distance at which as many animals are seen beyond it as are missed up to it, to deduce the effective area sampled during the line transect survey (Buckland et al. 2001). This value accounts for the missing individuals and is used to derive corrected animal densities.

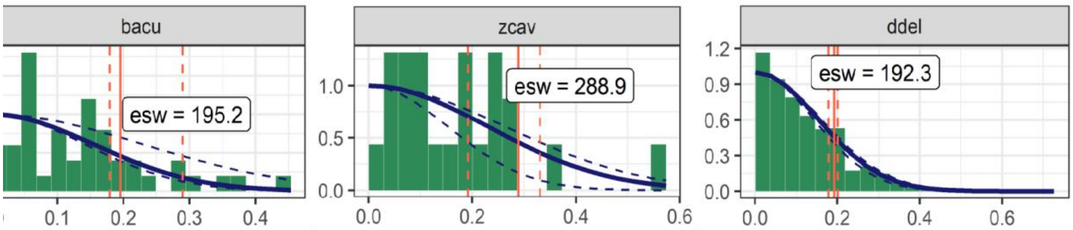


Figure 15: Fit of detection functions (dark blue) on histograms of the number of sightings per km to the transect line on three different species: the minke whale (bacu), the Cuvier's beaked whale (zcav) and the common dolphin (ddel). This figure was adapted from Plard et al. 2024 and the data were collated during the Cetambicion project. Red lines show the estimated ESW. Dotted lines show uncertainty (95% credible intervals) around the mean (plain line) for ESWs and for detection functions (blue).

This analysis gives robust results for common species with many detections but results are more uncertain for rare species (Buckland et al. 2015, Miller et al. 2017, Figure 15). To increase the accuracy and precision of estimated ESW, I will use a new methodology (Plard et al. 2024) based on pooling information from multiple species and surveys. Pooling observations with expected similar detection functions is a statistical technique to increase precision (Marques et al. 2007). However when pooling multiple survey and species, one must also account for the heterogeneity in detection probabilities among survey and species (see pooling robustness, Buckland et al. 2015, Rexstad et al. 2023) as pooling species or survey with heterogeneous detection functions would result in higher bias in ESW. Sighting data from multiple surveys and species shared common information about the influence of the increasing distance from the transect line and the environmental conditions on the probability to detect individuals. In this new methodology, this information will be used to increase precision in ESW using fusion effects (Malsiner-Walli et al. 2018, Plard et al. 2024). Fusion effects are state of the art statistical methods that allow the clustering of homogeneous categories of one variable automatically (Malsiner-Walli et al. 2018, Miller and Harrison 2018, Hu et al. 2022). Implemented in distance sampling models, this new method allows grouping surveys and/or species with homogeneous detection probabilities automatically while keeping apart heterogeneous ones. This methodology has been tested using simulation analyses and showed that in all cases, results using fusion effects were as or more precise and accurate than common distance sampling models (Plard et al. 2024). An example of the results obtained using fusion effects in distance sampling models vs. common distance sampling models is presented in Figure 16.

