PhD

3.º
CICLO

FCUP
2020

Distribution and habitat modelling for cetacean species in the eastern north Atlantic Ocean

Ana Mafalda Tomás Correia

U.PORTO
FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

Distribution and habitat modelling for cetacean species in the eastern north Atlantic Ocean

Ana Mafalda Tomás Correia

Tese de Doutoramento apresentada à
Faculdade de Ciências da Universidade do Porto
Biologia
2020

Distribution and habitat modelling for cetacean species in the eastern north Atlantic Ocean

Ana Mafalda Tomás Correia
Doutoramento em Biologia
Departamento de Biologia
2020

Orientador
Isabel Sousa-Pinto, Professora, CIIMAR and FCUP

Coorientador
Graham Pierce, Professor, CSIC, University of Aberdeen and UAveiro
Massimiliano Rosso, Researcher, CIMA Research Foundation

Ao meu macaquinho,
que um dia também cruzes mares e oceanos.

Vou-me embora, vou partir mas tenho esperança
de correr o mundo inteiro, quero ir
quero ver e conhecer rosa branca
e a vida do marinheiro sem dormir
E a vida do marinheiro branca flor
que anda lutando no mar com talento
adeus adeus minha mãe, meu amor
eu hei-de ir hei-de voltar com o tempo
- Vitorino (na voz do Cmdt Madeira)

Acknowledgements

Happiness only real when shared.
-Chris McCandless-

To people,

À minha orientadora, Isabel, obrigada pela confiança em mim, a aposta incondicional no meu trabalho e a rede de apoio que montou, para que eu não tivesse medo de arriscar! Um obrigada que se estende no tempo, desde 2011!

Ao meu ~~(co-)~~ orientador, Massi, obrigada por me conduzires, pela parceria, amizade, apoio e por criares novas portas e janelas, novos caminhos, desvios e atalhos! Obrigada por me apresentares o mar! Mais ainda por me acolheres na tua família e me dares a conhecer a Isa, e os vossos miúdos tão queridos, Júlia e Tiago, verdadeiras fontes de alegria, paz e amor!

To my ~~(co-)~~ supervisor, Graham, thank you for believing in that girl, so inexperienced and unaware. You saw only the will, and for that, I am grateful. Thank you for the time, the patience, the wisdom and priceless advice!

To all my dearest marine mammal observers: this is also your work! This thesis is yours and of each one of you. Together, we created and we were CETUS. Together, we proved that the sum of the parts is a million times better than all the parts alone. I am proud of each one of you, wonderful biologists, all spread around the world. Thank you all: Olga Azevedo, Zara Valquíria, Andreia Pereira, Pedro Fernandes, Cláudia Fernandes, Sofia Esteves, Joana Romero, Ayça Eleman, Anicee Lombal, Julien Freyer, Lara Meraz, Alberto Pérez, Bárbara Matos, Catarina Morgado, Juliana Moron, Nicolas Blanc, Laura González, Oleh Danduriants, Ágatha Gil, Amélie Gadbois, Catarina Fonseca, Patrícia Carvalho, Corina Peter, Julia Heiler, Jack Ball, Paula Carreño, Roosevel Gutierrez, Paula Silvar, Santiago Otero, Ryan Mathews, Valéria Aragão, Francisco Cardeal, Mariana Morais, Amber Coleman, Alexandra Pires, Rhiannon Nichol, Torcuato Mantas, Annalisa Sambolino, April Eassom, Sara Pastorino, Anja Badenas, Leonardo Bernisone, Raul Valente, Tanja Shwanck, Lídia Pedrera, Rosie Bowyer, Maria Rívas, Rebeca Velasco, António Nogueira, Ana Ayres, Mieke Weyn, Fadia Al Abbar, María Jiménez, Svenja Halfter, Verónica Belchior, Evangelia Mpinou, Diana Fonseca, Carmen Escobar, Tara Callahan, Esteban Iglesias, Philippa Wright and Victoria Hope.

À TRANSINSULAR, às tripulações, a todos… Todos vocês têm nome neste trabalho. Uma vida não chega para expressar a minha gratidão a esta empresa e a cada um de vós. Vocês foram (e são!) a minha casa, o meu lar, a minha família. O camarote foi tantas vezes o meu refúgio e vocês o meu desabafo. Deixo um agradecimento especial aos que acompanharam de perto: Bruno Passarinho (e respetiva Santa Sara e piratas!), Cmdt Labrincha, Cmdt Luís, Cmdt Hélio, Cmdt João Paulo, Márcio, Marco, Alcindo, Helder, Chefe Júlio, Chefe Paiva, Chefe Rico.

Não posso deixar de fora um agradecimento muito muito especial ao meu primeiro comandante, aquele que acreditou e viu a força e potencial do CETUS (quando ainda nem se chamava CETUS!), aquele que tantas vezes me ligou para saber se "os dados estavam bons", aquele que viu baleias e golfinhos e que, quando eles não aparececiam, dizia "Don't worry, be happy", aquele que me deu a conhecer "a vida de marinheiro sem dormir" e que me falou dos cabos, das vigias, das cartas, das luzes, dos relógios… aquele que me desejou todo o sucesso com a maior das genuidades. Ao Cmdt Madeira, obrigada! Obrigada por tudo!

Also, a big thank you to all researchers that became friends, some even before this PhD started, and somehow played a major role in this work: Ana, Filhipe das pilhoto, Rita, Rui Caldeira, Paola Tepsich, Aurélie. Thank you also to their families that often opened their homes to me.

Ao meu macaquinho, que chegou já quase neste último ano de doutoramento. Tomás, obrigada por seres a maior felicidade da minha vida! Claro, obrigada por dormires noites completas, por comeres tão bem e me deixares terminar esta etapa… Mas obrigada também por tornares a minha (nossa) vida num caos tão bem-vindo, lindo e alegre! Num mundo tão caótico mas tão "país das maravilhas"! Obrigada por tudo que me trouxeste, mostraste e ensinaste. Há quem diga que me tornaste mais calma e centrada, que me equilibraste! Eu acredito e acredito ainda que foste o melhor que podia ter acontecido agora (e sempre)!

Ao meu porto de abrigo, o Marcos. O Marcos está comigo desde 2007, o Marcos é pai comigo, mas o Marcos é também o meu melhor amigo e meu melhor colega de trabalho. O Marcos conhece-me como ninguém e, para minha surpresa, gosta de mim… O Marcos faz tudo por mim e apoia-me incondicionalmente ainda que eu nem sempre agradeça. E o Marcos quer que eu trabalhe e senta-se ao meu lado a ajudar-me, mesmo não sendo biólogo! Então, o Marcos formata o meu computador, ajuda-me com uma base de dados, escreve um paper comigo, é meu co-autor, é condutor designado do CETUS e quando não é, empresta o carro que às vezes chega riscado, mas nunca faz mal! O Marcos recebe voluntários do CETUS em nossa casa por vários dias e, genuinamente, não se importa! O Marcos fica com o Tomás quando eu faço directas a trabalhar e é o melhor pai do mundo. Mas o Marcos também me arrasta para descansar, para férias, cinema e jantares fora, para experiências novas e alucinantes! O Marcos ainda não sabe, mas sem ele, eu era uma grande barata tonta! Por isso, Marcos, tens toda a minha (eterna) gratidão (volta aqui quando eu me esquecer da palavra obrigada)!

Aos meus pais, Nida e Tó, os melhores de sempre. Sempre aqui, e em todo o lado! De sorriso rasgado ou sobrolho franzido com conselho na ponta da língua, mas sempre de braços abertos. Sempre prontos, sempre disponíveis, sempre alerta e preocupados. Obrigada por sempre me deixarem ir, sem nunca me deixarem sem chão! A vocês devo tudo!

Aos meus avós, Laurinda e Domingos, símbolos de amor e força, incríveis, lamechas e lindos; aos meus avós-estrela, em especial, a Zelinha que é minha inspiração; aos meus primos e tios loucos; às minhas primas-irmãs Susana e Cati; e à minha família mais recente, o Mindo, a Ana e a irmã-amiga Leninha (aí, força da natureza!). Aos mais pequenos, agradeço a energia contagiante: Mafaldinha, Tomás, Kika, Bernardo, Tomé, Dri, Elen, Rafa e Santiago. Sei que tenho a família mais linda (e louca!) do mundo e isso deu-me toda a segurança que precisava para poder desaparecer de tempos a tempos sabendo que lá estariam quando eu voltasse (e garantiam que o meu lar estaria no sítio)!

Às maravilhas da minha vida, as minhas afilhadas lindas, agradeço os risos, a alegria, a felicidade e ansiedade da vossa chegada, do vosso regresso, a montanha-russa de emoções que me dá tanta energia para começar mais uma semana de trabalho. Obrigada por me fazerem querer viajar mar fora para ter histórias, mas também regressar para vos contar! Obrigada minhas queridas Mafaldinha (e tu que apareces aqui 2x?), Benedita, e os meus mais recentes amores, Marias Clara e Carolina.

Aos amigos-família, irmãos do coração, obrigaram por serem e estarem! Os de infância e os que apareceram depois sem eu contar. Os que vieram e ficaram. Aos Bilarrealenses, as minhas Girls e Boys de sempre: Ana, Sara, Catarina, Inês, Maria, Mingas, Bap, Lobo, Óscar, David. Também à minha Matilde tripeira e louca! Aos principes rebentos, porque saber deles me trouxe tanta alegria numa fase tão difícil deste doutoramento: obrigada Bea, obrigada Sebastião, obrigada Martim e Miguel, obrigada Manel, obrigada Bernardo! Às algas e T-Rexs da minha vida: tias Coelhinho, Lu e Ju, obrigada pela dança, pelas loucuras, e por me fazerem rir até doerem as bochecas (e comer até a calça desapertar!). Aos da Confederação, em especial Rosarinho e Miguel, obrigada pelo palco, por me fazerem sentir uma pessoa completa com tanta vida para além do trabalho! À minha Laurinha espanhola e açoreana que passou de voluntária a colega e depois a grande grande amiga! Sempre pronta para uma ajuda, um apoio, uma palavra encorajadora! Finalmente, à minha irmã que ao sê-lo só desde 2015, é-o tanto! Obrigada Ágatha! Tu sim, elevaste o CETUS e elevaste este trabalho! E a mim! Oh, tanto a mim! Tu sim, és tão autora desta tese! Obrigada pelos desenhos magníficos que deram cor a este trabalho! Obrigada mana chata!

Ao Raul, colega e amigo brilhante, pronto para apoiar em qualquer tarefa, disponível para o melhor e também para o mais chato. Obrigada pelo apoio e pela ajuda constante!

Ao LBC, o meu trabalho-casa sempre turbulento, confuso e tão bom! A todos colegas-amigos que me receberam e saíram, aos que estiveram desde sempre, e aos que vieram depois mas ficaram, obrigada pelo apoio! Em especial, obrigada à minha (Amaz)Sónia, Diogo, Raquel, Patrícia, Helena, Tânia, Débora, João, Isabel e Catarina.

Ao Honolulu, agradeço o cantinho sempre guardado para mim, o conforto do espaço e da comida, a alegria das meninas e as sopas do Tomás!

E obrigada às baleias e aos golfinhos! Por aparecerem e trazerem tanta alegria ao meu mundo!

To institutions,

I thank my host institutions, the Interdisciplinary Centre of Marine and Environmental Research (CIIMAR, Porto), the Instituto de Investigaciones Marinas - Consejo Superior de Investigaciones Científicas (IMM – CSIC, Vigo), the Faculty of Sciences of University of Porto (FCUP) and the TRANSINSULAR / ETE Group.

Funding

This study was conducted within a PhD program from the Faculty of Sciences of the University of Porto, Portugal, hosted by the Centre of Marine and Environmental Research (CIIMAR – Porto, Portugal) and funded by the Portuguese national funding agency for science, research and technology (FCT) under the grant SFRH/BD/100606/2014. CETUS Project is led by CIIMAR / University of Porto in partnership with the cargo ship company TRANSINSULAR / ETE Group.

Resumo

Os cetáceos assumem um papel muito relevante na conservação dos ecossistemas marinhos por se tratarem de espécies chave, guarda-chuva, indicadoras de biodiversidade, sentinelas e emblemáticas. No entanto, a conservação de cetáceos é limitada pelo conhecimento escasso da sua ocorrência e padrões de distribuição. As lacunas de informação são mais notórias em alto-mar, onde o esforço de investigação e amostragem é reduzido dada a logística complexa e dispendiosa. Face a este desafio, os investigadores têm vindo a utilizar plataformas de oportunidade (POs) para amostrar áreas mais remotas e menos acessíveis.

No Nordeste Atlântico (NEA), já foi registada uma grande diversidade de espécies de cetáceos. Contudo, registos da ocorrência destas espécies são maioritariamente restritos a áreas costeiras. Dados de ocorrência de cetáceos recolhidos entre 2012 e 2017 abordo de POs foram usados para estudar a distribuição e o habitat dos cetáceos no NEA, numa área delimitada pelas costas Ibérica e do noroeste Africano e pelos arquipélagos da Macaronésia. Uma grande diversidade de cetáceos foi registada em alto-mar. Houve oito espécies encontradas com maior frequência: *Delphinus delphis*, *Stenella frontalis*, *Stenella coeruleoalba*, *Tursiops truncatus*, *Ziphius cavirostris*, *Globicephala* sp., *Physeter macrocephalus* e *Balaenoptera acutorostrata*, sendo que *D. delphis* foi a mais frequentemente avistada. Modelos de nicho ecológico foram desenvolvidos tendo em conta fatores que afetam a deteção dos animais, variáveis espacio-temporais, processos oceanográficos e estruturas topográficas: *D. delphis* preferiu áreas a norte em águas costeiras e pouco profundas, principalmente na Península Ibérica e nos Açores; *Stenella* sp. apresentaram um nicho ecológico amplo, ocorrendo maioritariamente em águas oceânicas e nos arquipélagos; o *T. Tursiops* esteve associado às plataformas continentais; *Z. cavirostris* preferiu águas oceânicas a norte, estando associado às montanhas submarinas; *Globicephala* sp. e *P. macrocephalus* ocorreram mais a sul; e *B. acutorostrata* preferiu águas oceânicas a norte, em áreas próximas às montanhas submarinas. Campanhas dedicadas são necessárias em áreas onde se determinou uma maior probabilidade de habitats favoráveis à ocorrência das espécies. É essencial um maior esforço de investigação para entender a estrutura populacional e a conetividade entre as regiões, os movimentos das espécies e os efeitos dos impactos antropogénicos à escala da bacia oceânica.

Para ir de encontro às atuais prioridades na conservação marinha, estes dados deverão ser utilizados para planeamento espacial marinho e para a modelação da distribuição de cetáceos em cenários futuros de alterações climáticas.

Palavras-chave: cetáceos; nordeste Atlântico; distribuição; modelação de nicho ecológico; conservação.

Abstract

Cetacean species assume very important roles for the conservation of marine ecosystems as keystone, umbrella, biodiversity indicators, sentinel and flagship species. However, conservation of cetaceans is hindered by the lack of information on their occurrence and distribution patterns. The most important knowledge gaps are in high-seas where research effort is limited given the complex and expensive logistics of offshore surveys. To overcome this challenge, reseachers have been relying on observation platforms of opportunity (OPOs) to collect data in more remote, less accessible areas.

In the eastern North Atlantic (ENA), a great diversity of cetacean species has been recorded but occurrence records are mostly restricted to coastal areas. Data on cetacean occurrence collected between 2012 to 2017 abroad OPOs was used to study cetacean distribution and habitat within the ENA, in the area delimited by the coasts of the Iberian Peninsula and northwest Africa and the archipelagos of Macaronesia. A great biodiversity of cetacean species was reported in the high seas. Eight species were recorded more often: *Delphinus delphis*, *Stenella frontalis*, *Stenella coeruleoalba*, *Tursiops truncatus*, *Ziphius cavirostris*, *Globicephala* sp., *Physeter macrocephalus* and *Balaenoptera acutorostrata*, with *D. delphis* being the most frequently sighted. Ecological niche modelling was conducted taking into account detectability factors, spatiotemporal variables, oceanographic processes and topographic structures: *D. delphis* preferred areas located in the north and in coastal shallow waters, mostly in the Iberian Peninsula and the Azores archipelago; *Stenella* sp. presented a wide ecological niche, occurring mostly in oceanic waters and in the archipelagos; *T. Tursiops* was associated with continental platforms; *Z. cavirostris* preferred northern oceanic waters and was associated with seamounts; *Globicephala* sp. and *P. macrocephalus* occurred further south; and *B. acutorostrata* preferred northern oceanic waters in areas closer to the seamounts. Dedicated survey effort is needed in priority areas where suitable habitats were predicted with higher probability and further research is needed to better understand the population structure and connectivity between regions, species movements and effects of anthropogenic impacts at the basin-scale.

To address current priorities in marine conservation, the present dataset should be used for spatial marine planning and to model cetacean distribution under future climate change scenarios.

Keywords: cetaceans; eastern north Atlantic; distribution; ecological niche modelling; conservation.

Table of Contents

List of Figures

Chapter I

Chapter II

Chapter III

Chapter IV

Chapter V

Chapter VI

Chapter VII

List of Tables

Chapter I

Chapter II

Chapter III

Chapter IV

Chapter V

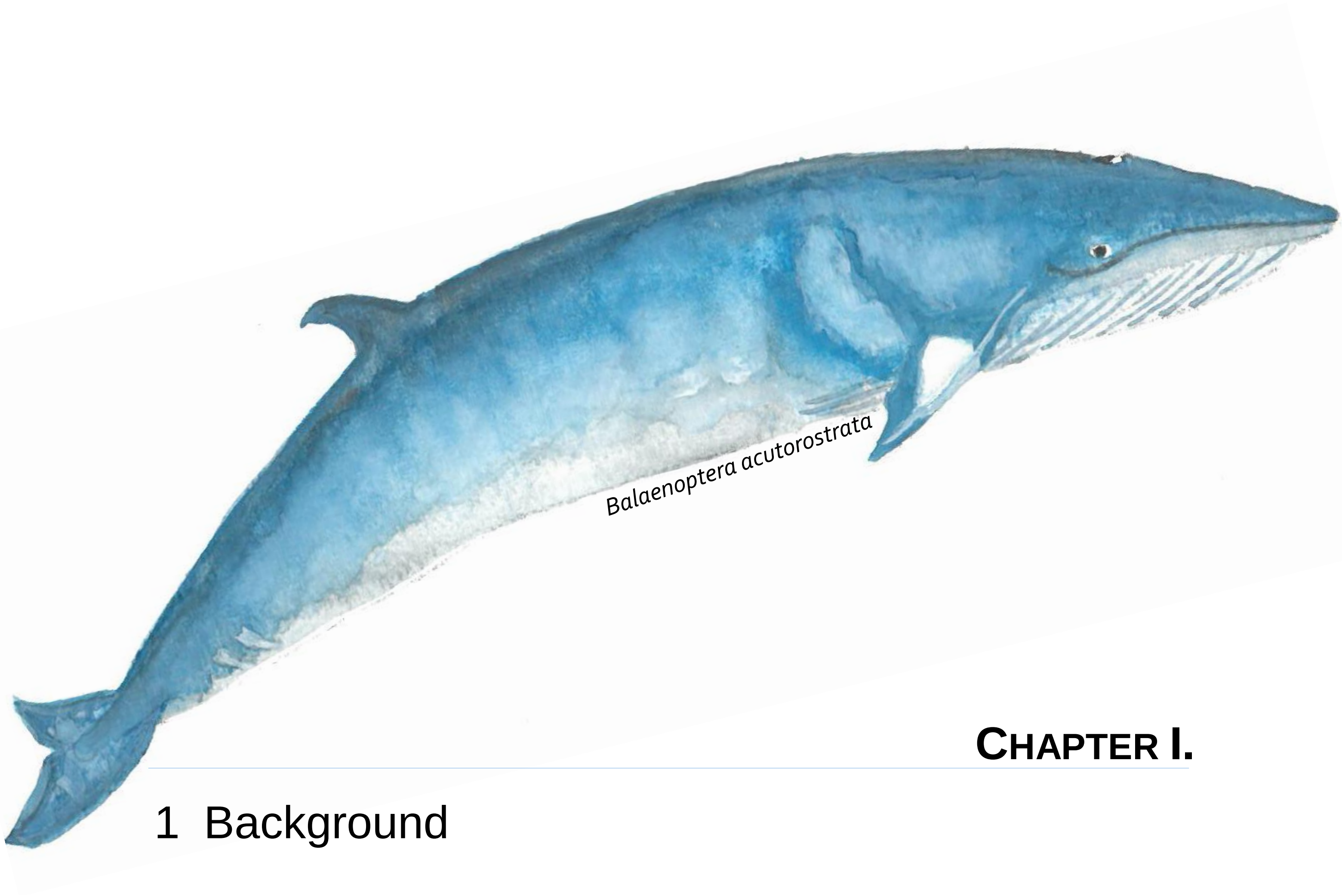
Chapter VI

Chapter VII

List of Abbreviations

º – degree
ºC – degree Celsius
% – percent
AC – Azores current
ACCOBAMS – Agreement on the Conservation of Cetaceans of the Black Sea, Mediterranean Sea and contiguous Atlantic area
AIC – Akaike Information Criterion
AIS – Automatic Identification System
ALME – African Large Marine Ecosystem
ASCOBANS – Agreement on the Conservation of Small Cetaceans of the Baltic, North East Atlantic, Irish and North Seas
AUC – Area Under the Curve
AV – available habitat
Az – Azores archipelago
BA – *Balaenoptera acutorostrata*
BD – bottlenose dolphin
BL – baleen whle
BK – beaked whale
BODC – British Oceanographic Data Centre
BRT – boosted regression trees
C – Caniçal
CBD – Convention on Biological Diversity
CC – Canary current
CCS – Canarian current system
CD – common dolphin
CHL / chl – chlorophyll-a
CI – Canary Islands / confidence interval (context dependant)
CIIMAR – Interdisciplinary Centre of Marine and Environmental Research
cm – centimeters
CMS – Convention on Migratory Species
CSV – Comma Separated Values
CTA – classification tree analysis
CV – Cape Verde
D – surveyed distance
DD – *Delphinus delphis*
dist_coast / dist.coast – distance to the coast
dist_sm / dist.sm – distance to seamounts
DT – delayed time
DwC – Darwin core
DwC-A – Darwin core archive
δ – change
EBSAS – Ecologically and Biologically Significant Marine Areas
edf – effective degrees of freedom
Ef – effort
EEZ – Exclusive Economic Zone
e.g. – for example
EMEPC – Estrutura de Missão para a Extensão da Plataforma Continental
EMODnet – European Marine Observation and Data Network
eMoF – ExtendedMeasurementOrFact
ENA – eastern north Atlantic
ENM – ecological niche model
EPSG – European Petroleum Survey Group
ER – encounter rate
EU – European Union
FAO – Food and Agricultural Organization
FCT – Portuguese Foundation for Science and Technology
GAM – generalized additive models
GBM – generalized boosted models
GCS – Geographic Coordinate System
GCV – generalised cross-validation
GEBCO – General Bathymetric Chart of the Oceans
GES – good environmental status
GLM – generalized linear models
GMT – Greenwich Mean Time
GPS – Global Positioning System
Gsp. – *Globicephala* sp.
h – height
H – Horta
ICES – International Council for the Exploration of the Sea
lon – longitude
LP – Las Palmas
m – meters
m/s – meters per second
max – maximum
Mad – Madeira Island
MARS – multivariate adaptive regression splines
Maxent – maximum entropy
MFS – most frequently sighted
MGET – Marine Geospatial Ecology Tools
mg/m^3 – milligrams per cubic meter
mgcv – Mixed GAM Computation Vehicle with Automatic Smoothness Estimation (R package)
min – minimum
mm – millimeter
MMO – marine mammal observer
MODIS – Moderate-Resolution Imaging Spectroradiometer
MOU – Memorandum of Understanding
MPA – Marine Protected Area
NRT – near-real time
MSFD – Marine Strategy Framework Directive
MSLA / msla – Mean Sea Level Anomalies
µm – micrometer
NASA – National Aeronautics and Space Administration
N – north
NE – northeast
NEA – nordeste Atlântico
NI – non-identified
nmi – nautical miles
No / N / n – number
ns – non-sighting
NTW – northeasterly trade winds
NW – northwest
NWA – northwest Africa
O – Oporto
OBIS – Ocean Biogeographic Information System
OBIS-ENV-DATA – Ocean Biogeographic Information System Environmental Data
OBIS - SEAMAP – Ocean Biogeographic Information System Spatial Ecological Analysis of Megavertebrate Populations
OPO – Observation Platform of Opportunity
OSPAR – Oslo and Paris Conventions
P – Palmeira
PC – Portuguese current
PD – Ponta Delgada
PEEZ – Portuguese Exclusive Economic Zone
PM – *Physeter macrocephalus*
PO – Plataforma de Oportunidade
PV – Praia da Vitória
resid – residuals
RF – random forests
ROC – Receiver Operating Characteristic
RSF – Resource Selection Functions
S – Sal
s – sighting / smooth (context dependant)
SA – surveyed area
SAC – Special Area of Conservation
SC – *Stenella coeruleoalba*
SCI – Site of Community Importance
SD / std / σ – standard deviation
SE / se – standard error
sec / s – second
SF – *Stenella frontalis*
sp – species
SST / sst – sea surface temperature
SW – sperm whale
Tot – total
TT – *Tursiops truncatus*
UBRE – Unbiased Risk Estimator
UNEP – United Nations Environment Programme
URI – Unique Resource Identifiers

ICNF – Instituto da Conservação da Natureza e das Florestas
i.e. – meaning
IFAN – Institut Fondamental d'Afrique Noir
IP – Iberian Peninsula
IPT – Integrated Publishing Toolkit
ISO – International Organization for Standardization
IUCN – International Union for Conservation of Nature
IW – international waters
km – kilometers
L – Lisbon
lat – latitude
UTM – Universal Transverse Mercator
VIF – Variance Inflation FactorUTM
vis – visibility
VLIZ – Vlaams Instituut voor de Zee
W – west
WGS84 – World Geodetic System 1984
WKT – well-known text
WoRMS – World Register of Marine Species
WWF – World Wildlife Fund
ZC – *Ziphius cavirostris*



CHAPTER I.

1 Background

1.1 State of the Art

1.1.1 Cetacean ecology – their role in the aquatic ecosystems

Cetacean species are distributed across the entire globe, being encountered from the poles to the equator and from rivers to high-seas. They perform latitudinal and longitudinal movements, with baleen whales undertaking long seasonal migrations from cold feeding grounds to warm breeding areas (Cawardine, 2000). But what is their ecological role in the aquatic ecosystems?

If we strictly consider the term "ecological role" as any functional significance of a species/taxon, then all cetaceans influence, somehow, the ecosystems they inhabit. But, to determine their role (or their importance), then we need to understand how the ecosystems would respond to changes in cetacean distribution or abundance (Bowen, 1997). Using empirical data, Pershing *et al.* (2010) proved that whaling, which has led to a decrease of 66% to 90% of the worldwide whale numbers, has altered the ocean's ability to store and sequester carbon.

There are four main ecological pathways by which cetaceans influence ecosystems (Roman *et al.*, 2014):

1. Consumers. The infraorder Cetacea includes both Odontocetes and Mysticetes, hence comprising major consumers at most trophic levels, from zooplankton to predatory fish (Bowen, 1997). Cetaceans exert a strong pressure on marine communities through predation and indirect food-web interactions (trophic cascades). On the other hand, bottom-up control is also relevant, with cetaceans depending on areas with high biodiversity value or great biomass (Sergio *et al.*, 2008). Given their size and biomass requirements, cetaceans are an important storage of energy and nutrients (Roman *et al.*, 2014).
2. Prey. Cetaceans are important predators in aquatic ecosystems but they can also be important prey (for example, for killer whales or sharks). The loss of such important sources of nutrients and energy leads to major changes in the function and structure of the ecosystem (Roman *et al.*, 2014).
3. Detritus. Cetacean carcasses are important sources of detritus in the ecosystems, with sinking carcasses providing energy-rich habitats to benthic communities (Smith, 2007). Also, stranded carcasses are sources of biomass in the sea-land interfaces (Roman *et al.*, 2014). These are important processes for the carbon sequestration (Pershing *et al.*, 2010).
4. Nutrient vectors. Transport of nutrients occurs both horizontally, through high to low latitudinal whale migrations, and vertically, through the "whale pump" mechanism.

Migratory movements from productive high-latitude feeding grounds to oligotrophic low-latitude breeding areas, leads to the transport of limiting nutrients through urine, carcasses and placentas (Roman *et al.*, 2014). On the other hand, vertical movements result in the ocean mixing of the water column, especially relevant in highly-stratified water columns; and the transport of limiting nutrients, nitrogen and iron, to the surface due to whales feeding at the bottom and defecating at surface. Hence, cetaceans play an important role in the transport and recycling of carbon and limiting nutrients in the ocean (Roman *et al.*, 2014).

Besides enhancing biodiversity and productivity, cetaceans provide other ecosystem services, for example, in climate regulation through carbon sequestration and "whale pump" and in tourism with activities such as whale-watching (Roman *et al.*, 2014). Finally, as top-predators, cetaceans are strategic species in marine conservation as they gather several important characteristics (Sergio *et al.*, 2006; 2008):

1. Keystone species. They have a relevant ecological role in the ecosystems.
2. Umbrella species. Their area requirements support populations of other less area-demanding species.
3. Biodiversity indicators. They usually select sites with complex habitats associated with high biodiversity value.
4. Sentinels. As long-lived animals with low densities and low fecundity, they are particularly susceptible to changes in the ecosystems, especially human-induced. They are also bioaccumulators of toxic pollutants.
5. Flagship species. They are charismatic species with a great fundraising potential.

1.1.2 The eastern north Atlantic – cetacean diversity and habitat complexity

The eastern north Atlantic, in the area delimited west by the Macaronesian archipelagos (Azores and Cape Verde) and east by the mainland coasts of the Iberian Peninsula and northwest Africa (Figure 1), holds a great diversity of cetaceans, both coastal and oceanic, resident and migratory species. In total, 17 species have been recorded along the continental Portuguese coast (Brito & Sousa, 2011; Moura *et al.*, 2017), 26 in Madeira (Alves *et al.*, 2018a; Freitas *et al.*, 2012), 28 in Azores (Silva *et al.*, 2014; Tobeña *et al.*, 2016), 28 in the Canary Islands (Carrillo *et al.*, 2010; Pérez-Vallazza *et al.*, 2008), 24 in Cape Verde (Hazevoet & Wenzel, 2000; Hazevoet *et al.*, 2010) and 36 in the northwestern African coast (Djiba *et al.*, 2015; Perrin & Van Waerebeek, 2012; Robineau & Vely, 1998; Weir & Pierce, 2013).

Such diversity is a result of the wide latitudinal and longitudinal range and habitat complexity within the area. Topographically, there are several seamounts, hills, canyons and abyssal plains (Figure 1). Seamount chains rise from over 5000 m depths to less than 100 m. Four oceanic archipelagos with narrow continental platforms and abrupt slopes contrast with mainland regions with large continental platforms extending further offshore. Oceanographically, there are strong upwelling systems in Iberian and African coasts with a marked seasonality. There are also smaller upwelling phenomena associated with the seamounts and oceanic islands. Moreover, six major currents influence the dynamism of the area: Portugal Current, Azores Current, Canary Current, Mauritania Current, North Equatorial Current and North Equatorial Counter-Current (Mason, 2009).

Without any physical barriers, cetaceans cross administrative boundaries and range over several Exclusive Economic Zones (EEZs). Alves *et al.* (2018b) documented movements of *Globicephala macrorhynchus* between Madeira, Azores and Canary Islands, proving the connectivity between Macaronesian archipelagos. Also, baleen whales cross the area while migrating from productive, cold water, high-latitude feeding grounds to oligotrophic, warm water, low-latitude breeding grounds (Valente *et al.*, 2019). Published knowledge on cetacean distribution though is mostly limited to coastal areas where most research effort has been undertaken. The lack of data in offshore waters hinders cetacean conservation and management in the area (Directive 2008/56/EC; Santos & Pierce, 2015).

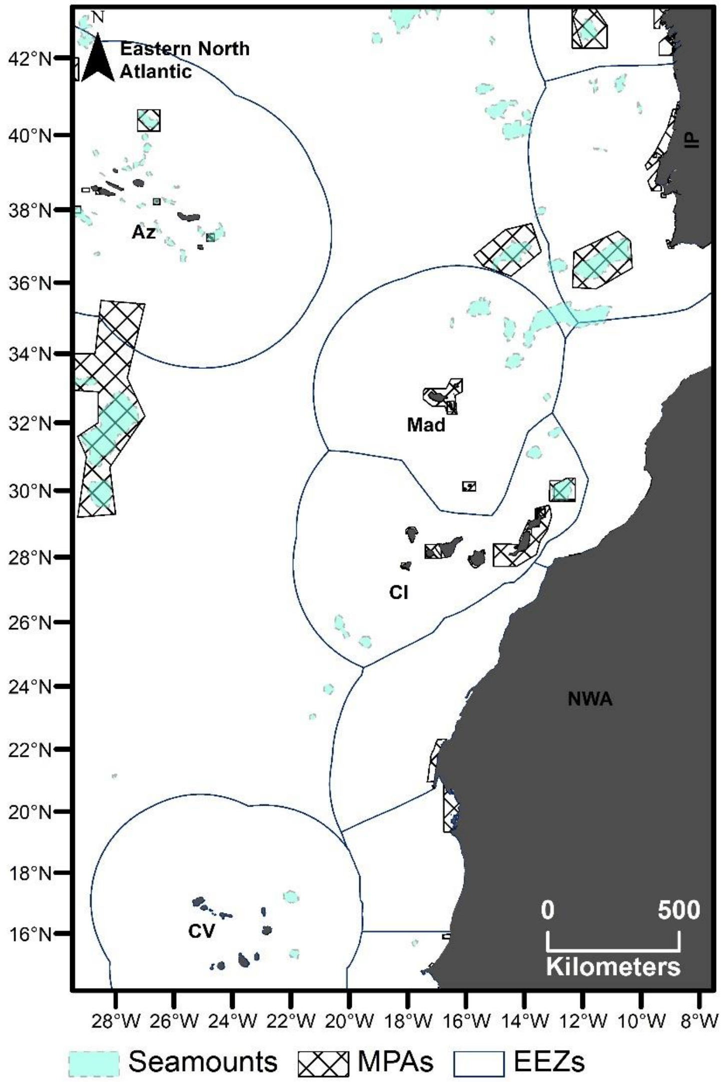


Figure 1. Map of the eastern north Atlantic, delimited west by the Macaronesian archipelagos (Azores and Cape Verde) and east by the mainland coasts of the Iberian Peninsula and northwest Africa, with the seamounts, Marine Protected Areas (MPAs) and Exclusive Economic Zones (EEZs). Az – Azores archipelago; IP – Iberian Peninsula; Mad – Madeira island; CI – Canary Islands; NWA – Northwest Africa; CV – Cape Verde archipelago. MPAs data collected from www.protectedplanet.net.

1.1.3 Data collection – filling gaps in data-poor areas

Lack of data in high-seas is not a specificity of the eastern north Atlantic (Correia *et al.*, 2015; Hammond *et al.*, 2013; Jungblut *et al.*, 2017). To collect data in oceanic waters is a major challenge, given the logistics and financial investment needed (Alves *et al.*, 2018a; Correia *et al.*, 2015; Kiszka *et al.*, 2007; Moura *et al.*, 2012; Tobeña *et al.*, 2016; Viddi *et al.*, 2010). Moreover, when the areas to be surveyed cross national boundaries or are beyond national jurisdiction, the need for international agreements complicates the implementation of monitoring programmes (García-Baron *et al.*, 2019).

Data deficiency is one of the main issues in cetacean conservation. Over 35% of cetacean species are data deficient (www.iucnredlist.org) and often “data deficient” is treated as “no concern" or leads to delayed management actions (Parsons *et al.*, 2016). In international waters, only 1.2% are protected areas, while within EEZs the percentage reaches the 16.8% (UNEP-WCMC *et al.*, 2018). Moreover, most of these protected areas do not have a defined management plan (at least, reported) possibly due to the lack of baseline data (http://natura2000.eea.europa.eu/; www.protectedplanet.net). It is then urgent to find solutions to survey data-poor areas.

In recent years, the use of Observation Platforms of Opportunity (OPOs) for research purposes has increased and researchers have worked on protocols to collect long-term standardized data (www.cetusproject.com; cetaceos.webs.ull.es/bioecomac/cetavist, www.monicet.net). Cetacean monitoring from OPOs consists on taking advantage of existing structures, which main aim is not research-related, to collect data on cetaceans. The platforms range from whale-watching boats (e.g., García *et al.*, 2018), touristic cruises or ferries (e.g., Tepsich *et al.*, 2014), fishing vessels (e.g., Silva *et al.*, 2014), oceanographic vessels (e.g., Campbell *et al.*, 2015), seismic surveys (e.g., Gray & Van Waerebeek, 2011) or cargo ships (e.g., Correia *et al.*, 2015, 2019). Though on the whale-watching the aim is to find cetaceans, it is for touristic purposes and not data collection, hence this is still considered an OPO.

When depending on OPOs, the survey design has to be adjusted to the platform characteristics and the company's agenda. However, the costs of monitoring are substantially lower and the survey coverage is increased, both spatially and temporally. As such, although the use of OPOs does not allow a "custom-made" survey design, it is a very cost-effective method for cetacean monitoring. The data can either be collected opportunistically (e.g., by the company's operators) or by dedicated researchers. While the costs are reduced with data collected opportunistically, there are some disadvantages: usually there is no information on survey effort, sampling is not performed according to a standard protocol and many occurrences may be missed. If data is collected by dedicated researchers, then survey effort can be recorded and sampling follows a standard protocol. This provides more reliable information and allows a greater number of analysis: for example, the assessment of relative abundances and the use of certain ecological niche modelling techniques that require presence/(pseudo)absence data (see section below) (Correia *et al.*, 2015, 2019; Kiszka *et al.*, 2007; Moura *et al.*, 2012).

Even with the possibility to use OPOs to cross large areas, a full coverage of the global ocean is virtually impossible. The vast majority of marine ecosystems is still lacking baseline data or is insufficiently surveyed to permit a good level of understanding of species diversity and distribution patterns. Identifying priority areas with knowledge gaps, where monitoring efforts are required, is essential (Kaschner *et al.*, 2012; Mannocci *et al.*, 2018). So, how can we prioritize areas to survey? How can we select areas that might hold higher biodiversity or suitable habitats for certain species? Solution lies on analytical methods such as the ecological niche modelling techniques.

1.1.4 Ecological niche modelling – concepts, techniques, applications

Modelling techniques have been widely used to forecast cetacean distribution and understand its relationship with habitat characteristics (Table 1). Essentially, ecological niche models predict different types of the realised niche of a species (Sillero, 2011). The realised niche is the portion of the fundamental niche constrained by species competition, being the fundamental niche the “n-dimensional volume in the environmental space where a species can maintain a viable population and persist along time” (Hutchinson, 1957; Pearson, 2007). The predictions can be closer to the potential niche, i.e. “the intersection between the fundamental niche space and the available environmental space” (as defined by Jackson & Overpeck, 2000), or to the occupied niche, i.e. the portion of the realised niche constrained also by geographical and historical factors (as defined by Pearson, 2007). The main result of the models is the identification of suitable habitats, potentially leading to the production of habitat suitability maps (Sillero, 2011). Sillero (2011) provides an analysis on the classification of ecological niche models and concepts of the different types of niches (fundamental, potential, realised, occupied).

Modelling techniques range from environmental models, classification techniques (classification tree analysis (CTA)), regression-based models (Generalized linear models (GLM), Generalized additive models (GAM), Generalized boosted models (GBM), multivariate adaptive regression splines (MARS)) and machine learning (Maximum entropy (Maxent), boosted regression trees (BRT), random forests (RF)) (Derville *et al.*, 2017; Elith & Leathwick, 2009). Derville *et al.* (2017) compared the performance of several modelling techniques using an endangered population of humpback whales (*Megaptera novaeangliae*) as a case study. GAMs and Maxent models provided the best complexity trade-off, accurate predictions and ecological insight. While regression techniques require presence/(pseudo)absence data, Maxent models can be used with presence-only data, usually undertaking a presence/background approach.

When working with data obtained from cetacean monitoring, one can never have true absences as the occurrences may be missed. With visual data, false absences happen when cetaceans pass by the observer while underwater or visibility conditions preclude detection. With acoustic monitoring, detection is affected by distance to the source, weather, background noise or absence of vocalizations (Sousa-Lima *et al.*, 2013). The alternatives are to use presence/pseudo-absence or presence-only data. Presence-only techniques, although allowing the use of opportunistically collected data with a good predictive performance (do Amaral *et al.*, 2015; Friedlaender *et al.*, 2011; Thorne *et al.*,2012), are complex and sometimes a "black-box", limiting their application in studies focusing on ecological insights (Elith & Leathwick, 2009). Models using effort-based data can provide results for better ecological interpretation of the species habitat characteristics (Brotons *et al.*, 2004; Praca *et al.*, 2009; Tepsich *et al.*, 2014). Survey effort information allows the definition of pseudo-absences within the sampled environmental space, accounting for the effort.

Besides the modelling technique and approach, it is very important to decide on the most appropriate variables and spatio-temporal scales to include in the models. It is desirable that selected variables have a relevant ecological justification to be included. However, other variables such as spatial and temporal factors may be considered to account for biogeography and seasonality (Elith & Leathwick, 2009). The spatio-temporal resolution of the predictors determines the extent to which results should be interpreted and also which processes are being modelled (large or small scale processes). Multi-scale models have been suggested for a better understanding of the cetacean-habitat relationship (Fernandez *et al.*, 2018; García *et al.*, 2018). Overall, the best model approach and methodology must be selected given the data available, sampled area and the aims of the study (Redfern *et al.*, 2006). Table 1 summarizes recent publications on ecological niche modelling of cetacean species.

Table 1. Recent studies (2015-2019) on ecological niche modelling of cetaceans – summary of the modelling approaches used.