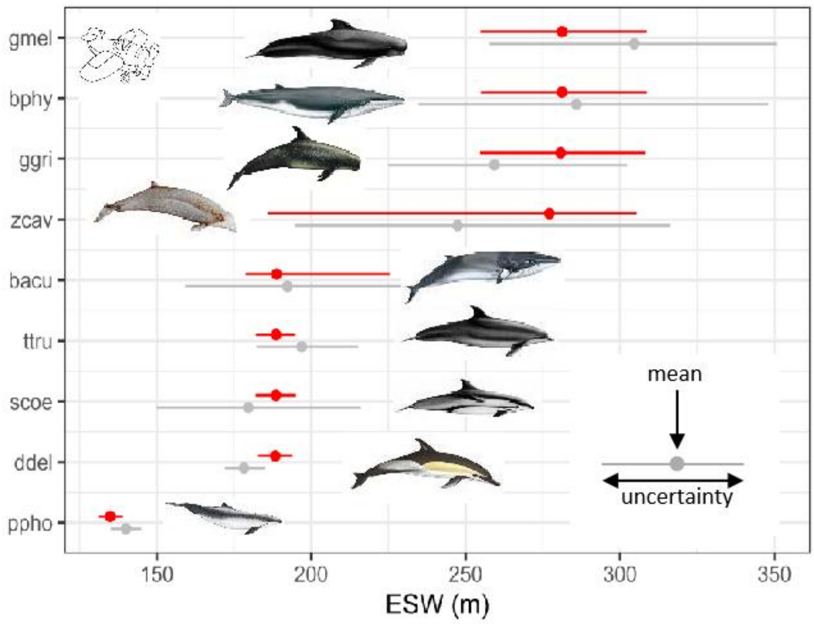


Figure 16: Example of ESW of 9 species estimated from a distance sampling model using fusion effect (red) and a common distance sampling model (grey). Adapted from Plard et al. 2024.

In the current analysis, I will analyse separately ship and plane surveys to account for the heterogeneous detection probabilities in aerial and ship surveys. Because this call aims to predict the densities of rare species (sperm whale, Cuvier's beaked whale and fin whale), I will include all data collected in Beaufort lower than 6 and I will add Beaufort as an explanatory variable in the model to account for variable probability detection in variable marine conditions. I will also include the subjective conditions variable as a continuous variable from 1 to 4 to account for the effect of bad environmental conditions on detection probabilities. Finally I will include all cetacean species included in the collated prepared dataset in distance sampling models and I will use fusion effects on species and survey to make the most of available data from all species and improve accuracy and precision of estimated ESW of rare species.

6/ Method to estimate abundance and density maps

The results of the distance sampling model will be the densities of cetaceans on each segment of the effort data deduced from the number of detected animals and ESW. Because distance sampling models assume perfect detection on the transect line, we need to further correct the estimated animal densities by bias availability (Marsh & Sinclair 1989, Buckland et al. 2004, Barlow 2015) before using them to estimate abundance and density maps. The most common reason why cetacean individuals are not available to be detected is them being submerged when diving. This availability bias is particularly important in species such as sperm whale, fin whale and Cuvier's Beaked whale. Previous estimate of bias availability in the literature (Virgili et al. 2019, Sigournay et al. 2020, Okamura et al. 2012, Laran etal. 2017) and estimated in similar surveys using robust protocol (SCAN IV, Gilles et al. 2023) will be used to correct animal densities.

Finally, corrected animal densities will be used in density surface models (DSM, Elith, J. & Leathwick 2009, Buckland et al. 2015, Miller et al. 2013) to estimate the relationship between cetacean densities and environmental variables (listed in table 2). Because we expect complex relationships between environmental variables and cetacean densities (see environmental variables part), I will use additive generalized linear models (GAMs) as they are able to account for highly flexible relationships (Wood 2006). For each species, a set of models including a combination of selected environmental variables will be run. This combination will include a maximum of four covariates to avoid excessive complexity of models and difficulty in their interpretation (Mannocci et al., 2014), and this combination will exclude correlated covariates. The five best fitting model will be selected using the Leave-one-out Information Criterion (LOOIC) developed by Vehtari et al. (2017, 2020). To avoid losing any important effect and getting results reflecting a model choice only, the five best selected models will be averaged using a stacking method (Yao et al. 2018). This method combines and weights the predictions of different models to get averaged predictions. Densities of each species in the Pelagos sanctuary will be predicted using these five staged best models from the environmental covariates describing the Pelagos Sanctuary area. To reflect the uncertainty in predictions, coefficient of variation of animal densities will also be estimated.

7/ Conclusion

Most requested data collected using a distance sampling protocol have been gathered, but other will be added once received (surveys ICCAT and Pelagos 2009 in summer). The results of this intermediate report will thus evolve.