Source	Modelling technique	Predictors	Main application
do Amaral *et al.*, 2015	Maxent	Topographic, oceanographic	Estimate potential distribution
Arcangeli *et al.*, 2016	GAM	Topographic	Compare presence, distribution and habitat use between two periods
Barragán-Barrera *et al.*, 2019	Maxent	Topographic, oceanographic	Determine potential distribution
Breen *et al.*, 2017	Maxent	Topographic, oceanographic	Risk assessment – interaction with fisheries
Correia *et al.*, 2015[a]	Environmental Envelope, GAM	Spatial, topographic, oceanographic	Assess species – habitat relationship
Correia *et al.*, 2019[b]	GAM	Spatio-temporal, detectability, topographic, oceanographic	Assess species – habitat relationship, test of different type of predictors
Derville *et al.*, 2017	GLM, GAM, Maxent, SVM, BRT	Temporal, topographic, oceanographic	Compare model approaches
Fernandez *et al.*, 2018	Maxent	Topographic, oceanographic	Test of different spatial and temporal scales
Fiedler *et al.*, 2018	GAM, Maxent	Topographic, oceanographic	Compare presence/absence with presence-only models

García-Baron *et al.*, 2019	GAM	Topographic, oceanographic	Assess critical conservation areas
García *et al.*, 2018	GAM	Topographic, oceanographic	Characterize habitat preferences and analyze the temporal distribution
Passadore *et al.*, 2018	Ensemble (GAM+ GBM+CTA+RF+ Maxent)	Anthropogenic, topographic, oceanographic, biological	Investigate the spatio-temporal distribution
Prieto *et al.*, 2017	Maxent	Topographic, oceanographic	Determine habitat-use and niches overlap
Redfern *et al.*, 2017	GAM	Topographic, oceanographic	Exploration of methods to deal with data-poor areas
Storrie *et al.*, 2018	Maxent	Topographic, oceanographic	Assess key habitats and explore spatial overlap
Tobeña *et al.*, 2016	Maxent	Topographic, oceanographic	Estimate potential distribution and species richness
Valente *et al.*, 2019[c]	GAM	Spatio-temporal	Assess distribution patterns

[a,b,c] With CETUS dataset: [a] Chapter IV; [b] Chapter V; [c] Attachment 1.

1.1.5 Protecting cetaceans – marine management and conservation mechanisms

The eastern north Atlantic, in the area delimited west by the Macaronesian archipelagos (Azores and Cape Verde) and east by the mainland coasts of the Iberian Peninsula and northwest Africa, is partially covered by the following conservation agreements and mechanisms (Figure 1):

- **Oslo and Paris Conventions (OSPAR)**. A “mechanism by which 15 Governments & the EU cooperate to protect the marine environment of the North-East Atlantic.” (www.ospar.org);
- **International Council for the Exploration of the Sea (ICES)**. “An intergovernmental marine science organization, meeting societal needs for impartial evidence on the state and sustainable use of our seas and oceans.” (www.ices.dk);
- **Agreement on the Conservation of Cetaceans of the Black Sea, Mediterranean Sea and contiguous Atlantic area (ACCOBAMS)** (http://www.accobams.org) and **Agreement on the Conservation of Small Cetaceans of the Baltic, North East Atlantic, Irish and North Seas (ASCOBANS)** (www.ascobans.org). Both agreements were made under the auspices of the **Convention on Migratory Species (CMS)**. CMS is “an environmental treaty under the aegis of the United Nations Environment Programme” and “provides a global platform for the conservation and sustainable use of migratory animals and their habitats. CMS brings together the States through which migratory animals pass, the Range States, and lays the legal foundation for

internationally coordinated conservation measures throughout a migratory range." (www.cms.int).

Since part of the area is within the national jurisdiction of the European Member States, it is thus covered by the **Habitats Directive**. The Habitats Directive was signed in 1992, aiming "to promote the maintenance of biodiversity, taking account of economic, social, cultural and regional requirements". Within the directive, all cetacean species are under the Annex IV and require a strict protection regime across their entire range in European waters. Bottlenose dolphins (*Tursiops truncatus*) and harbour porpoises (*Phocoena phocoena*) are also under Annex II and thus core areas of their habitat have to be designated Sites of Community Importance (SCIs). The designated protected areas under this directive are included in the **Natura 2000**, the "largest coordinated network of protected areas in the world" (www.ec.europa.eu).

Several protected areas exist within the EEZs, either associated with coastal ecosystems or with seamounts (Figure 1). Recently, the proposals of the project MarPro to create the SCI "Maceda-Praia da Vieira" in the northern coast of Portugal and to extend the SCI "Costa Sudoeste" located in the southwest of Portugal were approved (Resolução do Conselho de Ministros n.º 17/2019; Resolução do Conselho de Ministros n.º 18/2019). These proposals target important areas for bottlenose dolphins and harbour porpoises occurring in the Portuguese mainland coast. Another recent SCI is "Cetáceos da Madeira", located around the Madeira Island. It is the biggest protected area focused on cetaceans in the North Atlantic. In areas beyond national jurisdiction, the Josephine Seamount, located between mainland Portugal and the Madeira island, was designated an OSPAR high-seas marine protected area (Figure 1).

While Habitats Directive focus solely on certain priority species and habitats, the **Marine Strategy Framework Directive (MSFD)**, adopted in 2008, is more holistic and aims "to protect more effectively the marine environment across Europe" (www.ec.europa.eu). The goal is to achieve a Good Environmental Status (GES), settled according to 11 descriptors, by the year 2020. Several indicators of the MSFD are related to cetacean species (for the descriptors 1, 4, 8, 10, 11). The implementation, status, achievements and challenges of the MSFD, focused on cetacean biodiversity, are discussed under Santos & Pierce (2015).

Alongside conservation efforts, the **Convention on Biological Diversity (CBD)** is designating areas for their ecological and biological importance, the **Ecologically and Biologically Significant Marine Areas (EBSAS).** In the eastern north Atlantic, the Instituto da Conservação da Natureza e das Florestas (ICNF) has proposed the designation of the

Madeira-Tore as an EBSA. Madeira-Tore is a complex of a total of 17 seamounts located between mainland Portugal and the Madeira island (Figure 1).

With the growing impact of human activities in areas beyond national jurisdiction, the European Commission adopted a legislation for **Maritime Spatial Planning** that "works across borders and sectors to ensure human activities at sea take place in an efficient, safe and sustainable way". The target is to define the maritime spatial plans by 2021 (www.ec.europa.eu).

Even though important achievements are being attained for the conservation of marine ecosystems, a greater effort will be needed to comply with most of the defined targets by the aforementioned mechanisms. For example, the Strategic Plan for Biodiversity 2010-2020 established that by 2020, 10% of the global ocean should be protected, however, at the moment the percentage falls on the 7.4% (according to the last report, UNEP-WCMC *et al.*, 2018). When comparing protection in areas within the EEZs with areas beyond national jurisdiction, there is a big gap: 16.8% against 1.2% (UNEP-WCMC *et al.*, 2018). Moreover, in many cases, protected areas do not have a management plan defined, or at least, reported (www.protectedplanet.net). In the case of cetacean species, lack of data hinders conservation and the design of management strategies. Knowledge on the occurrence, distribution patterns and ecological niches of cetacean species as well as the identification of data gaps in need of research efforts are crucial to implement effective conservation measures and management plans (Kaschner *et al.*, 2012; Mannocci *et al.*, 2018; Santos & Pierce, 2015).

1.2 Objectives

The aim of this work was to study cetacean distribution and habitat within the eastern north Atlantic, in the area delimited by the coasts of the Iberian Peninsula and northwest Africa and the archipelagos of Macaronesia (study area – Figure 1). The specific objectives were:

1. Provide effort-related data on cetacean occurrence, addressing specific knowledge gaps such as in the high-seas;
2. Identify the spatial and temporal patterns of cetacean distribution, comparing profiles of species diversity among the different sub-regions;
3. Apply ecological niche modelling techniques to: a) study the habitat characteristics and preferences of cetacean species; b) understand the effects of spatio-temporal, detectability and environmental factors on cetacean occurrence; c) fit the models using different spatial and temporal scales of the oceanographic variables; d) produce habitat suitability maps at basin-scale by extrapolating the model results obtained with data collected in the surveyed transects;

4. Highlight areas of special interest for monitoring efforts, e.g., regions with data gaps, with special concerns regarding management and conservation, with high diversity or abundance of cetaceans;
5. Address specific research questions and pinpoint priorities related with cetacean conservation;

1.3 Dissertation Outline

This dissertation is organized as a cohesive and structured compilation of scientific articles accepted, submitted or prepared to submit to international peer-reviewed journals (each one a chapter). **Chapter I** is a conceptual Background that frames and provides context to the subsequent chapters.

The **Chapter II** is a data descriptor published in the journal “Scientific Data” (Correia *et al.*, 2019a). It describes the dataset that was used in the present work (CETUS dataset, 2012-2017), published through the Flanders Marine Institute (VLIZ) IPT portal and distributed by the European Marine Observation and Data Network (EMODnet) and the Ocean Biogeographic Information System (OBIS) (Data citation 1). The methodology of data collection at-sea, namely of the cetacean occurrence, is described in detail, as well as the first steps of data processing, technical verification and storage. As such, this chapter was included as a methodology section common to the remaining chapters.

In the **Chapter III**, the entire dataset from 2012 to 2017 was analysed in order to provide a detailed spatio-temporal description of the cetacean occurrence in the eastern north Atlantic, with a focus on the eight most frequently sighted species: *Delphinus delphis*, *Stenella frontalis*, *Stenella coeruleoalba*, *T. truncatus*, *Ziphius cavirostris*, *Globicephala* sp., *Physeter macrocephalus* and *Balaenoptera acutorostrata*. The area was divided in sub-regions according to administrative boundaries (EEZs and international waters), and the profile of species diversity, as well as encounter rates, were compared between them. This paper is “Under Review” in the journal “Diversity and Distributions”.

Chapter IV is focused on cetacean distribution and occurrence range in the northwestern Africa, between Morocco and Liberia. This area was selected for detailed analysis on cetacean diversity due to existing conservation concerns, mainly related with poorly managed fisheries, that affect cetacean stocks. As such, there is a need to keep an updated inventory of cetacean species and their distribution range. Here, a thorough bibliographic revision on cetacean occurrence data published together with data collected within CETUS (from 2012 to 2016) is provided. Species diversity, range states and number of groups are presented for each state.

Data gaps for species and ranges are highlighted. This manuscript was prepared to be submitted as a “Review” paper.

After looking at spatio-temporal distribution patterns, the first results of habitat preferences and modelling are presented under **Chapter V**. This paper was published in the “Journal of Marine Systems” in 2015 (Correia *et al.*, 2015). The preliminary results were obtained using cetacean occurrence data collected along the route from mainland Portugal to Madeira Island, from 2012 to 2013. Habitat preferences and ecological niche modelling using GAMs are presented for the species *D. delphis*, *T. truncatus* and *P. macrocephalus*; and the groups beaked and baleen whales.

For the most frequently sighted species in the area, *D. delphis*, a separate chapter, **Chapter VI**, was prepared to present an in-depth approach on ecological niche modelling, using data collected between 2012 and 2016. Models were conducted to test the effect of different types of variables on *D. delphis* distribution: detectability, spatio-temporal and environmental predictors. Several spatial and temporal scales of oceanographic variables were tested. A final model combining all variables was developed to present distribution patterns and hotspot areas of *D. delphis* occurrence in the eastern north Atlantic. This paper was published in the “Journal of Marine Biological Association of the United Kingdom” in 2019 (Correia *et al.*, 2019b).

Under **Chapter VII**, using the entire dataset 2012-2017, ecological niche modelling techniques were applied to the eight most frequently sighted species: *D. delphis*, *S. frontalis*, *S. coeruleoalba*, *T. truncatus*, *Z. cavirostris*, *Globicephala* sp., *P. macrocephalus* and *B. acutorostrata*. Two different model techniques, GAM and Maxent, were complemented to describe ecological drivers of cetacean species and deliver maps of habitat suitability. Objective recommendations for research and conservation and management efforts within the area are provided. This manuscript was prepared to be submitted as an “Original research” paper.

Finally, **Chapter VIII** are the Final remarks that highlight the main conclusions from the previous chapters. An overall overview of the work was undertaken, and main results summarized, in order to identify substantial advances made in the field of study. Moreover, future perspectives and research priorities, specifically concerning cetacean conservation, are identified.

All the data analysed in this thesis was collected within CETUS Project. CETUS is a cetacean monitoring programme that started in 2012, led by Isabel Sousa-Pinto and the Interdisciplinary Centre of Marine and Environmental Research (CIIMAR – Porto, Portugal). In partnership with TRANSINSULAR, a Portuguese company for maritime transports, cetacean monitoring is performed from cargo ships along routes between mainland Portugal to the Macaronesian archipelagos and the northwestern coast of Africa. I was involved in this project from the beginning, being in charge of the logistics, partnerships, designing the sampling protocol, training observers, leading the surveys, boarding, processing and analysing the data. Data collected was not only used for the present thesis, but also for several undergraduation and ERASMUS reports, my master thesis plus two other (Correia, 2013; Gil, 2018; Valente, 2017). Besides the scientific papers here presented, CETUS data was used in two other published articles (Gil *et al.*, 2019; Valente *et al.*, 2019), a book (Moura *et al.*, 2017) and 16 conference abstracts. Out of these, I am either first author (six conference abstracts) or co-author (two papers, a book and ten conference abstracts). Finally, data was included within two international reports, from ASCOBANS and ICES – WGMME, and used in the process of the EBSAS definition for Portuguese waters (Attachment 1).

1.4 References

Alves F., Alessandrini A., Servidio A., Mendonça A.S., Hartman K.L., Prieto, R, *et al.* (2018a). Complex biogeographical patterns support an ecological connectivity network of a large marine predator in the north-east Atlantic. *Diversity and Distributions*, 25(2): 269-284.

Alves F., Ferreira R., Fernandes M., Halicka Z., Dias L. and Dinis A. (2018b). Analysis of occurrence patterns and biological factors of cetaceans based on data from platforms of opportunity: Madeira Island as a case study. *Marine Ecology*, e12499.

do Amaral K.B., Alvares D.J., Heinzelmann L., Borges-Martins M., Siciliano S. and Moreno I.B. (2015). Ecological niche modeling of Stenella dolphins (Cetartiodactyla: Delphinidae) in the southwestern Atlantic Ocean. *Journal of Experimental Marine Biology and Ecology*, 472: 166-179.

Arcangeli A., Campana I., Marini L. and MacLeod C.D. (2016). Long-term presence and habitat use of Cuvier's beaked whale (*Ziphius cavirostris*) in the Central Tyrrhenian Sea. *Marine Ecology*, 37: 269-282.

Barragán-Barrera D.C., do Amaral K.B., Chávez-Carreño P.A., Farías-Curtidor N., Lancheros-Neva R., Botero-Acosta N., *et al.* (2019). Ecological Niche Modeling of Three Species of *Stenella* Dolphins in the Caribbean Basin, With Application to the Seaflower Biosphere Reserve. *Frontiers in Marine Science*, 6: 10.

Bowen W.D. (1997). Role of marine mammals in aquatic ecosystems. *Marine Ecology Progress Series*, 158: 267-274.

Breen P., Brown S., Reid D. and Rogan E. (2017). Where is the risk? Integrating a spatial distribution model and a risk assessment to identify areas of cetacean interaction with fisheries in the northeast Atlantic. *Ocean and Coastal Management*, 136: 148-155.

Brito C. and Sousa A. (2011). The Environmental History of cetaceans in Portugal: Ten Centuries of Whale and Dolphin Records. *PLoS ONE*, 6(9): e23951.

Brotons L., Thuiller W., Araújo M.B. and Hirzel, A.H. (2004). Presence−absence versus presence-only modelling methods for predicting bird habitat suitability. *Ecography*, 27: 437-448.

Campbell G.S., Thomas L., Whitaker K., Douglas A.B., Calambokidis J. and Hildebrand J.A. (2015). Inter-annual and seasonal trends in cetacean distribution, density and abundance off southern California. *Deep Sea Research Part II: Tropical Studies in Oceanography*, 112: 143-157.

Carrillo M., Pérez-Vallazza C.P. and Alvarez-Vázquez R. (2010). Cetacean diversity and distribution off Tenerife (Canary Islands). *Marine Biodiversity Records*, 3:1-9.

Cawardine M. (2000). *Whales, Dolphins and Porpoises*. Dorling Kindersley Limited. London, Great Britain.

Correia A.M. (2013). Cetacean monitoring in Northeastern Atlantic Ocean: Occurrence and distribution of cetacean species in the Canary Basin. Master thesis (MSc) in Ecology, Environment and Territory. Sciences Faculty of Porto University, Porto, Portugal. 93 pp.

Correia A.M., Tepsich P., Rosso M., Caldeira R. and Sousa-Pinto I. (2015). Cetacean occurrence and spatial distribution: Habitat modelling for offshore waters in the Portuguese EEZ (NE Atlantic). *Journal of Marine Systems*, 143: 73-85.

Correia A.M., Gandra M., Liberal M., Valente R., Gil A., Rosso M., *et al.* (2019a). A dataset of cetacean occurrences in the Eastern North Atlantic. *Scientific Data*, 6: 177.

Correia A.M., Gil A., Valente R., Rosso M., Pierce G.J. and Sousa-Pinto I. (2019b). Distribution and habitat modelling for short-beaked common dolphins (*Delphinus delphis*) in Eastern North Atlantic Ocean. *Journal of the Marine Biological Association UK*, 1-15.

Derville S., Torres L.G., Iovan C. and Garrigue C. (2018). Finding the right fit: Comparative cetacean distribution models using multiple data sources and statistical approaches. *Diversity and Distributions*, 24(11): 1657-1673.

Djiba A., Bamy I.L., Bilal A.S.O. and Van Waerebeek K. (2015). Biodiversity of cetaceans in coastal waters of northwest Africa: new insights through platform-of-opportunity visual surveying in 2011-2013. *IOC Technical Series*, 115: 283-297.

Elith J. and 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.

Fernandez M., Yesson C., Gannier A., Miller P.I. and Azevedo J.M.N. (2018). A matter of timing: how temporal scale selection influences cetacean ecological niche modelling. *Marine Ecology Progress Series*, 595: 217-231.

Fiedler P.C., Redfern J., Forneu K.A., Palacios D.M., Sheredy C., Rasmussen K., *et al.* (2018). Prediction of Large Whale Distributions: A Comparison of Presence–Absence and Presence-Only Modeling Techniques. *Frontiers in Marine Science*, 5: 419.

Freitas L., Dinis A., Nicolau C., Ribeiro C. and Alves F. (2012). New records of cetacean species for Madeira Archipelago with an updated checklist. *Boletim do Museu Municipal do Funchal*, 62: 25-43.

Friedlaender A., Johnston D.W., Fraser W., Burns J., Halpin P. and Costa D. (2011). Ecological niche modeling of sympatric krill predators around Marguerite Bay, western Antartic Peninsula. *Deep-Sea Research*, 58: 1729-1740.

García-Barón I., Authier M., Caballero A., Vázquez J. A., Santos M. B., Murcia J. L., and Louzao, M. (2019). Modelling the spatial abundance of a migratory predator: A call for transboundary marine protected areas. *Diversity and Distributions*, 25(3): 346-360.

García L.G., Pierce G.J., Autret E., Torres-Palenzuela J.M. (2018). Multi-scale habitat preference analyses for Azorean blue whales. *PLoS ONE* 13(9): e0201786.

Gray H. and Van Waerebeek K.V. (2011). Postural instability and akinesia in a pantropical spotted dolphin, Stenella attenuata, in proximity to operating airguns of a geophysical seismic vessel. *Journal for Nature Conservation*, 19: 363-367.

Gil Á. (2018). Cetáceos na Zona Económica Exclusiva Continental Portuguesa: distribuição espaço-temporal e registo de novas ocorrências. Master thesis (MSc) in Ecology and Environment. Sciences Faculty of Porto University, Porto, Portugal. 99 pp.

Gil Á., Correia A. and Sousa-Pinto I. (2019). Records of harbour porpoise (Phocoena phocoena) in the mouth of the Douro River (Northern Portugal) with presence of an anomalous white individual. *Marine Biodiversity Records*, 12: 1.

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

Hazevoet C.J. and Wenzel W. (2000). Whales and dolphins (Mammalia, Cetacea) of the Cape Verde Islands with special reference to the Humpback Whale *Megaptera novaeangliae* (Borowski, 1871). *Contributions to Zoology* 69(3): 197-211.

Hazevoet C.J., Monteiro V., López P., Varo N., Torda G., Berrow S. and Gravanita B. (2010). Recent data on whales and dolphins (Mammalia: Cetacea) from the Cape Verde Islands, including records of four taxa new to the archipelago. *Zoologia Caboverdiana* 1: 75–99.

Hutchinson G.E. (1957). Concluding remarks. In: Cold Spring Symposium on Quantitative Biology, vol.22, pp. 415-427.

Jackson S.T. and Overpeck J.T. (2000). Responses of plant populations and communities to environmental changes of the late Quaternary. *Paleobiology*, 26: 194-220.

Jungblut S., Nachtsheim D.A., Boos K. and Joiris C.R. (2017). Biogeography of top predators – seabirds and cetaceans – along four latitudinal transects in the Atlantic Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography*, 141: 59-73.

Kaschner K., Quick N.J., Jewell R., Williams R. and Harris C.M. (2012). Global Coverage of Cetacean Line-Transect Surveys: Status Quo, Data Gaps and Future Challenges. *PLoS ONE*, 7(9): e44075.

Kiszka J., Van Canneyt O., Macleod K., Walker D. and Ridoux, V. (2007). Distribution, encounter rates and habitat characteristics of toothed cetaceans in the Bay of Biscay and adjacent waters from platform of opportunity data. *ICES Journal of Marine Science*, 64: 1033-1043.

Mannocci L., Roberts J.J., Halpin P.N., Authier M., Boisseau O., Bradai M.N., *et al.* (2018). Assessing cetacean surveys throughout the Mediterranean Sea: a gap analysis in environmental space. *Scientific Reports*, 8: 3126.

Mason E. (2009). High-resolution Modelling of the Canary Basin Oceanic Circulation. Doctorate thesis (PhD) in Oceanography, Universidad de Las Palmas de Gran Canaria, Las Palmas de Gran Canaria. 245 pp.

Moura A.E., Sillero N. and Rodrigues A. (2012). Common dolphin (*Delphinus delphis*) habitat preferences using data from two platforms of opportunity. *Acta Oecológica*, 38: 24-32.

Moura A.E., Silva S.E., SPEA, Correia A.M., Sousa-Pinto I., Gil A., *et al.* (2017). In: Bencatel J., Álvares F., Moura A. E. and Barbosa, A. M. (eds.) (2017). *Atlas de Mamíferos de Portugal*. Universidade de Évora, Portugal.

Parsons E.C.M. (2016). Why IUCN should replace ‘Data Deficient’ conservation status with a precautionary ‘Assume Threatened’ status—A cetacean case study. *Frontiers in Marine Science*, 3: 193.

Passadore C., Möller L.M., Diaz-Aguirre F. and Parra G.J. (2018). Modelling dolphin distribution to inform future spatial conservation decisions in a marine protected area. *Scientific reports*, 8(1): 15659.

Pearson R.G. (2007). Species’ Distribution Modeling for Conservation Educators and Practitioners. *The American Museum of Natural History*, 3: 54-89.

Pérez-Vallazza C., Álvarez-Vázquez R., Cardon L., Pintado, C. and Hernández-Brito J. (2008). Cetacean diversity at the west coast of La Palma island (Canary Islands). *Journal of the Marine Biolical Association of the United Kingdom*, 88: 1289-1296.

Pershing A.J., Christensen L.B., Record N.R., Sherwood G.D. & Stetson P.B. (2010). The impact of whaling on the ocean carbon cycle: why bigger was better. *PloS ONE*, 5: e12444.

Perrin W.F. and Van Waerebeek K. (2012). *The small-cetacean fauna of the west coast of Africa and Macaronesia: diversity and distribution*. Convention on the Conservation of Migratory Species of Wild Animals, Bonn, Germany, CMS Technical Series, 26, 7-17.

Praca E., Gannier A., Das K. and Laran S. (2009). Modelling the habitat suitability of cetaceans: example of the sperm whale in the northwestern Mediterranean Sea. *Deep-Sea Research I*, 56: 648-657.

Prieto R., Tobeña M. and Silva M.A. (2017). Habitat preferences of baleen whales in a mid-latitude habitat. *Deep Sea Research Part II: Tropical Studies in Oceanography*, 141: 155-167.

Redfern J.V., Ferguson M.C., Becker E.A., Hyrenbach K.D., Good C., Barlow J., *et al.* (2006). Techniques for cetacean-habitat modeling: a review. *Marine Ecology Progress Series*, 310: 271-295.

Redfern J.V., Moore T.J., Fiedler P.C., de Vos A., Brownell R.L., Forney K.A., *et al.* (2017). Predicting cetacean distributions in data-poor marine ecosystems. *Diversity and Distributions*, 23(4): 394-408.

Robineau D. and Vely M. (1998). Ces cétacés des côtes de Mauritanie (Afrique du Nordouest). Particularltés et variations spatio-temporelles de repartition: role des facteurs oceanographiques. *Revue d Ecologie (Terre Vie)*, 53.

Roman J., Estes J.A., Morissette L., Smith C., Costa D., McCarthy J., *et al.* (2014). Whales as marine ecosystem engineers. *Frontiers in Ecology and the Environment*, 12.

Santos M.B. and Pierce G.J. (2015). Marine mammals and good environmental status: science, policy and society; challenges and opportunities. *Hydrobiologia* 750: 13-41.

Sergio, F., Newton, I., Marchesi, L. and Pedrini, P. (2006). Ecologically justified charisma: preservation of top predators delivers biodiversity conservation. *Journal of Applied Ecology,* 43, 1049-1055.

Sergio F., Caro T., Brown D., Clucas B., Hunter J., Ketchum J., *et al.* (2008). Top Predators as Conservation Tools: Ecological Rationale, Assumptions, and Efficacy. *Annual Review of Ecology, Evolution, and Systematics*, 39: 1-19.

Sillero N. (2011). What does ecological modelling model? A proposed classification of ecological niche models based on their underlying methods. *Ecological Modelling*, 222: 1343-1346.

Silva M.A., Prieto R., Cascão I., Seabra M.I., Machete M., Baumgartner M.F. and Santos R.S. (2014). Spatial and temporal distribution of cetaceans in the mid-Atlantic waters around the Azores. *Marine Biology Research*, 10:123-137.

Sousa-Lima R.S., Norries T.F., Oswald J.N. and Fernandes D.P. (2013). A Review and Inventory of Fixed Autonomous Recorders for Passive Acoustic Monitoring of Marine Mammals. *Aquatic Mammals*, 39(1): 23-53.

Smith C.R. (2007). Bigger is better: the role of whales as detritus in marine ecosystems. *Whales, whaling and ocean ecosystems*, 284-299.

Storrie L., Lydersen C., Andersen M., Wynn R.B. and Kovacs K.M. (2018). Determining the species assemblage and habitat use of cetaceans in the Svalbard archipelago, based on recorded observations from 2002-2014. *Polar Research*, 37 (1): 1463065.

Tepsich P., Rosso M., Halpin P.N. and Moulins A. (2014). Habitat preferences of two deep-diving cetacean species in the northern Ligurian Sea. *Marine Ecology Progress Series*, 508: 247-260.

Thorne L.H., Johnston D.W., Urban D.L., Tyne J., Bejder L., Baird R.W., *et al*. (2012). Predictive modeling of spinner dolphin *Stenella longirostris* resting habitat in the Main Hawaiian Islands. *PLoS ONE*, 7: e43167.

Tobeña M., Prieto R., Machete M. and Silva M.A. (2016). Modeling the potential distribution and richness of cetaceans in the Azores from fisheries observer program data. *Frontiers in Marine Science*, 3: 202.

UNEP-WCMC, IUCN & NGS (2018). *Protected Planet Report 2018*. UNEP-WCMC, IUCN and NGS: Cambridge UK; Gland, Switzerland; and Washington, D.C., USA.

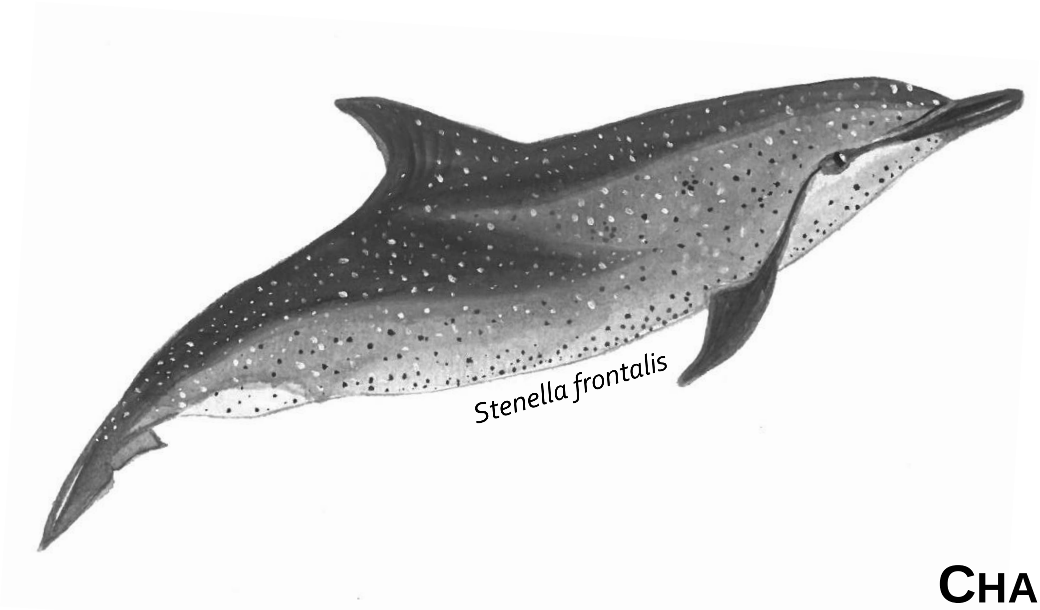
Valente R. (2017). Looking for the migratory whales: Routes of the baleen whales in the Macaronesia. Master thesis (MSc) in Biodiversity, Genetics and Evolution. Sciences Faculty of Porto University, Porto, Portugal. 142 pp.

Valente R., Correia A.M., Gil Á., González L.G. and Sousa-Pinto I. (2019). Baleen whales in Macaronesia: occurrence patterns revealed through a bibliographic review. *Mammal Review*, 49(2): 129-151.

Viddi F.A., Hucke-Gaete R., Torres-Florez J.P. and Ribeiro S. (2010). Spatial and seasonal variability in cetacean distribution in the fjords of northern Patagonia, Chile. *ICES Journal of Marine Science*, 67: 959-970.

Weir C.R. and Pierce G.J. (2013). A review of the human activities impacting cetaceans in the eastern tropical Atlantic. *Mammal Review*, 43: 258-274.

Data Citation 1: Correia, A.M., Gandra, M., Liberal, M., Valente, R., Gil, A., Rosso, M., Pierce, G.J. and Sousa-Pinto, I, CIIMAR - UP. (2019). CETUS: Cetacean monitoring surveys in the Eastern North Atlantic. Marine Data Archive. https://doi.org/10.14284/350.



CHAPTER II.

Paper published in the journal Scientific Data, 2019. doi: 10.1038/s41597-019-0187-2

2 A dataset of cetacean occurrences in the Eastern North Atlantic

Ana M. Correia[1,2], Miguel Gandra[1,3], Marcos Liberal[4], Raul Valente[1,2], Ágatha Gil[1], Massimiliano Rosso[1,5], Graham J. Pierce[6,7,8], Isabel Sousa-Pinto[1,2]

[1]Interdisciplinary Centre of Marine and Environmental Research (CIIMAR). 4450-208 Matosinhos, Portugal.
[2]Department of Biology, Faculty of Sciences, University of Porto (FCUP). 4169-007 Porto, Portugal.
[3]Centre of Marine Sciences (CCMAR), University of the Algarve, Campus de Gambelas, 8005-139 Faro, Portugal.
[4]Fraunhofer AICOS. 4200-135 Porto, Portugal.
[5]CIMA Research Foundation. 17100 Savona, Italy.
[6]Instituto de Investigacións Mariñas (CSIC). 36208 Vigo, Pontevedra, Spain.
[7]Oceanlab, University of Aberdeen. AB41 6AA, UK.
[8]CESAM and Department of Biology, University of Aveiro. 3810-193 Aveiro, Portugal.

2.1 Abstract

The CETUS project is a cetacean monitoring program that takes advantage of cargo ships to undertake survey routes between Continental Portugal, Macaronesian archipelagos and West Africa. From 2012 to 2017, over 50 volunteers participated in the program, actively surveying more than 124.000 km, mostly beyond national jurisdictions in the high seas, for which little or no previous data existed. In total, the collection comprises 3058 georeferenced transect lines and 8913 positions, which are associated with 2833 cetacean sightings, 362 occurrences of other pelagic megafauna, 5260 estimates of marine traffic and 8887 weather observations. This dataset may provide new insights into the distribution of marine mammals in the Eastern North Atlantic and was published following the OBIS-ENV-DATA format (with the most recent biodiversity data standards at the time of writing). Consequently, it may serve as a model for similar visual line transect data collections yet to be published.

2.2 Background & Summary

The assessment of the distribution of pelagic apex predators is one of the research priorities for marine management and conservation as these predators are frequently key species in marine ecosystems (Heithaus *et al.*, 2008). However, given their occurrence in the open ocean and their transboundary movements, assessing their distribution is logistically challenging and, consequently, the high seas generally remain poorly surveyed (Hammond *et al.*, 2013; Jungblut *et al.*, 2017; Williams *et al.*, 2006). Moreover, the few apex predator occurrences recorded in the open ocean are mostly opportunistic, lacking associated data such as search effort, weather conditions or the presence of marine traffic (Richardson *et al.*, 2012). At the time of writing, 691 datasets containing 772827 records on cetacean occurrence were compiled in the OBIS-SEAMAP portal (Ocean Biogeographic Information System Spatial Ecological Analysis of Megavertebrate Populations; http://seamap.env.duke.edu). However, until 2016, only 5% of all the marine mammal records were from areas beyond national jurisdiction (i.e., beyond the ~ 200 nautical mile limit of National Exclusive Economic Zones - EEZ) and nearly 35% of all observations were within the United States EEZ. Moreover, occurrences were also unevenly distributed among species, with few or no records for nearly 26% of cetacean species (Miloslavich *et al.*, 2016). This reflects the heterogeneity in sampling effort and the limitations imposed by logistic, economic and weather constraints.

The CETUS project aims to address these gaps and to study the distribution, habitat characteristics and diversity of cetaceans in the Eastern North Atlantic, with a focus on less surveyed areas, in particular the high seas. CETUS is a cetacean monitoring program that started in 2012, led by the Interdisciplinary Centre for Marine and Environmental Research

(CIIMAR). Through a partnership with TRANSINSULAR, a Portuguese company for maritime transport, marine mammal observers (MMOs) board on cargo ships on routes between Continental Portugal, Macaronesian archipelagos and West Africa, to provide new insights into the distribution and occurrence of cetaceans in the Eastern North Atlantic. Use of these so-called "platforms of opportunity" enables the sampling of large geographical areas during extended periods of time at relatively low cost and they are therefore widely used in situations where it is difficult to mount a dedicated marine campaigns (Aïssi *et al.*, 2015; Alves *et al.*, 2018; Correia *et al.*, 2015; Kiszka *et al.*, 2007; Morgado *et al.*, 2017; Moura *et al.*, 2012; Tobeña *et al.*, 2016; Viddi *et al.*, 2010). In addition to the presence of cetacean species and the occurrence of other pelagic megafauna, data on survey effort, weather conditions and marine traffic, among other variables, were collected during the surveys. In total, from 2012 to 2017, observers boarded 430 trips (a trip being a journey from one port to another), surveyed more than 124 000 km on-effort, and registered 8913 positions associated with cetacean sightings (2833 records), other pelagic megafauna occurrences (362 records), marine traffic (5260 records) and weather conditions (8887 records) (Figure 1).

Part of the information available in the dataset presented here (Data citation 1) was already used for developing cetacean habitat models (Correia *et al.*, 2015; 2019), to study migratory movements of baleen whales (Valente *et al.*, 2019) and to assess their abundance in the North of Continental Portugal (Gil *et al.*, 2018). The monitoring program is still on-going and collecting data with the aim of building the first long-term, wide-range, open-source dataset on cetacean occurrence and distribution in the Eastern North Atlantic. This will allow the study of species distribution patterns, trends in relative abundance, migratory routes and basin-scale habitat use, and will provide important baseline data available to decision-makers and for future management and conservation initiatives.

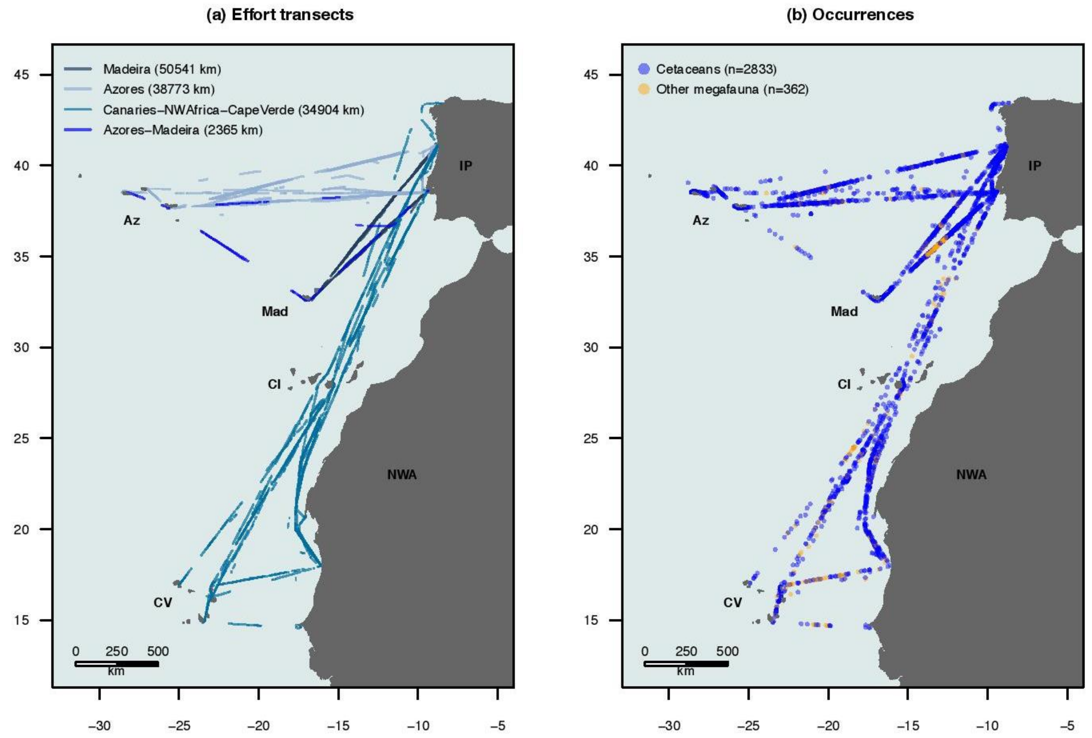


Figure 1. Study area with the on-effort transects and positions included in the dataset. On-effort transects (a); Occurrences of cetaceans and other pelagic megafauna (b). IP – Iberian Peninsula; NWA – Northwest Africa; Az – Azores; Mad – Madeira; CI – Canary Islands; CV – Cape Verde. Map coordinates are presented in the decimal format.

2.3 Methods

2.3.1 Geographic location

The dynamic oceanographic processes and complex topographic structures in the Eastern North Atlantic result in a wide variety of habitats which, in turn, support high levels of marine biodiversity (Caldeira *et al.*, 2002; Mason, 2009; Sala *et al.*, 2013). Between 2012 and 2017, four different routes were monitored under the CETUS project, all starting in Continental Portugal: Portugal - Madeira (from 2012), Portugal - Azores (from 2014), Portugal - Cape Verde (with stopovers in Canary Islands, Mauritania and Senegal, from 2015) and Portugal - Azores - Madeira (in 2017). The surveyed area ranged in latitude from the north of the Iberian Peninsula (43.4408º N) to Dakar (14.5748º N); and in longitude from northwest Spain (8.3093º W) to the Azores (31.1475º W) (Figure 1). From the literature, a total of 17 cetacean species has been recorded along the Continental Portuguese coast (Brito & Sousa, 2011; Moura *et al.*, 2017), 26 in Madeira (Alves *et al.*, 2018, Freitas *et al.*, 2012), 28 in the Azores (Tobeña *et al.*, 2016, Silva *et al.*, 2014), 28 in the Canary Islands (Carrillo *et al.*, 2010; Pérez-Vallazza *et al.*, 2008), 24 in Cape Verde (Hazevoet & Wenzel, 2000; Hazevoet *et al.*, 2010) and 36 along the Northwestern African coast (Djiba *et al.*, 2015; Perrin & Van Waerebeek, 2012; Robineau & Vely, 1998; Weir & Pierce, 2014).

2.3.2 Field surveys

Since 2012, volunteers have been selected to participate as MMOs in the CETUS surveys. From 2014, this selection has been made through an international call that prioritizes volunteers with previous experience in sea-surveys and in marine mammal identification. Nonetheless, each year the MMO team receives an intensive course on line-transect survey protocol and on marine mammal identification before embarking. Each ship then receives a team of two observers, one of them always being experienced in the CETUS protocol (i.e., someone who has previously participated in CETUS campaigns). Each team stays on-board the same ship and route for, at least, one round trip starting in Continental Portugal (boarding and disembarking in the ports of Leixões-Porto or Lisbon).

The sampling protocol follows the standard for line-transect cetacean surveys (Correia *et al.*, 2015; 2019; Hammond *et al.*, 2013; Tepsich *et al.*, 2014) and is performed from sunrise to sunset. Observers stand in the wings of the navigation bridge (at a height of between 13.5 and 16 metres above sea level, considering maximum draught, and depending on the ship). Sporadically, when weather conditions are unsuitable (i.e., strong winds or moderate rain) but there is still adequate visibility, the monitoring is carried from inside the navigation bridge, to ensure continuation of the survey effort. Each MMO stands on one side of the vessel and they switch every 60 minutes (approximately) to avoid fatigue and data-biases. Moreover, in turns, both take one-hour breaks for meals and two optional rests of up to 40 minutes (one in the morning and another during the afternoon).