The current collated data reveal high variation in temporal coverage. These results recommend building for each species an average map pooling all years together. Doing that, the gap analysis shows that a robust map could be built for each season except spring. This season will remain poorly informed by the data. These results mainly apply to common species such as the bottlenose dolphin. However, the densities of rarer species including the sperm whale, and the Cuvier's beaked whale might be predicted with high uncertainty form the DSM due to the low number of sightings in the collated data. Depending on the results and their uncertainty, we can consider looking for and requesting additional specific dataset to increase the number of sightings of these rare species. These dataset might increase the number of observed detections but also increase the heterogeneity of data. In particular they can be very difficult to use if detection probability cannot be estimated from the protocol or if effort was not quantified.

From these analyses, the following deliverables should be produced per season (except spring) and for the following species: bottlenose dolphin, striped dolphin, long-finned pilot whale, Risso dolphin, sperm whale, fin whale and Cuvier's beaked whale:

- Maps of extrapolations/interpolations
- Abundance estimates in the Pelagos Sanctuary and the 95% credible interval
- Map of densities: average and coefficient of variation in the Pelagos Sanctuary

Similar deliverables will be produced at the scale of the Western Mediterranean region.

References

Astarloa, A., Glennie, R., Chust, G., García-Baron, I., Boyra, G., Martínez, U., Rubio, A. & Louzao, M. (2021) Niche segregation mechanisms in marine apex predators inhabiting dynamic environments. Diversity and Distributions 27, 799–815

Authier, M., Saraux, C., & Péron, C. (2017). Variable selection and accurate predictions in habitat modelling: A shrinkage approach. Ecography, 40, 549–560.

Barlow, J. (2015) Inferring trackline detection probabilities, g(0), for cetaceans from apparent densities in different survey conditions. Marine Mammal Science 31, 923–943.

Barlow, D., Bernard, K., Escobar-Flores, P., Palacios, D., Torres, L. (2020) Links in the trophic chain: Modeling functional relationships between in situ oceanography, krill, and blue whale distribution under different oceanographic regimes. Marine Ecology Progress Series, 642, 207–225.

Benoit-Bird, K.J., & Au, W.W.L. (2003). Prey dynamics affect foraging by a pelagic predator (Stenella longirostris) over a range of spatial and temporal scales. Behavioral Ecology and Sociobiology, 53, 364–373.

Bouchet, P.J.; Miller, D.L.; Roberts, J.J.; Mannocci, L.; Harris, C.M. & Thomas, L. (2019) From Here and Now to There and Then: Practical Recommendations for Extrapolating Cetacean Density Surface Models to Novel Conditions. University of Saint Andrews

Buckland, S.T.; Anderson, D.R.; Burnham, K.P. & Laake, J.L. (1993) Distance Sampling - Estimating Abundance of Biological Populations. Chapman & Hall, 1st Edition.

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

Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L. & Thomas, L. (2004) Advanced distance sampling: estimating abundance of biological populations. Oxford University Press Oxford.

Buckland, S.T., Rexstad, E.A., Marques, T.A. & Oedekoven, C.S. (2015) Distance sampling: methods and applications, vol. 431. Springer.

Croll, D.A., Tershy, B.R., Hewitt, R. P., Demer, D.A., Fiedler, P.C., Smith, S.E., Armstrong, W., Popp, J. M., Kiekhefer, T., Lopez, V.R., Urban, J., Gendron, D. (1998) An integrated approch to the foraging ecology of marine birds and mammals. Deep-Sea Research Part II: Topical Studies in Oceanography, 45, 1353–1371.

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

Field, S.A., Tyre, A.J. & Possingham, H.P. (2005) Optimizing allocation of monitoring effort under economic and observational constraints. The Journal of Wildlife Management 69, 473–482.

Frederiksen, M., Edwards, M., Richardson, A.J., Halliday, N.C., & Wanless, S. (2006). From plankton to top predators: Bottom-up control of a marine food web across four trophic levels. Journal of Animal Ecology, 75, 1259–1268.