Monitoring is performed from the front of the vessel with a field of view of approximately 180º, with each MMO covering 90º (except at mealtimes and resting periods, in which case the lone MMO at the survey stand covers the entire 180º from one of the sides range). Binoculars (Porro Prism marine binoculars, with a compass and a distance scale with seven or eight reticles, 7 x 50 mm) are used for occasional scans and to support species identification and group size estimation. Given the nature of the project with the monitoring being performed from a platform of opportunity, it was not possible to have a second independent observer to validate the data collected. Only two observers were allowed to board on each ship (all performing different routes in different periods) and both would stay at the highest point of observation possible performing the monitoring protocol together in order to guarantee a thorough scan of the 180º ahead.

The route of the ship during sampling and the positions marked by the MMOs are recorded using a tablet with an inbuilt GPS and running the application MyTracks (https://my-tracks.pt.aptoide.com). This application registers, among other variables, the date and time (programmed for GMT+0), the speed (in m/s) and direction (in º) of the vessel and the GPS coordinates (in decimal format, WGS84 coordinate system). The recommended recording settings are maintained: recording every 10 seconds or 10 metres (whichever the smallest) and with minimum precision of 50 m. These settings were occasionally changed in order to overcome battery life issues. The application, although working efficiently at sea, on rare occasions, generates errors in the date and time recording (with the time going forwards and backwards). This issue demands a careful verification process during data entry. An adaptation of the protocol was made during the first year of the project (2012 campaigns), in which a Garmin GPS (with similar settings) was used, with positions being annotated by hand and later imported into Microsoft Excel spreadsheets. During this year, the use of alternative survey stands (e.g. deck of the ship) was explored. Besides data on cetacean occurrences, information about the weather conditions, marine traffic and the presence of other pelagic megafauna is also collected. To record weather conditions, observers assess sea state (using the Douglas scale), wind speed (using the Beaufort scale), visibility (on a categorical scale of values from 1 – 10 thus covering visibility ranges from 0 m to more than 50000 m, estimated based on the definition of the horizon line and reference points at a known range, e.g., ships with an AIS system) and the occurrence of rain. This information is registered at the beginning and end of each survey leg (a survey leg being a continuous period of sampling) and every time there are significant changes in the conditions. For the marine traffic, small and big vessels (less than and over 20 m in length), detected with or without binoculars, all around the ship´s position, are registered at the beginning and end of each survey, at every sighting of cetacean species and every hour. For pelagic megafauna other than cetaceans, the data are always collected opportunistically, as sampling effort is dedicated uniquely to cetacean species. In these cases, only taxonomic information and the number of individuals (as well as optional comments about the sighting, e.g., animals behaviours, presence of calves or others) are registered. Whenever MMOs cannot gain access to the survey stand (e.g. during safety drills, manoeuvres) or when weather conditions are unfavourable for cetacean monitoring (e.g. at Beaufort or Douglas values >4, visibility <1 km or heavy rain), the sampling effort stops and any data collected until effort resumes are considered opportunistic (off-effort).

Whenever a cetacean species is sighted, both observers gather on the relevant side of the boat and mark the end of an on-effort transect in order to correctly collect the data on the occurrence. After registering the sighting, a new on-effort transect starts. Species identification is attempted to the species level, although the identity assigned is always at the taxonomic level at which the MMOs are confident of their identification. Angle of the sighting and vessel direction (angle to the bow) is measured with the binoculars’ compass. Then, these measures are used to calculate the horizontal angle in degrees between the ship’s route and the line to the animal or group of animals (bearing). This bearing ranges between 0 and 360 degrees and is measured in a clockwise direction starting from the ship's heading (i.e., 0 degrees). As the binoculars´ compasses can be unreliable on platforms containing ferrous metals, the vessel heading is also measured using the binoculars´ compass whenever an animal is sighted. Then, this value is compared with the direction of the route as measured using the GPS, to obtain the approximate error of the compass and correct the registered horizontal angle (during data processing). Additionally, MMOs also measure the vertical angle to the animals by using the reticular-scale in the binoculars (horizontal equally-spaced marks inscribed in the lens). This range estimation method, involves placing the uppermost reticule on the horizon and counting down to the sighted animals. Together with an estimated observation height, these measurements can be then used to calculate the approximate distance to the animals (based on simple trigonometry). It is important to note, however, that for these calculations a mean estimate of the height of the eye-level of the observers should be added to the platform height measures supplied in the dataset, since exact observation heights were not kept. For group size measures, the observers provide an estimate of the minimum, maximum and assumed (best estimate) number of individuals in a sighting. Moreover, whenever possible, information on the heading of the group and its behaviour towards the ship (i.e. approaching, indifferent or avoiding) is also collected.

2.3.3 Data processing

After collection, all data between 2012 and 2017 obtained from the GPS recordings and stored in CSV files were imported into Microsoft Excel 2016 spreadsheets and processed on a survey-by-survey basis. Data recorded on paper sheets (in 2012), were entered into the spreadsheets by hand. Data-cleaning procedures were carried out throughout, involving, for example, the conversion of variables to the metric system and correcting the bearing to the sighted animals using the estimated error of the binoculars, as mentioned above. After these operations, data were uploaded to a MySQL database (https://www.mysql.com) for permanent storage and to easily perform queries for verification, validation and export of subsets. All records were then imported into ArcGIS (https://www.esri.com) in order to visualize and correct occasional inaccuracies in the coordinates, and to create the on-effort transect lines with the Data Management tool "Points to Line". The resulting polyline shapefiles were then used to calculate the effort distances in km (using the Mercator projection) and to describe the transects in the "well-known text" (WKT) format.

2.3.4 Data management and standardization

In order to comply with the current biodiversity data standards and provide as much information as possible, the entire dataset was reprocessed based on the recently developed OBIS-ENV-DATA format (De Pooter *et al.*, 2017). This new structure was designed for sampling-event datasets, and thus enables the capture of much more detailed information than is found in the widespread "occurrence-only" datasets. It allows the inclusion of important details about the nature of the sampling / observation methods, as well as providing the opportunity to record a multiplicity of biological or environmental measurements collected together alongside the occurrence data. Its underlying conceptual model is a star-like schema (Darwin Core Archive), where one core data file is associated with one or more extension data files through common database keys, i.e. ID fields. In the present format, the core data file holds information about the sampling events (i.e. geographic coordinates, date, protocol, etc.). The occurrences (taxonomical information) and biological / environmental measurements are stored in two separate files: the Occurrence and the ExtendedMeasurementOrFact (eMoF) extensions.

The dataset was exported from the MySQL database and restructured into the appropriate relational format, using custom-written routines in the R environment (R Core Team, 2012). All variables were renamed to match the corresponding Darwin Core (DwC) Terms, with a total of 42 controlled terms being used to describe the collected information (for the full list please see Figure 2).

To reduce data redundancy and facilitate interpretation, an event hierarchy with successive "one-to-many" relationships was created (Figure 2) comprising four different levels:

1. Here we included events describing the 4 main vessel routes along which surveys were conducted, representing the boundary of the surveyed area in the *footprintWKT* field.
2. Here we described all individual trips, with information about the arrival/departure ports being added to the *eventRemarks* field and the corresponding coordinates being added to the *footprintWKT* as a "MultiPoint" geometry. Additional details about the vessels (name, IMO number and platform height, i.e., distance between the sea-level and the wings of the navigation bridge considering maximum draught), together with the name of the MMOs on-board, were added to the eMoF extension
3. Within each cruise, two types of events were included: the on-effort transects and the off-effort positions (opportunistic sightings or locations where weather/marine traffic measurements were recorded outside effort time). The on-effort transect events include an approximate representation of the survey track in WKT format, as well as sampling dates converted to the ISO 8601 standard. Since most of the transects were approximately linear, the segments were simplified using the Douglas-Peuker algorithm

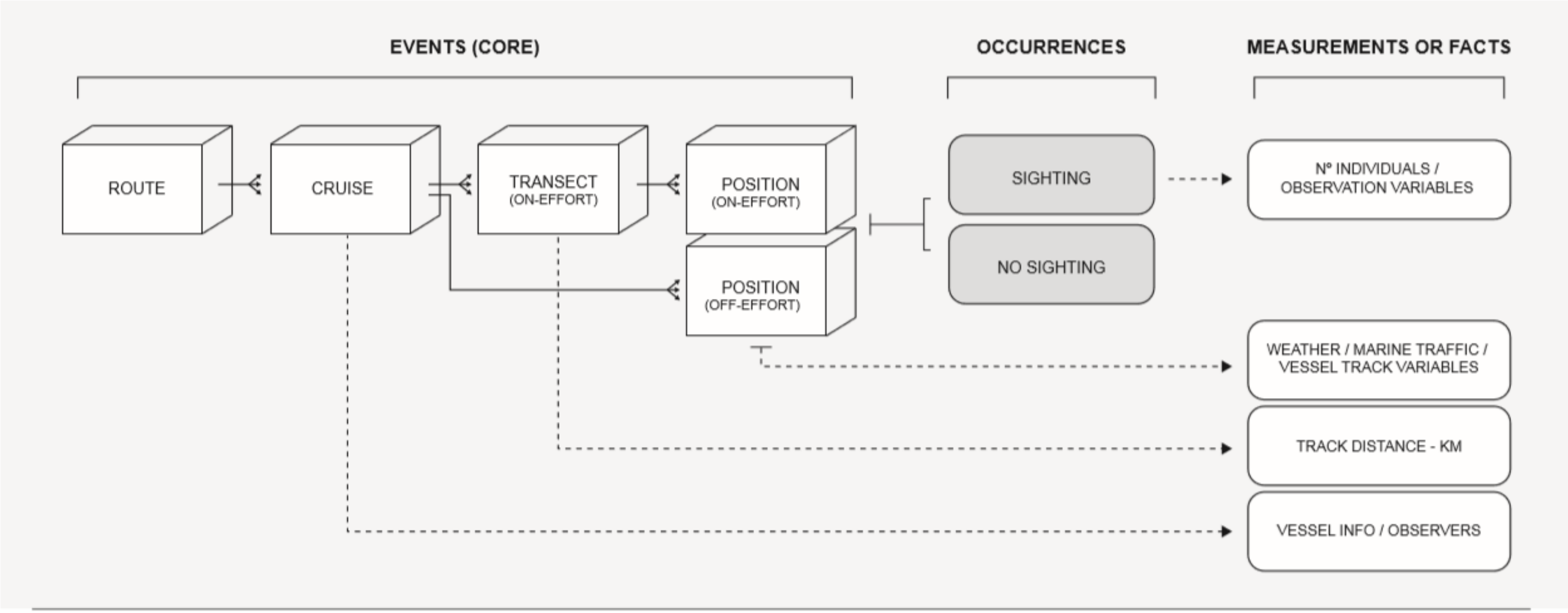
(though the *gSimplify* function available in "rgeos" package; Bivand *et al.*, 2018) with the objective of decreasing the length of the WKT strings. In some cases, this process resulted in minor differences (mean diff. <1 km) between the distances originally recorded and those resulting from the simplified segments. In order to preserve the maximum accuracy possible, the effort distances (which are added to both the *samplingEffort* field and the eMoF extension) refer to the original high-resolution trajectories and not to the simplified geometries included in the WKT field. On 14 occasions, only a single GPS position could be logged during the transect, either due to a sighting occurring immediately after the start of the monitoring effort (survey effort stops upon a sighting) or due to rapid changes in weather conditions that hindered the ability of the MMOs to proceed with the survey. In such cases the WKT portrays only the recorded point and the associated track distance was set to zero. Whenever observations were not registered during a transect, we added the coordinates of the centre of the segment to the *decimalLongitude* and *decimalLatitude* terms and calculated the distance between this point and the start or end of the line to include in the *coordinateUncertaintyInMeters* term. This step was required since "absence" records must be associated with specific coordinates in the current repositories.

4. Again associated with each transect, we added the on-effort positions. Both the on and off-effort positions include detailed geographic coordinates (*decimalLongitude* and *decimalLatitude*) in the EPSG:4326 spatial reference system (WGS84), information about the georeferencing precision (*coordinateUncertaintyInMeters*) and, similarly to the transects, the date in the ISO 8601 format. Information on the vessel speed and heading at each position was added to the eMoF file with the corresponding *eventID* code. To provide improved readability and make the structure more easily perceivable, all event identifiers (*eventID* term) were generated based on the position of each record within the hierarchy, by repeating the preceding parent events on the child records (separated by ":") and using sequentially assigned numbers, together with a "-OC" suffix to identify positions associated with cetacean sightings and a "-OO" suffix to identify positions associated with other megafauna occurrences.

After completing the structure of the core file, occurrence records were subsequently created by associating each position to the observed taxon and creating absence records whenever cetacean species were not detected. The WoRMS (World Register of Marine Species) webservice was used to validate the scientific names and to extract all the additional taxonomic variables (*scientificNameID*, *taxonRank*, *kingdom*, *phylum*, *class*, *order*, *family* and *genus*). When available, information about taxa associated with the occurrence and their behaviour towards the ship was added to the *associatedTaxa* and *behaviour* DwC terms respectively.

Group size estimates (best, maximum and minimum) were added to the eMoF extension, together with information on the bearing to the sighted animals, the number of reticules below horizon and the type of binoculars used, to allow distance-based calculations (Thomas *et al.*, 2002). However, for these calculations, the user of the dataset has to be aware of the limitations: height of the platform provided is based on maximum draught (which is not always the case), eye-level height will be an average height, possible errors during angle measurements of the observer and of the GPS. The best estimate of the number of individuals was also included in the field *organismQuantity*, for cross-compatibility with other dataset formats.

Finally, all remaining meteorological (weather) and marine traffic measurements were linked to the correspondent events and grouped together in the eMoF file. Whenever possible, the corresponding Unique Resource Identifiers (URIs) were added to the *measurementTypeID* and *measurementUnitID* fields using the NERC Vocabulary Server developed by the British Oceanographic Data Centre (BODC), as recommended for semantic standardization.



DARWIN CORE TERMS

- eventID
- type
- eventDate
  (interval)
- footprintWKT
  (bounding box)
- waterBody
- eventRemarks
  (route description)
- institutionCode
- datasetName
- modified

- eventID
- parentEventID
- type
- eventDate
  (interval)
- footprintWKT
  (departure/arrival points)
- samplingProtocol
- samplingEffort
- eventRemarks
  (departure/arrival port)
- institutionCode
- datasetName
- modified

- eventID
- parentEventID
- type
- eventDate
- footprintWKT
  (survey track)
- samplingEffort
- institutionCode
- datasetName
- modified

- eventID
- parentEventID
- type
- eventDate
- decimalLongitude
- decimalLatitude
- geodeticDatum
- coordinateUncertaintyInMeters
- eventRemarks
  (on-effort/off-effort)
- institutionCode
- datasetName
- modified

- occurrenceID
- eventID
- scientificName
- scientificNameID
- taxonRank
- kingdom
- phylum
- class
- order
- family
- genus
- basisOfRecord
- occurrenceStatus
- organismQuantity
- organismQuantityType
- associatedTaxa
- behaviour
- recordedBy

- measurementID
- eventID
- occurrenceID
- measurementType
- measurementTypeID
- measurementValue
- measurementValueID
- measurementUnit
- measurementUnitID
- measurementRemarks

Figure 2. Simplified dataset structure, based on the OBIS-ENV-DATA format. The events hierarchy is outlined under the Event Core. The included Darwin Core (DwC, https://dwc.tdwg.org/terms) fields are depicted below each category.

2.3.5 Data Records

The final dataset was published through the Flanders Marine Institute (VLIZ) IPT portal and distributed by EMODnet and OBIS (Data citation 1) and can be downloaded as a self-contained file (Darwin Core Archive, DwC-A). The present data descriptor is based on version 1.0 of the dataset.

In total, the dataset contains 12405 events, 9440 occurrences and 86022 measurements or facts, spanning the period from 2012 to 2017 and including nearly 500 days of on-effort surveys. Overall, the collection comprises 44 taxa, 30 of which belong to the infraorder Cetacea (approximately 89% of the sightings). The remaining occurrences, opportunistically registered, consisted mostly of marine turtles (families Cheloniidae and Dermochelyidae, accounting for nearly 9% of the observations) and elasmobranchs (55 sightings, approximately 2% of the observations). Due to occasional visibility limitations inherent to shipboard surveys and in order to ensure a high accuracy of identification, only 46% of the sightings were listed to the species level (Table 1). In total, 30 different species were identified, with more than one third of these being classified as vulnerable, near threatened or endangered by the IUCN Red List (https://www.iucnredlist.org). All records are georeferenced in decimal degrees and include the corresponding dates, resolved to day. Moreover, all on-effort occurrences are associated with the corresponding geospatial transect lines, as well as several observation and weather-related variables to allow end-users to conduct additional analyses.

To our knowledge, this is one of the first cetacean survey datasets to be made available in the recent OBIS-ENV-DATA format and thus we believe it can provide a model to be considered when similar visual line transect data collections are assembled in the future.

Table 1. Number of occurrences of each recorded taxa.

Taxa	Taxon rank	Number of occurrences
Elasmobranchii	Class	25
Myliobatiformes	Order	1
Cetacea	Infraorder	270
Mysticeti	Superfamily	279
Cheloniidae	Family	173
Delphinidae	Family	750
Istiophoridae	Family	1
Myliobatidae	Family	16
Sphyrnidae	Family	11
Ziphiidae	Family	121
Globicephala	Genus	59
Kogia	Genus	5
Morus	Genus	2
Thunnus	Genus	2
Balaenoptera acutorostrata	Species	92
Balaenoptera borealis	Species	4

Balaenoptera edeni	Species	6
Balaenoptera musculus	Species	3
Balaenoptera physalus	Species	33
Caretta caretta	Species	108
Chelonia mydas	Species	1
Delphinus delphis	Species	394
Dermochelys coriacea	Species	3
Grampus griseus	Species	8
Hyperoodon ampullatus	Species	4
Lagenodelphis hosei	Species	1
Manta birostris	Species	2
Megaptera novaeangliae	Species	9
Mesoplodon densirostris	Species	5
Mola mola	Species	16
Monachus monachus	Species	1
Orcinus orca	Species	5
Peponocephala electra	Species	4
Phocoena phocoena	Species	4
Physeter macrocephalus	Species	152
Pseudorca crassidens	Species	13
Stenella attenuata	Species	8
Stenella clymene	Species	16
Stenella coeruleoalba	Species	154
Stenella frontalis	Species	226
Stenella longirostris	Species	6
Steno bredanensis	Species	4
Tursiops truncatus	Species	134
Ziphius cavirostris	Species	64
TOTAL	**44 taxa**	**3195 occurrences**

The number of occurrences is presented by taxa recorded to the highest possible level. The table is organized by taxon rank of the records and alphabetically within.

2.3.6 Technical Validation

Every year, within CETUS Project, observers that board on the cargo ships receive an intensive training on both the sampling protocol and marine mammals' identification. Moreover, during selection process, applicants are evaluated and selected according to their previous experience on cetacean identification and fieldwork at sea; and the interest of the observers to participate in CETUS as part of an internship for academic purposes is encouraged to guarantee personal interest in the surveys. Within the boarding team, at least one of the observers has experience in the survey protocol (i.e., boarded before with the CETUS Project). Observers only register the identifications to the taxonomic level they are confident with, hence, the accuracy of the collected data is assured.

During data processing and development of the final dataset, verifications and validations were made at several stages: during the digitalization of the data to the excel files, within the MySQL database, in ArcGIS and in R, after structuring the final dataset. Up to the ArcGIS stage, these verifications were made every year (i.e., after compiling the data for that year). Posteriorly, a thorough validation was undertaken by the members of the research team. This included checking mismatching codes, cross-crossing dates and coordinates across the dataset schema (i.e., route-cruises-segments-positions), confirming segment and effort distances, verifying taxonomy and standardizing nomenclature.

2.4 Acknowledgements

A special acknowledgement to all CETUS volunteers for their contribution and dedication during the monitoring campaigns. We are extremely grateful to TRANSINSULAR that provided all the logistic support for data collection, and to the ship crews for their hospitality.

This work was conducted within a PhD program under the grant SFRH/BD/100606/2014 attributed by the Portuguese national funding agency for science, research and technology (FCT).

2.5 References

Aïssi M., Arcangeli A., Crosti R., Yahia M.N.D., Loussaief B., Moulins A., *et al.* (2015). Cetacean Occurrence and Spatial Distribution in the Central Mediterranean Sea Using Ferries as Platform of Observation. *Russian Journal of Marine Biology*, 5: 343-350.

Alves F., Ferreira R., Fernandes M., Halicka Z., Dias L. and Dinis A. (2018b). Analysis of occurrence patterns and biological factors of cetaceans based on data from platforms of opportunity: Madeira Island as a case study. *Marine Ecology*, e12499.

Bivand R. and Rundel C. (2018). Rgeos: Interface to Geometry Engine-Open Source (GEOS). R package version 0.4-2.

Brito C. and Sousa A. (2011). The Environmental History of cetaceans in Portugal: Ten Centuries of Whale and Dolphin Records. *PLoS ONE*, 6(9): e23951.

Caldeira R.M.A., Groom S., Miller P., Pilgrim D. and Nezlin N.P. (2002). Sea-surface signatures of the Island mass effect phenomena around Madeira Island, Northeast Atlantic. *Remote Sensing of Environment*, 80: 336-360.

Carrillo M., Pérez-Vallazza C.P. and Alvarez-Vázquez R. (2010). Cetacean diversity and distribution off Tenerife (Canary Islands). *Marine Biodiversity Records*, 3:1-9.

Correia A.M., Tepsich P., Rosso M., Caldeira R. and Sousa-Pinto I. (2015). Cetacean occurrence and spatial distribution: Habitat modelling for offshore waters in the Portuguese EEZ (NE Atlantic). *Journal of Marine Systems*, 143: 73-85.

Correia A.M., Gil A., Valente R., Rosso M., Pierce G.J. and Sousa-Pinto I. (2019). Distribution and habitat modelling for short-beaked common dolphins (*Delphinus delphis*) in Eastern North Atlantic Ocean. *Journal of the Marine Biological Association UK*, 1-15.

De Pooter D., Appeltans W., Bailly N., Bristol S., Deneudt K., Eliezer M., *et al.* (2017). Toward a new data standard for combined marine biological and environmental datasets-expanding OBIS beyond species occurrences. *Biodiversity Data Journal*, 5: e10989.

Djiba A., Bamy I.L., Bilal A.S.O. and Van Waerebeek K. (2015). Biodiversity of cetaceans in coastal waters of northwest Africa: new insights through platform-of-opportunity visual surveying in 2011-2013. *IOC Technical Series*, 115: 283-297.

Freitas L., Dinis A., Nicolau C., Ribeiro C. and Alves F. (2012). New records of cetacean species for Madeira Archipelago with an updated checklist. *Boletim do Museu Municipal do Funchal*, 62: 25-43.

Gil Á. (2018). Cetáceos na Zona Económica Exclusiva Continental Portuguesa: distribuição espaço-temporal e registo de novas ocorrências. Master thesis (MSc) in Ecology and Environment. Sciences Faculty of Porto University, Porto, Portugal. 99 pp.

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

Hazevoet C.J. and Wenzel W. (2000). Whales and dolphins (Mammalia, Cetacea) of the Cape Verde Islands with special reference to the Humpback Whale *Megaptera novaeangliae* (Borowski, 1871). *Contributions to Zoology* 69(3): 197-211.

Hazevoet C.J., Monteiro V., López P., Varo N., Torda G., Berrow S. and Gravanita B. (2010). Recent data on whales and dolphins (Mammalia: Cetacea) from the Cape Verde Islands, including records of four taxa new to the archipelago. *Zoologia Caboverdiana* 1: 75–99.

Heithaus M.R., Frid A., Wirsing A.J. and Worm B. (2008). Predicting ecological consequences of marine top predator declines. *Trends in Ecology & Evolution,* 23: 202-210.

Jungblut S., Nachtsheim D.A., Boos K. and Joiris C.R. (2017). Biogeography of top predators – seabirds and cetaceans – along four latitudinal transects in the Atlantic Ocean. Deep Sea Research Part II: Topical Studies in Oceanography, 141: 59-73.

Kiszka J., Van Canneyt O., Macleod K., Walker D. and Ridoux, V. (2007). Distribution, encounter rates and habitat characteristics of toothed cetaceans in the Bay of Biscay and adjacent waters from platform of opportunity data. *ICES Journal of Marine Science*, 64: 1033-1043.

Miloslavich P., Archambault P., Bax N.J., Berghe E.V., Boustany A., Brandt A., *et al.* (2016). *Extent of assessment of marine biological diversity – Chapter 35*. New York, NY: World Ocean Assessment, United Nations.

Mason E. (2009). High-resolution Modelling of the Canary Basin Oceanic Circulation. Doctorate thesis (PhD) in Oceanography, Universidad de Las Palmas de Gran Canaria, Las Palmas de Gran Canaria. 245 pp.

Morgado C., Martins A., Rosso M., Moulins A. and Tepsich P. (2017). Fin Whale Presence and Distribution in the Pelagos Sanctuary: Temporal and Spatial Variability Along 2 Fixed-Line Transects Monitored in 2009-2013. *International Journal of Marine and Environmental Sciences*, 1: 1-14.

Moura A.E., Sillero N. and Rodrigues A. (2012). Common dolphin (*Delphinus delphis*) habitat preferences using data from two platforms of opportunity. *Acta Oecológica*, 38: 24-32.

Moura A.E., Silva S.E., SPEA, Correia A.M., Sousa-Pinto I., Gil A., *et al.* (2017). In: Bencatel J., Álvares F., Moura A. E. and Barbosa, A. M. (eds.) (2017). *Atlas de Mamíferos de Portugal.* Universidade de Évora, Portugal.

Pérez-Vallazza C., Álvarez-Vázquez R., Cardon L., Pintado, C. and Hernández-Brito J. (2008). Cetacean diversity at the west coast of La Palma island (Canary Islands). *Journal of the Marine Biolical Association of the United Kingdom*, 88: 1289-1296.

Perrin W.F. and Van Waerebeek K. (2012). *The small-cetacean fauna of the west coast of Africa and Macaronesia: diversity and distribution*. Convention on the Conservation of Migratory Species of Wild Animals, Bonn, Germany, CMS Technical Series, 26, 7-17.

R Core Team (2012). *R: A Language and Environment for Statistical Computing* https://www.R-project.org/R Foundation for Statistical Computing, Vienna, Austria.

Richardson J., Wood A.G., Nei A., Nowacek D. and Moore M. (2012). Changes in distribution, relative abundance, and species composition of large whales around south georgia from opportunistic sightings: 1992 to 2011. *Endangered Species Research,* 19: 149-156 (2012).

Robineau D. and Vely M. (1998). Ces cétacés des côtes de Mauritanie (Afrique du Nordouest). Particularltés et variations spatio-temporelles de repartition: role des facteurs oceanographiques. *Revue d Ecologie (Terre Vie)*, 53.

Sala, I., Caldeira, R.M.A., Estrada-Allis, S.N., Froufe, E. and Couvelard, X. (2013). Lagrangian transport pathways in the northeast Atlantic and their environmental impact. Limnology and *Oceanography: Fluids and Environments*, 3: 40-60.

Silva M.A., Prieto R., Cascão I., Seabra M.I., Machete M., Baumgartner M.F. and Santos R.S. (2014). Spatial and temporal distribution of cetaceans in the mid-Atlantic waters around the Azores. *Marine Biology Research*, 10:123-137.

Tepsich P., Rosso M., Halpin P.N. and Moulins A. (2014). Habitat preferences of two deep-diving cetacean species in the northern Ligurian Sea. *Marine Ecology Progress Series*, 508: 247-260.

Thomas L., Buckland S.T., Burnham K.P., Anderson D.R., Laake J., Borchers D.L. and Strindberg S. (2006). Distance Sampling. *Encyclopedia of Environmetrics,* 1: 544-552.

Tobeña M., Prieto R., Machete M. and Silva M.A. (2016). Modeling the potential distribution and richness of cetaceans in the Azores from fisheries observer program data. *Frontiers in Marine Science*, 3: 202.

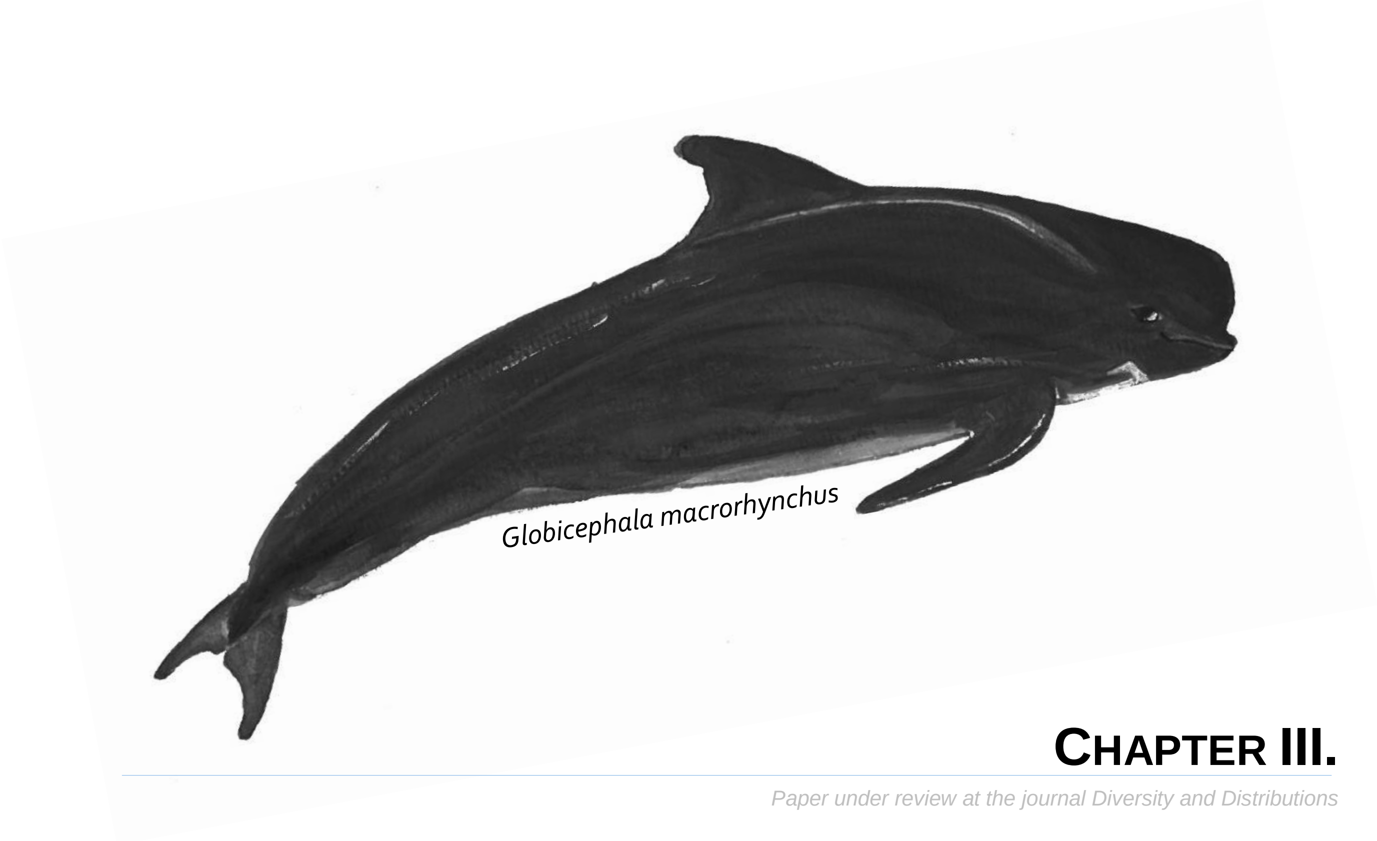
Valente R., Correia A.M., Gil Á., González L.G. and Sousa-Pinto I. (2019). Baleen whales in Macaronesia: occurrence patterns revealed through a bibliographic review. *Mammal Review*, 49(2): 129-151.

Viddi F.A., Hucke-Gaete R., Torres-Florez J.P. and Ribeiro S. (2010). Spatial and seasonal variability in cetacean distribution in the fjords of northern Patagonia, Chile. *ICES Journal of Marine Science*, 67: 959-970.

Williams R., Hedley S.L. and Hammond P.S. (2006). Modeling distribution and abundance of antarctic baleen whales using ships of opportunity. *Ecology and Society,* 11.

Weir C.R. and Pierce G.J. (2013). A review of the human activities impacting cetaceans in the eastern tropical Atlantic. *Mammal Review*, 43: 258-274.

Data Citation 1: Correia, A.M., Gandra, M., Liberal, M., Valente, R., Gil, A., Rosso, M., Pierce, G.J. and Sousa-Pinto, I, CIIMAR - UP. (2019). CETUS: Cetacean monitoring surveys in the Eastern North Atlantic. Marine Data Archive. https://doi.org/10.14284/350



CHAPTER III.

Paper under review at the journal Diversity and Distributions

3 Distribution of cetacean species at a large scale – connecting continents with the Macaronesian archipelagos in the eastern North Atlantic

Ana M. Correia[1,2], Ágatha Gil[1,2], Raul Valente[1,2], Massimiliano Rosso[1,3], Isabel Sousa-Pinto[1,2], Graham J. Pierce[4,5,6]

[1]Interdisciplinary Centre of Marine and Environmental Research (CIIMAR). 4450-208 Matosinhos, Portugal.
[2]Department of Biology, Faculty of Sciences, University of Porto (FCUP). 4169-007 Porto, Portugal.
[3]CIMA Research Foundation. 17100 Savona, Italy.
[4]Instituto de Investigacións Mariñas (CSIC). 36208 Vigo, Pontevedra, Spain.
[5]Oceanlab, University of Aberdeen. AB41 6AA, UK.
[6]CESAM and Department of Biology, University of Aveiro. 3810-193 Aveiro, Portugal.

3.1 Abstract

Aim: To describe distribution patterns and species richness of cetaceans along a wide geographical range using occurrence data from poorly studied oceanic areas, standardized to account for survey effort. Specific objectives were to compare species richness and relative abundances among sub-regions and distribution niches.

Location: Eastern North Atlantic

Time period: 2012-2017

Major taxa studied: Cetacea

Methods: Cetacean monitoring was performed by dedicated observers from cargo ships, used as platforms of opportunity, along routes between Iberian Peninsula, Macaronesia and northwestern Africa. We analyse spatial distribution of relative abundance (encounter rates), survey effort and species richness. We examined the dependence of the number of sightings and species richness on survey effort. The area was divided in sub-regions (according to the Exclusive Economic Zones and International waters) and relative abundances of the eight most frequently sighted species, as well as species richness, were compared among them. In addition, we describe niches in relation to sea depth, distance to coast, latitude and longitude.

Results: A total of 1989 sightings was logged and 26 cetacean species were identified. Species richness and relative abundances of the eight most common species differed substantially between sub-regions. Common and bottlenose dolphins distributed in shallow coastal waters contrasting with the oceanic distribution of Stenella dolphins. Cuvier's beaked whale and minke whale shared similar niches. Pilot and sperm whales were distributed in southern waters.

Main conclusions: A considerable amount of survey effort (about 3000 km per 100 km2 from this type of platforms) was needed to attain reliable estimates of species richness. Hence in less surveyed areas, species richness is likely to be underestimated. The offshore waters presented high species richness and several hotspots of cetacean abundance. This work provides new knowledge on cetacean distribution at a large scale in the eastern North Atlantic, relevant to future conservation management.

Keywords: Cetaceans, CETUS Project, distribution patterns, effort-based data, habitat range, high seas, relative abundances.

3.2 Introduction

Knowledge on distribution patterns of marine species is essential for efficient marine management and biodiversity conservation. While some areas are well-surveyed in space and time, the vast majority of the ocean is still lacking baseline data or is insufficiently surveyed to permit a good level of understanding of species diversity and distribution patterns. Hence, identifying priority areas with knowledge gaps, where monitoring efforts are required, is essential (Kaschner *et al.*, 2012; Mannocci *et al.*, 2018).

One of the priorities for research and monitoring in relation to marine conservation is the assessment of the distribution of pelagic top predators (Boyd *et al.*, 2007; Heithaus *et al.*, 2008; Parsons, 2016; Hazen *et al.*, 2019). These are key species for the maintenance of the structure and functioning of marine ecosystems (Sergio *et al.,* 2006; Sergio *et al.,* 2008). Often, knowledge on their range of occurrence is lacking or insufficient, as the range is frequently very wide and includes oceanic waters where research campaigns are expensive and logistically challenging (Alves *et al.*, 2018a; Correia *et al.,* 2015; Kiszka *et al.,* 2007; Moura *et al.,* 2012; Tobeña *et al.,* 2010).

Observation platforms of opportunity (OPOs) have been widely used to monitor cetacean presence, allowing the sampling of remote areas, such as the high seas, and monitoring over long-periods of time. This methodology has limitations, e.g. heterogeneous effort conditioned by the routes, schedules and logistics of the platform of opportunity, usually with a restricted spatial coverage of the study area. However, it is frequently the most cost-effective method to generate baseline data, allowing the collection of valuable data that would otherwise be difficult or impossible to obtain (Aïssi *et al.,* 2015; Alves *et al.,* 2018; Correia *et al.,* 2015; Evans & Hammond, 2004; Kiszka *et al.,* 2007; Morgado *et al.,* 2017; Moura *et al.,* 2012; Tobeña *et al.,* 2016; Viddi *et al.,* 2010).

The CETUS Project is a cetacean monitoring programme that records cetacean species occurrence in the eastern North Atlantic (ENA). Since 2012, cargo ships from a Portuguese maritime transport company, TRANSINSULAR, have been used as OPOs. On-board observers are trained in cetacean detection and identification, and do not have other duties. Moreover, data collected are effort-based as the survey effort is also recorded, which is fundamental to provide reliable relative abundance estimates and unbiased information on distribution, especially when effort is highly heterogeneous and survey activity is conditioned by the weather (Correia *et al.,* 2015; Evans & Hammond, 2004). The project has resulted in a large dataset of cetacean occurrence records within a vast area of the ENA (Correia *et al.,* 2019a).

The ENA is a topographically and oceanographically complex system (Caldeira & Sangrà, 2012; Mason, 2009; Sala *et al.,* 2013), with high cetacean diversity recorded both in Macaronesian islands and in the open ocean. In total, 17 species have been recorded along the continental Portuguese coast (Brito & Sousa, 2011; Moura *et al.,* 2017), 17 in the northwest Spain (Abollo *et al.,* 1998; Covelo *et al.,* 2016; Goetz *et al.,* 2015; Díaz López & Methion, 2019), 26 in Madeira (Alves *et al.,* 2018a; Freitas *et al.,* 2012), 28 in Azores (Silva *et al.,* 2014; Tobeña *et al.,* 2016), 28 in the Canary Islands (Carrillo *et al.,* 2010; Pérez-Vallazza *et al.,* 2008), 24 in Cape Verde (Hazevoet & Wenzel, 2000; Hazevoet *et al.,* 2010) and 36 along the northwestern African coast (Djiba *et al.,* 2015; Perrin & Van Waerebeek, 2012; Robineau & Vely, 1998; Weir & Pierce, 2013). The wide latitudinal and longitudinal range as well as the long temporal frame covered by the CETUS surveys, combined with the habitat variability and cetacean diversity in the area, offer the potential to analyse distribution patterns at a large scale.

We provide a descriptive analysis of spatial and temporal patterns in cetacean distribution and species richness, using effort-based data collected along cruises within the ENA, from 2012 to 2017, with high survey effort performed in open ocean. Moreover, we identified areas with the highest relative abundance and species richness, which may be priority areas for future research and conservation efforts, and compared species richness and relative abundances among sub-regions (Exclusive Economic Zones (EEZ) of Iberian Peninsula, Azores, Madeira, Canaries, Cape Verde, northwestern Africa, and International waters).

3.3 Material and Methods

3.3.1 Study area

The Canary Basin, an oceanic region located in the ENA, is characterized by a complex geography, including the existence of several archipelagos (Azores, Madeira, Canaries, and Cape Verde) that emerge from deep waters, structures such as seamounts, and a rugged coastline along the continents of Europe and Africa. It is also a very dynamic region affected by several important oceanographic features, including the North Atlantic subtropical gyre, bounded by the Azores Front that separates the anticyclonic eastern subtropical gyre from the northern cyclonic subpolar gyre, and the Cape Verde Frontal Zone that separates the nutrient-rich South Atlantic Central Waters from the cooler North Atlantic Central Waters (Zenk *et al.*, 1991). In addition, north-easterly trade winds help maintain the strong upwelling system in northwest Africa, one of the major Eastern Boundary Upwelling Systems (EBUS) of the world (Mason, 2009). These biologically productive marine regions cover less than 1% of the world's ocean but support up to 20% of the world's capture fisheries (Pauly, Christensen, 1995)

The transects sampled within CETUS cross a broad range of ocean habitats, including different topographic systems (continental platform, abyssal plains, steep slope, seamounts and canyons) and a diversity of oceanographic features, including four major currents (Portugal, Azores, Canary and Mauritania currents) and several mesoscale eddies (Mason, 2009).

To analyse cetacean occurrence by sub-regions within the area, we defined the spatial limits for each sub-region of analysis based on the EEZs (Iberian Peninsula, Azores, Madeira, Canaries, Cape Verde, northwestern Africa), delimiting also the International waters (Figure 1).