Freeman, M.M. (2008) Challenges of assessing cetacean population recovery and conservation status. Endangered Species Research 6, 173–184

Forney K.A. (2000) Environmental Models of Cetacean Abundance: Reducing Uncertainty in Population Trends. Conservation Biology, 14, 1271–1286.

Gilles A, Authier M,Ramirez-Martinez NC, Araújo H,Blanchard A, Carlström J, ..., Hammond PS (2023) Estimates of cetacean abundance in European Atlantic waters in summer 2022 from the SCANS-IV aerial and shipboard surveys

Guisan, A., & Zimmermann, N.E. (2000). Predictive habitat distribution models in ecology. Ecological Modelling, 135, 147–186.

Hammond, P.S., Macleod, K., Berggren, P., Borchers, D.L., Burt, L., Cañadas, A., Desportes, G., Donovan, G.P., Gilles, A., Gillespie, D. et al. (2013) Cetacean abundance and distribution in European Atlantic shelf waters to inform conservation and management. Biological Conservation 164, 107–122.

Hammond, P.S.; Lacey, C.; Gilles, A.; Viquerat, S.; Börjesson, P.; Herr, H.; MacLeod, K.; Ridoux, V.;

Santos, M.B.; Scheidat, M.; Teilmann, J.; Vingada, J. & Øien, N. (2021a) Estimates of Cetacean Abundance in European Atlantic Waters in Summer 2016 from the SCANS-III Aerial and Shipboard Surveys. Sea Mammal Research Unit, University of Saint Andrews, UK.

Hammond, P.S., Francis, T.B., Heinemann, D., Long, K.J., Moore, J.E., Punt, A.E., Reeves, R.R., Sepúlveda, M., Sigurðsson, G.M., Siple, M.C. et al. (2021b) Estimating the abundance of marine mammal populations. Frontiers in Marine Science 8, 1316.

Hastie, G.D., Wilson, B., Wilson, L.J., Parsons, K.M., & Thompson, P.M. (2004). Functional mechanisms underlying cetacean distribution patterns: Hotspots for bottlenose dolphins are linked to foraging. Marine Biology, 144, 397–403

Hijmans, R. J., Bivand, R., Forner, K., Ooms, J., Pebesma, E., & Sumner, M. D. (2022). Package 'terra'. *Maintainer: Vienna, Austria*.

Hu, G., Yang, H.C., Xue, Y. & Dey, D.K. (2022) Zero-inflated poisson model with clustered regression coefficients: Application to heterogeneity learning of field goal attempts of professional basketball players. Canadian Journal of Statistics, *51*, 157-172

Jewell, R., Thomas, L., Harris, C.M., Kaschner, K., Wiff, R., Hammond, P.S. & Quick, N.J. (2012) Global analysis of cetacean line-transect surveys: detecting trends in cetacean density. Marine Ecology Progress Series 453, 227–240.

Laran, S., Authier, M., Blanck, A., Doremus, G., Falchetto, H., Monestiez, P., Pettex, E., Stephan, E., Van Canneyt, O. & Ridoux, V. (2017) Seasonal distribution and 42 960 abundance of cetaceans within French waters-Part II: The Bay of Biscay and the English Channel. Deep Sea Research Part II: Topical Studies in Oceanography 141, 31–40

Malsiner-Walli, G., Pauger, D. & Wagner, H. (2018) Effect fusion using model-based clustering. Statistical Modelling 18, 175–196.

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

Marques, T.A., Thomas, L., Fancy, S.G. & Buckland, S.T. (2007) Improving estimates of bird density using multiple-covariate distance sampling. The Auk 124, 1229–1243.