3.3.2 Data collection

Cetacean occurrence data were collected within the CETUS Project, a cetacean monitoring programme within the ENA. Dedicated and trained observers followed a standardised protocol for cetacean monitoring in line-transect surveys, aboard cargo ships from TRANSINSULAR (Correia *et al.*, 2015; Correia *et al.*, 2019b), which are used as OPOs. The company operates routes for cargo transport between Continental Portugal and Macaronesian archipelagos, with stop-overs in the northwest Africa. Between 2012 and 2017, three routes were monitored: Continental Portugal to Madeira (starting in 2012, hereafter Madeira route), Azores (starting in 2014, hereafter Azores route) and Cape Verde (with stopovers in Canary Islands, Mauritania and Senegal – starting in 2015, hereafter Cape Verde route). On two occasions in 2016, the Cape Verde route included a transect to the northwest of Spain, although the track was only crossed on effort only once, due to weather conditions). Each ship followed one of these routes and accommodated two MMOs. Observers stood on the wings of the navigation bridge (at an approximate height of 15 m, measured from the sea-level, considering maximum draught) looking for cetacean presence, from sunrise to sunset. Monitoring was performed from the front of the vessel with a field of view covering 180º. Normally, the two MMOs each covered 90º, from opposite sides of the vessel. When one MMO was resting, as detailed below, the lone MMO covered 180º. MMOs switched side every hour to reduce fatigue. Monitor was performed by naked eye and binoculars (7 x 50 mm, fitted with a scale and compass) were used for occasional scans (approximately every 5 minutes) and to support the collection of the data (e.g., to detect vessels and for species identification). Survey effort stopped at sea-state or wind-state higher than 4 (on the Douglas and Beaufort scales, respectively), when visibility was lower than 1 km, during heavy rain, and whenever observers were not allowed in the navigation bridge (e.g., during manoeuvres, safety drills or cleaning of the deck). MMOs rested in turns for an hour each at mealtimes (lunch and dinner), and optionally for additional periods of approximately 40 minutes (in the morning and in the afternoon). Sightings collected off effort (i.e. when survey effort had to stop for any of the aforementioned reasons) were considered to be opportunistic and were not included in the present analysis. Weather state was assessed at the beginning and end of the survey leg (defined as a continuous period of sampling, usually a day from sunrise to sunset), or whenever it changed significantly. Also, the number of vessels, by size category (small, medium or large), visible over 360 degrees around the observation stand, was registered at the beginning and end of the survey leg, every hour and following each sighting. Whenever a cetacean was spotted, if possible, the species was identified, and observers registered the distance and angle in relation to the ship (with the scale and compass of the binoculars), the number of animals within the group, their reaction (if any) to the ship and direction of travel. For group size, due to the occasional difficulty in determining the exact number of animals, the minimum and maximum numbers as well as a best estimate (based on the observers' perspective) were recorded. Sightings of other top predators (such as turtles, sharks and sunfishes) were also registered. The route was recorded using a tablet with an inbuilt GPS (points along the track were automatically added, with settings defined for records every 10 seconds or every 50 metres) and all the waypoints were marked. In the data analysis, the GPS position of the ship at the moment of the sighting was used, as well as the best estimate for the group size.

3.3.3 Data analysis

To assess relative abundances, encounter rates were calculated as the number of cetacean sightings (the all species total and by species) recorded on effort per 100 km. Yearly and monthly information on total effort, number of sightings, overall encounter rates and number of species, as well as encounter rates for each species by year and by each of the defined sub-regions, are provided in supplementary material (Supplementary file 1 and 2).

The overall encounter rate for the all the species total (total sightings of cetaceans on effort per 100 km)) was calculated for every cell in a grid of 100 x 100 km, as well as the total effort and total number of species identified (at least to the genus level). This was done for the whole study period over the surveyed calendar months (February and March, May to December). The 100 km grid was chosen after testing different spatial resolutions: this resolution provided a suitable sample size for statistical analysis, allowed the identification of broad scale patterns and was suitable for data visualization while also avoiding zero inflation. Distance surveyed on effort was calculated based on the tracks recorded by the GPS, by transforming the set of on effort points along the track into lines (the effort tracks) and measuring the distance covered by those lines.

In order to provide an indication of the adequacy of the current level of search effort, we used Generalized additive models (GAMs) (Hastie & Tibshirani, 1990) to model number of sightings and number of species in relation to effort, taking into account depth and distance to coast. These two environmental variables had strong effects on the distribution of the eight most sighted species as revealed by the Principal Component Analysis (PCA) described below.

All variables were estimated for all cells with non-zero effort in the 100 x 100 km grid, with effort being the total distance surveyed within the cells (kilometres surveyed by 100 km^2) and habitat variables extracted to the position of the centroid of the cell. For the GAM models, and considering that the response variables were counts, we first tested the Poisson distribution (with a log link function). We then checked for overdispersion. Dispersion was adequate for the number of species model (0.92) but there was overdispersion for the number of sightings model (2.56). As such, for the latter, we fitted a negative binomial distribution (with a log link function). The smoothers obtained essentially depict rarefaction curves.

Before fitting the models, we checked for correlation between explanatory variables (effort, depth and distance to coast) through Pearson correlation to exclude highly correlated variables from the same model (threshold of 0.75) (after Marubini *et al.*, 2009). Moreover, we assessed multiple correlation among explanatory variables through the Variance Inflation Factor (VIF, with a threshold of 3) (Zuur *et al.*, 2010). All Pearson correlations and VIF values were lower than the thresholds, so no variables were removed.

Model fitting started by including the three explanatory variables, considering only main effects, followed by backwards selection (Quian, 2009; Correia *et al.*, 2015; Correia *et al.*, 2019b). Best models were chosen by using the Akaike Information Criterion (AIC) as a measure of goodness of fit and at each step of model fitting, we kept the model with the lowest AIC value, i.e. comparing between models that differed in one explanatory variable (after removing the least significant one). When AIC values differed in less than 2, a Chi-squared test was applied. If AIC differences were not statistically significant (based on δAIC >2 or the chi-square test result), the simplest model was kept (following the principle of parsimony, e.g. Burnham & Anderson, 2002).

We verified that there was no influential data points or relationship between model residuals and the explanatory variables in the final best models. We then evaluated the models using two random subsets of data: fitting and evaluating sets (75% and 25% of the data, respectively). Prediction power of the models was determined using the Area Under the Curve (AUC) of the Receiving Operator Characteristic (ROC) plot (Beck & Shultz, 1986; Liu *et al.*, 2005). Random models have an AUC equal to 0.5; the closer an AUC is to 1, the higher discriminatory power of the model. Finally, we plotted the back transformed predicted values of number of sightings and number of species for the surveyed cells in the 100 x 100 km grid against the explanatory variables used in the final best models to confirm the relationships.

Models were developed using the ‘mgcv’ package in R 3.4.4. (R Core Team, 2018) with R Studio.

Those species (or genera) with a number of sightings higher than 30 were selected for further analysis, namely: common dolphin (*Delphinus delphis* Linnaeus, 1758), Atlantic spotted dolphin (*Stenella frontalis* Cuvier, 1829), striped dolphin (*Stenella coeruleoalba* Meyen, 1833), sperm whale (*Physeter macrocephalus* Flower, 1864), bottlenose dolphin (*Tursiops truncatus* Montagu, 1821), minke whale (*Balaenoptera acutorostrata* Lacépède, 1804), Cuvier’s beaked whale (*Ziphius cavirostris* Cuvier, 1823) and pilot whales (Globicephala sp. Lesson, 1828). Considering these taxa, the cetacean community composition (both relative abundances and percentage relative contribution), as well as the monthly presences were represented for each previously defined sub-region. Maps of sightings distribution along tracks were created for these eight taxa and are presented in supplementary material (Supplementary file 3).

To analyse the species niches according to their geographical distribution and coastal or oceanic occurrence, we considered four niche factors: depth, distance to coast, latitude and longitude. To delimit and characterize the surveyed area, a set of points was created, with a point generated every 5 km within effort tracks (Correia *et al.*, 2015). The niche factors were extracted to this set of points. Depth was obtained from bathymetry data in GEBCO (GEBCO, 2017), and distance to coast was calculated using ArcGIS 10.5 (ESRI, 2016). Summary statistics were calculated for the group size of each species, as well as for the niche factors at the position of the sightings. Quantiles of the distributions for each factor are presented for each species (see Supplementary file 4) and illustrated with boxplots.

We then followed the Fernández *et al.* (2013) approach and run a PCA. Prior to PCA, we verified Pearson correlations among niche factors. Considering a 0.75 threshold, depth and distance to coast were correlated (0.83). However, PCA analysis is commonly used to deal with correlated explanatory variables, hence both niche factors were included in the analysis. We performed a PCA which requires that all variables are on a comparable scale. Thus, we first standardized the data by subtracting the mean value of each variable for all data points and dividing by the standard deviation. Then, for the most important Principal Components (PC) (those that together account for more than 75% of the total accumulated variation explained), we used boxplot graphs to represent the quartiles of the PC scores (minimum, 25%, median, 75% and maximum values) for the eight most frequently sighted species. For overall and pairwise comparisons of the PC scores among species, we used Kruskal-Wallis and Mann-Whitney tests, respectively (test results are given in Supplementary file 5). To avoid type I errors on multiple pairwise comparisons, we applied the Bonferroni correction to the significance level accepted for the Mann-Whitney tests. Thus, we calculated the new significance level as 0.05 divided by the number of comparisons (28 comparisons that result on a significance level of 0.002).

Maps were created in ArcGIS 10.5 (ESRI, 2016) using a Mercator projection (EPSG: 4326), graphs in Microsoft Excel 2016 and statistical tests and boxplots were carried out using R Studio (R Development Core Team, 2012).

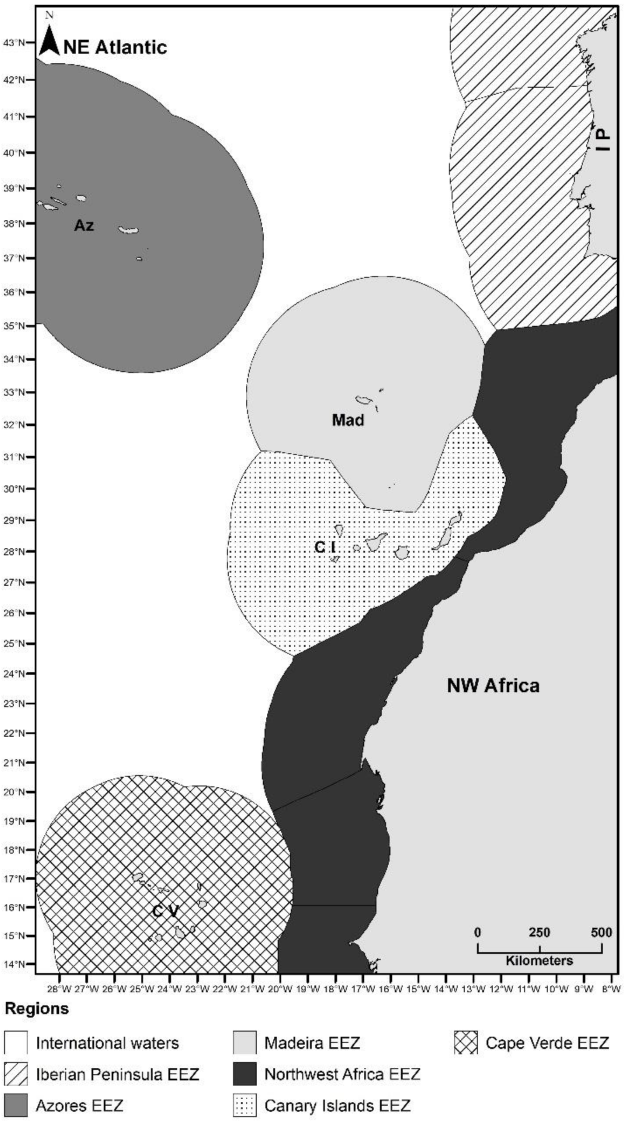


Figure 1. Sub-regions for the analysis, considering the limits of the Exclusive Economic Zones in the study area. EEZ – Exclusive Economic Zone; IP – Iberian Peninsula; Az – Azores archipelago; Mad – Madeira archipelago; CI – Canary Islands archipelago; CV – Cape Verde archipelago; NW – Northwest.

3.4 Results

3.4.1 Overall distribution of effort, encounter rates and species richness

Survey effort was concentrated in summer and early autumn (July to October), while other months (February, March, May, June, November and December) were surveyed in only one of the years and there was no survey effort in January or April. Yearly effort increased from 2012 to 2015, due to an increase in the number of routes being monitored: one route between 2012 and 2014, two from 2014 and 2015, and three from 2015 and 2017. Effort, number of sightings, encounter rates and number of species, all presented a high inter-annual variability (Supplementary file 1).

In total, 124 428 km were surveyed in the study area and 26 cetacean taxa were identified at least to the genus level, with 1989 sightings collected on effort, resulting in an overall encounter rate of 1.60 sightings per 100 km. Overall, the 8 most frequently sighted species contributed 45% of the sightings. These were: common dolphin, spotted dolphin, striped dolphin, sperm whale, bottlenose dolphin, mink whale, Cuvier's beaked whale and pilot whales. With 262 occurrences, common dolphins were the most frequently sighted species, comprising 12.9% of the sightings. This was also the species with the highest number of individuals recorded during a single sighting, with a record of 2500 animals in a group seen off Dakar (Senegal) on the 26th of July 2015 (Table 1 and Supplementary file 2).

Table 1. Summary table for the most frequently sighted species.

			Group size	
Taxa	**No. sightings**	**ER**	**Range**	**Mean ± SD**
Delphinus delphis	262	0.206	1-2500	27.44 ± 160.82
Stenella frontalis	167	0.131	1-130	19.31 ± 19.76
Stenella coeruleoalba	119	0.093	1-150	19.60 ± 21.85
Physeter macrocephalus	116	0.092	1-20	2.01 ± 2.17
Tursiops truncatus	92	0.071	1-130	10.44 ± 15.37
Balaenoptera acutorostrata	75	0.059	1-4	1.36 ± 0.65
Ziphius cavirostris	51	0.023	1-7	2.08 ± 1.50
Globicephala sp.	44	0.041	1-100	19.03 ± 21.45

For group size estimates, sightings with associated species are not considered as the number of animals assessed during surveys corresponds to the mixed group, hence was not representative of a single species. ER – Encounter rate (number of sightings per 100 km surveyed); SD – Standard deviation.

In general, the areas with the highest survey effort were in offshore waters between Continental Portugal and Madeira and Azores, where a high diversity of species (up to 11 species per 100 km^2) was observed (Figure 2). The highest encounter rates were registered elsewhere: e.g., close to continental Portugal and west Africa, and near the Macaronesian archipelagos (Figure 2).

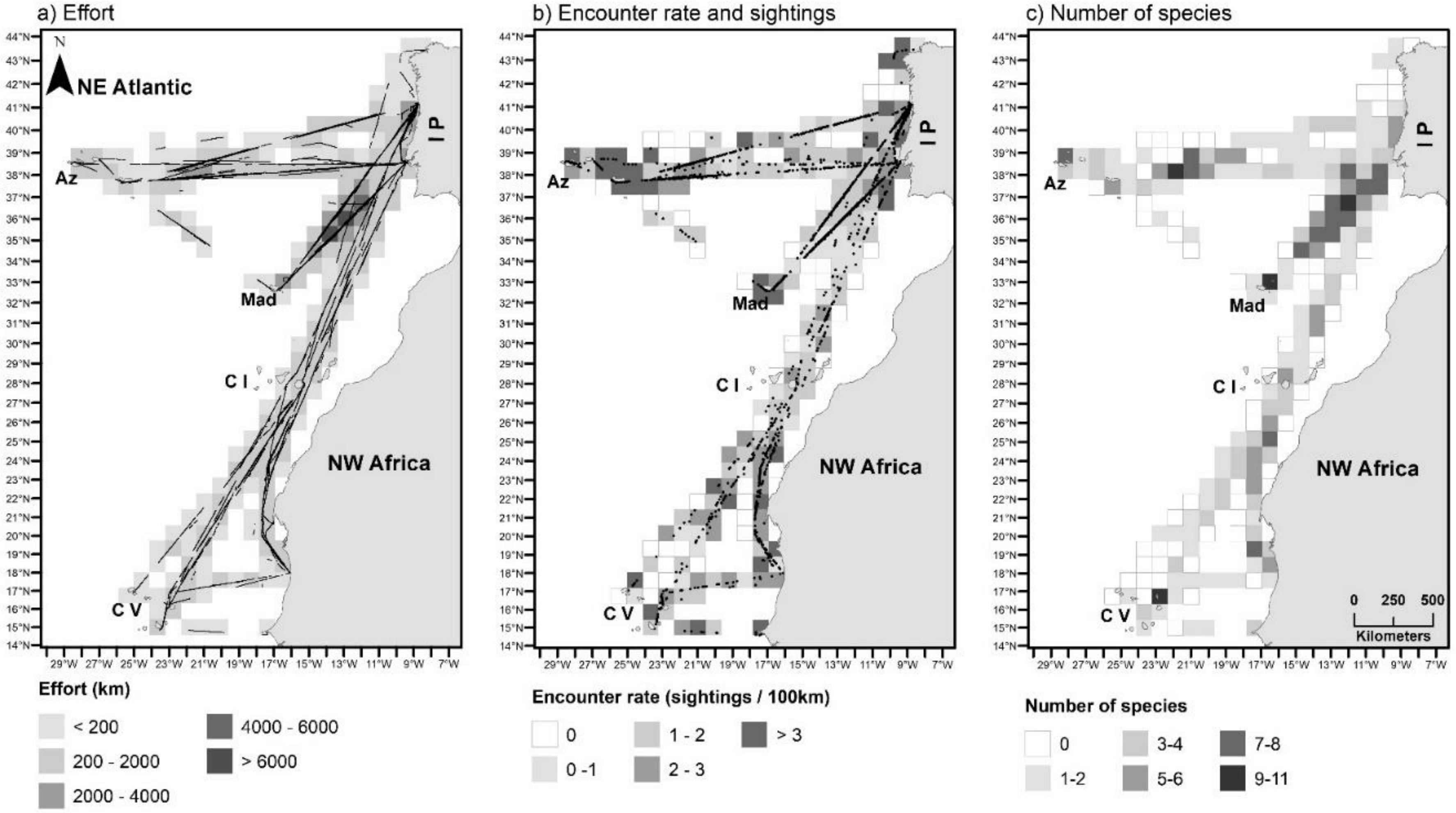
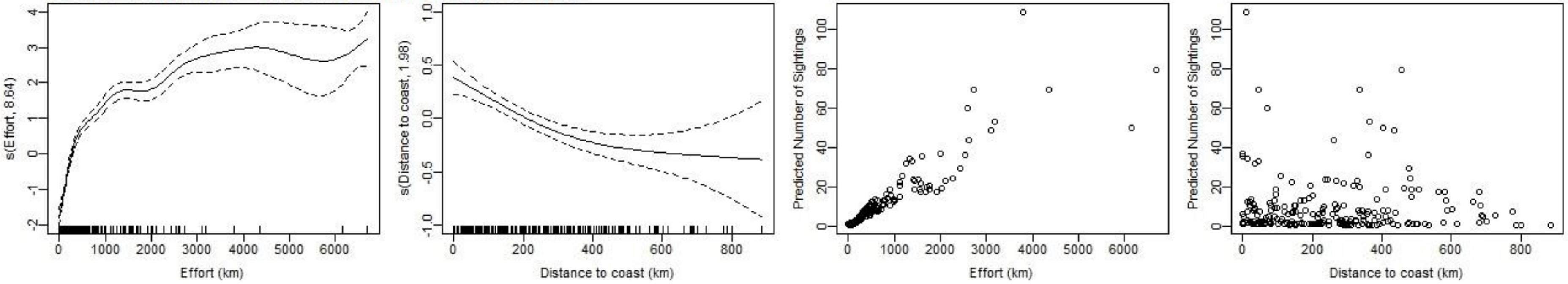


Figure 2. Spatial distribution of survey effort, encounter rate and number of species seen. a) Black lines represent effort tracks. b) Black dots represent sightings. Grid: 100 x 100 km. IP – Iberian Peninsula; Az – Azores archipelago; Mad – Madeira archipelago; CI – Canary Islands archipelago; CV – Cape Verde archipelago; NW – Northwest.

The number of sightings and species seen were highly influenced by survey effort. Number of sightings per grid cell generally increases with survey effort and decreases with distance to coast. Number of species per grid cell increased with effort up to around 3000 km per 100 km^2, after which they started to stabilize Species richness peaks at approximately 1500 m of depth. Confidence intervals are wide for the smoothers at high values of effort and distance to coast and for the models AUCs (Figure 3).

Number of Sightings ~ Effort + Distance to coast
P (Effort) < 0.001 | P (Distance to coast) < 0.001
Deviance explained = 83.9% | AUC (CI) = 0.60 (0.12 - 1.00)



Number of Species ~ Effort + Depth
P (Effort) < 0.001 | P (Depth) = 0.06
Deviance explained = 67.3% | AUC (CI) = 0.60 (0.35 - 0.84)

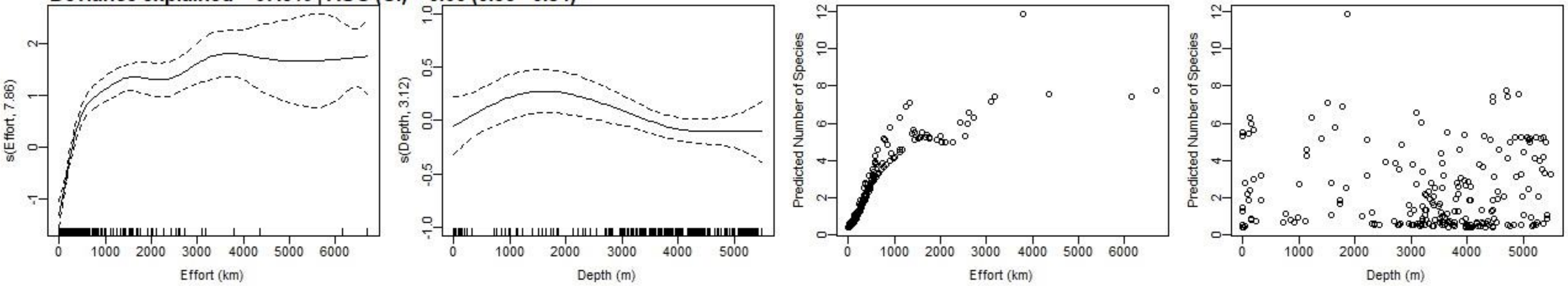


Figure 3. GAM models: fitted smoothers and predicted values for (upper panel) number of sightings (per grid cell) versus survey effort and distance to the coast and (lower panel) number of species seen (per grid cell) versus survey effort and water depth. Totals for each variable were calculated over the whole study period for each cell within a grid of 100 x 100 km cells.