Marsh, H. & Sinclair, D.F. (1989) Correcting for visibility bias in strip transect aerial surveys of aquatic fauna. The Journal of Wildlife Management 53, 1017–1024

McPherson, J.M. & Myers, R.A. (2009) How to infer population trends in sparse data: examples with opportunistic sighting records for great white sharks. Diversity and Distributions 15, 880–890

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

Miller, D.L., Rexstad, E., Thomas, L., Marshall, L. & Laake, J.L. (2017) Distance sampling in r. BioRxiv p. 063891.

Miller, J.W. & Harrison, M.T. (2018) Mixture models with a prior on the number of components. Journal of the American Statistical Association 113, 340–356.

Okamura, H., Minamikawa, S., Skaug, H. J., & Kishiro, T. (2012). Abundance estimation of long-diving animals using line transect methods. Biometrics, 68, 504-513.

Pace, D., Tizzi, R. & Mussi, B. (2015) Cetaceans value and conservation in the Mediterranean Sea. Journal of Biodiversity & Endangered Species 2015, 1-24.

Plard, F., Araújo, H., Astarloa, A., Louzao, M., Saavedra, C., Bonales, J.A.V., ... & Authier, M. (2024). Using fusion effects to decrease uncertainty in distance sampling models when collating data from different surveys. Marine Mammal Science.

Redfern, J. V., Ferguson, M., Becker, E. A., Hyrenbach, K. D., Good, C., Barlow, J., Kaschner, K., … & Werner, F. (2006) Techniques for Cetacean–Habitat Modelling. Marine Ecology Progress Series, 310, 271-295

Rexstad, E., Buckland, S., Marshall, L. & Borchers, D. (2023) Pooling robustness in distance sampling: Avoiding bias when there is unmodelled heterogeneity. Ecology and Evolution 13, e9684

Sigourney, D. B., Chavez-Rosales, S., Conn, P. B., Garrison, L., Josephson, E., & Palka, D. (2020). Developing and assessing a density surface model in a Bayesian hierarchical framework with a focus on uncertainty: insights from simulations and an application to fin whales (Balaenoptera physalus). PeerJ, 8, e8226.

Tew-Kai E, Quilfen V, Cachera M, Boutet M (2020) Dynamic Coastal-Shelf Seascapes to Support Marine Policies Using Operational Coastal Oceanography: The French Example. Journal of Marine Science and Engineering, 8, 585.

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Statistics and computing, 27, 1413-1432.

Vehtari A, Gabry J, Magnusson M, Yao Y, Bürkner P, Paananen T, Gelman A (2020). "loo: Efficient leave-one-out cross-validation and WAIC for Bayesian models." In R package version 2.4.1, .

Virgili, A., Authier, M., Boisseau, O., Cañadas, A., Claridge, D., Cole, T., Corkeron, P., Dorémus, G., David, L., Di-Méglio, N., Dunn, C., Dunn, T. E., García-Barón, I., Laran, S., Lauriano, G., Lewis, M., Louzao, M., Mannocci, L., Martínez-Cedeira, J., … Ridoux, V. (2019). Combining multiple visual surveys to model the habitat of deep-diving cetaceans at the basin scale: Large-scale modelling of deep-diving cetacean habitats. Global Ecology and Biogeography, 28, 300–314.

Wade, P. R. (1998) Calculating Limits To the Total Allowable Human-Caused Mortality of Cetaceans and Pinnipeds. Marine Mammal Science, 14, 1-37

Waggitt, J.J., Evans, P.G., Andrade, J., Banks, A.N., Boisseau, O., Bolton, M., ... & Hiddink, J.G. (2020). Distribution maps of cetacean and seabird populations in the North-East Atlantic. Journal of Applied Ecology, 57, 253-269.

Wencheng L.M. (2024) Package Gradient-Based Recognition of Spatial Patterns in Environmental Data Version 1.6.0.

Wood, S. N. (2017). Generalized additive models: an introduction with R. chapman and hall/CRC.

Yao Y., Vehtari A., Simpson D. & Gelman A., 2018. Using Stacking to Average Bayesian Predictive Distributions (with Discussion) Bayesian Analysis, 13: 917-100.