3.4.2 Analysis of the cetacean community composition by sub-region

In all sub-regions, the sightings of the eight most frequently sighted species make up 40% to 50% of total sightings, except in the Cape Verde EEZ, where sightings for other taxa represented about 74% of the total sightings. The cetacean community varied greatly among sub-regions. In the EEZs of the Iberian Peninsula and Azores, the most frequently encountered species was the common dolphin. However, while in Iberian waters the number of common dolphin sightings represented about half of the total sightings; in Azores, sightings of this species make up only around a quarter of the total for the most frequently sighted species. The Atlantic spotted dolphin was the most frequently sighted species in the Canary Islands EEZ, in the Madeira EEZ and in international waters. Its contribution to the total number of sightings was highest in the Canary Islands. Here, the sperm whale was the second most frequently encountered species. Sperm whales were the most frequently sighted species in the EEZs of the northwestern Africa. The encounter rate for pilot whales was highest in the Cape Verde EEZ, where they were the most frequently sighted species. The second highest encounter rate for this species was recorded in the EEZs of northwestern Africa. Presence of the species was minimal in the remaining sub-regions and they were never sighted in international water (Figure 4).

Taxa	IP EEZ	AZ EEZ	MAD EEZ	NWA EEZ	CI EEZ	CV EEZ	IW
Total ER	1.622	2.639	1.208	2.045	1.440	1.533	1.123
No. of species	19	14	12	21	12	14	16
Total effort (km)	44955	11938	18794	14282	5901	4957	23600

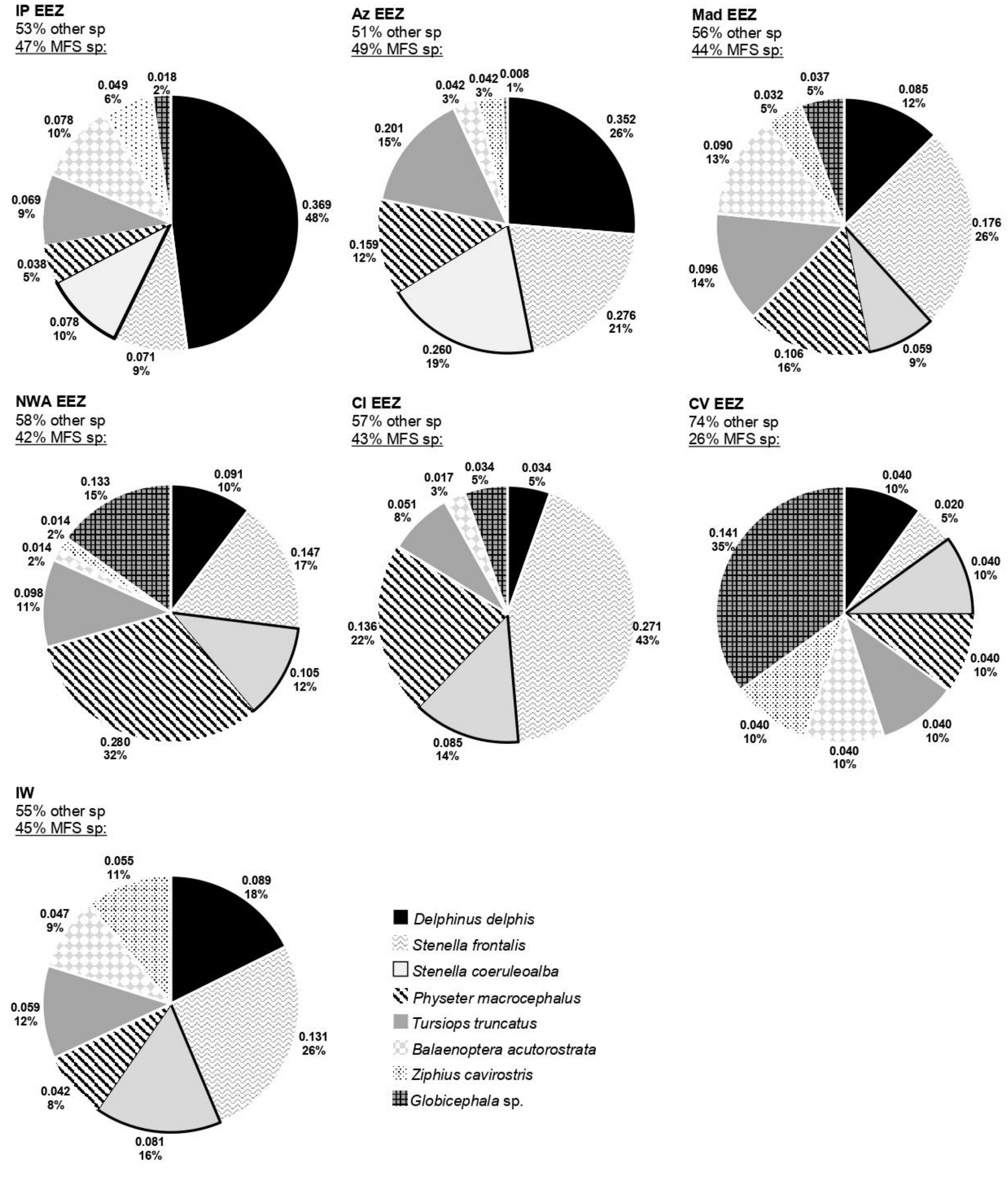


Figure 4. Cetacean community composition in each sub-region defined, highlighting encounter rates and percentage relative contribution for the eight most frequently sighted species. Pie charts illustrate the encounter rates and percentage of contribution of the most frequently sighted species (identified, at least, to the genus level) for each sub-region (defined in Figure 3). Occurrences with associated species were used to calculated the encounter rate of both taxa only if at least one of the taxa sighted was amongst the eight most frequently sighted species over the whole study area. ER – encounter rate (sightings per 100 km); sp – species; MFS – Most frequently sighted; EEZ – Exclusive Economic Zone; IP – Iberian Peninsula; Az – Azores archipelago; Mad – Madeira archipelago; NWA – Northwest Africa; CI – Canary Islands archipelago; CV – Cape Verde archipelago; IW – International waters.

The EEZs of the northwestern Africa had the highest number of species registered (21) and encounter rates of 11 out of these 21 species where the highest across all sub-regions. The highest overall encounter rate was registered in the Azores EEZ. In international waters, 16 species were recorded, and the overall encounter rate was about 1.12 sightings / 100 km. Almost 20% of the survey effort was undertaken within these waters (Figure 4 and Supplementary file 2).

Regarding temporal patterns, six of the eight most frequently sighted species were seen in international waters every month from July to October but were not seen outside this period. Of the two exceptions, sperm whales differed in that they were absent in September while pilot whales were never seen in international waters. In the Canary Islands and Cape Verde, effort was very low, and the presence of most species was restricted to few months, but the occurrence of the Atlantic spotted dolphin from June to November in the Canary Islands is noteworthy (Figure 5).

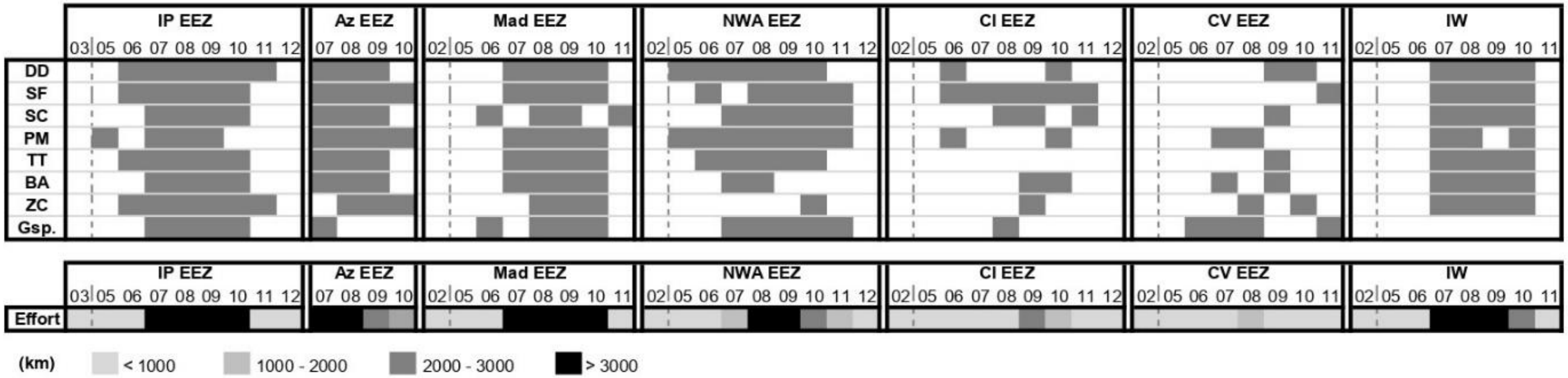


Figure 5. Temporal presence of the most frequently sighted species by sub-regions (defined in Figure 3). The occurrences where the species were associated with other taxa were considered. DD – Delphinus delphis; SF – Stenella frontalis; SC – Stenella coeruleoalba; PM – Physeter macrocephalus; TT – Tursiops truncatus; BA – Balaenoptera acutorostrata; ZC – Ziphius cavirostris; Gsp. – Globicephala sp. EEZ – Exclusive Economic Zone; IP – Iberian Peninsula; Az – Azores archipelago; Mad – Madeira archipelago; NWA – Northwest Africa; CI – Canary Islands archipelago; CV – Cape Verde archipelago; IW – International waters. For each sub-region, only surveyed months are shown. There were no surveys in January or April in any sub-region..

3.4.3 Habitat niches of the most frequently sighted species

Surveys covered a wide range of habitat characteristics in the area, and the most frequently surveyed areas were in deeper waters, at distances from the coast of up to 871 km, in northern latitudes and at longitudes ranging from 28.62º W to 8.33º W. Common dolphin was the species seen in the shallowest waters, closest to the coast, in northern and eastern areas (Figure 6 and Supplementary file 4).

In the PCA analysis, the first two PCs together explained of 78.3% of variation. The variables that contributed the most to PC1 were depth and distance to coast, while PC2 was mainly related to the geographical variables (latitude and longitude). Species with higher PC1 scores are found in deeper waters and further from the coast and the species with higher PC2 scores occur more in northern and eastern regions of the study area (Table 2).

Table 2. PCA results for niche factors of the most frequently sighted species.

	PC1	PC2	PC3	PC4
Depth	46.193	3.207	1.022	49.578
Distance to coast	46.852	2.498	0.342	50.309
Latitude	5.973	41.835	52.008	0.104
Longitude	0.983	52.460	46.548	0.009
Eigenvalue	1.880	1.250	0.700	0.170
Percentage of variation explained	47.009	31.242	17.504	4.244
Accumulated % variation explained	47.009	78.252	95.756	100.000

Eigen vectors and coefficients of each niche factor are indicated for each principal component. Those occurrences records where a species was associated with other taxa were included in the analysis (i.e., as occurrence records for all species in the mixed group).

Common and bottlenose dolphins share a similar niche according to the PC1 score but are significantly different according to PC2. In fact, both species were mostly found in shallower waters closer to the coast (with the bottlenose dolphin in deeper areas compared to the common dolphin), but common dolphins are found more in northern and eastern waters. Both species of the genus Stenella sp. shared similar niches on both PCs, with statistically significant differences from bottlenose and common dolphins, and occupy deeper waters further from the coast (Figure 6 and 7 and Supplementary files 4 and 5).

Pilot and sperm whales did not present statistically significant differences in scores for either PC1 or PC2, being distributed in more southern and western areas compared to the other species. Both species were encountered relatively close to the coast. Finally, Cuvier's beaked whale and minke whales were found in the deepest waters and furthest from the coast, sharing similar ranges across the two PCs (Figures 6 and 7 and Supplementary files 4 and 5).

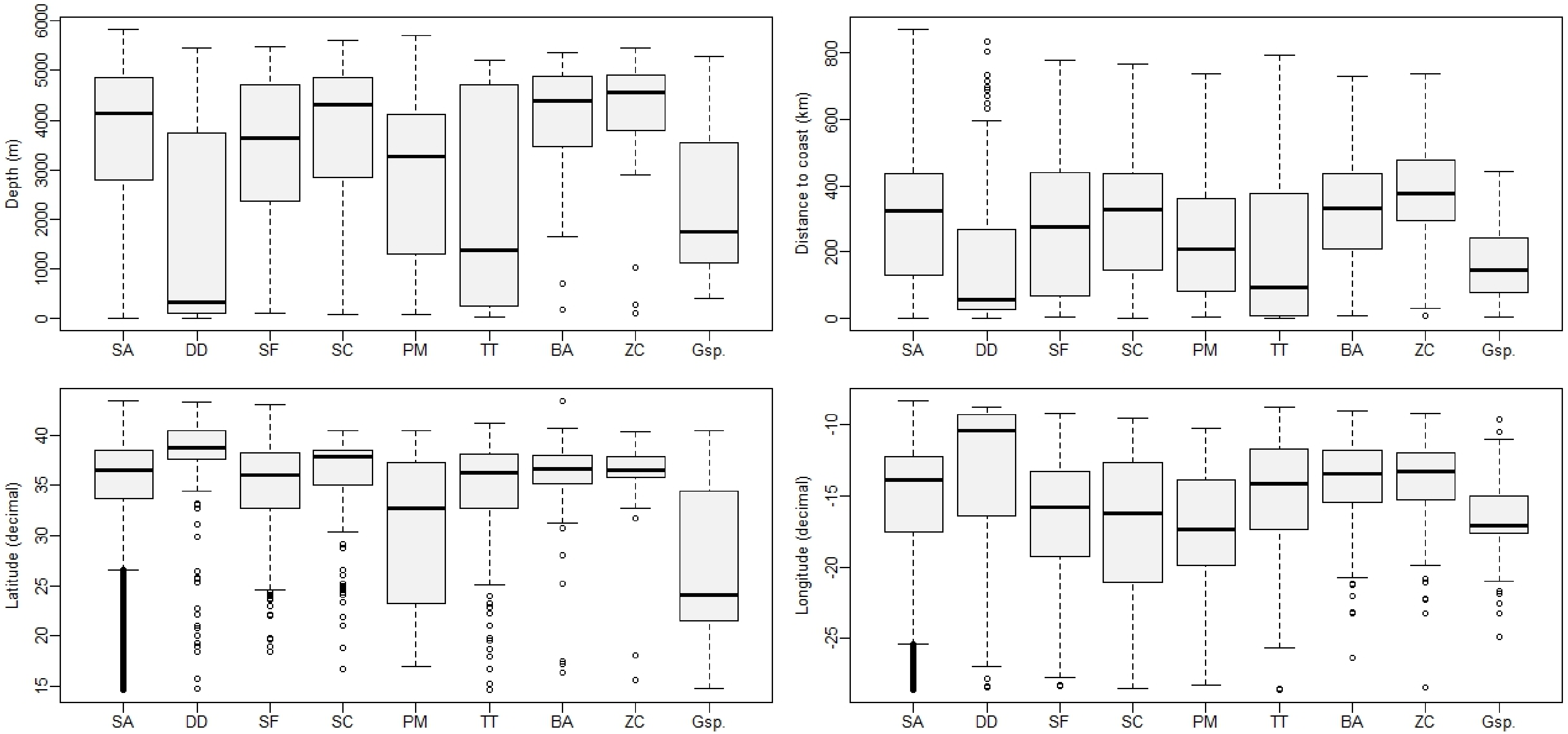


Figure 6. Boxplots of the niche factors for the eight most frequently sighted species. Values of the niche factors were extracted to the position of the occurrence records. The occurrences where the species were associated with other taxa were included in the analysis. The lower 25% and the upper 25% scores are represented by the whiskers and grey box represents the central 50% of the scores (with median represented by a black line). Outlier scores are illustrated by the circles. SA – Surveyed area; DD – *Delphinus delphis*; SF – *Stenella frontalis*; SC – *Stenella coeruleoalba*; PM – *Physeter macrocephalus*; TT – *Tursiops truncatus*; BA – *Balaenoptera acutorostrata*; ZC – *Ziphius cavirostris*; Gsp. – *Globicephala* sp.

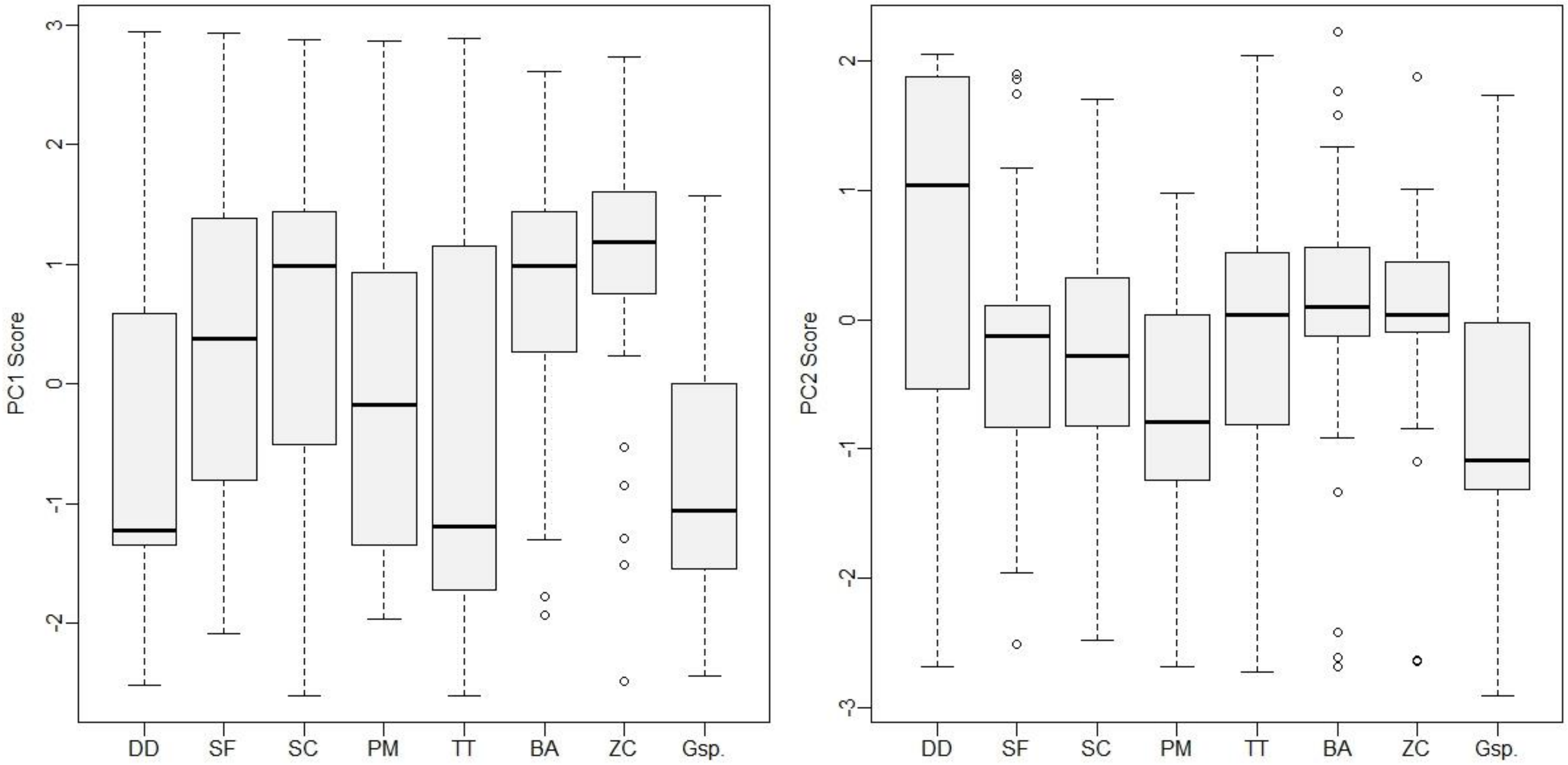


Figure 7. Boxplots of PC scores for the eight most frequently sighted species. The occurrences where the species were associated with other taxa were included in the analysis. The lower 25% and the upper 25% scores are represented by the whiskers and grey box represents the central 50% of the scores (with median represented by a black line). Outlier scores are illustrated by the circles. SA – Surveyed area; DD – *Delphinus delphis*; SF – *Stenella frontalis*; SC – *Stenella coeruleoalba*; PM – *Physeter macrocephalus*; TT – *Tursiops truncatus*; BA – *Balaenoptera acutorostrata*; ZC – *Ziphius cavirostris*; Gsp. – *Globicephala* sp.

3.5 Discussion

Results confirm the high cetacean diversity previously reported for continental shelf waters within ENA (e.g. Alves *et al.*, 2018a; Brito & Sousa, 2011; Carrillo *et al.*, 2010; Correia *et al.*, 2015; Djiba *et al.*, 2015; Freitas *et al.*, 2012; Hazevoet & Wenzel, 2000; Hazevoet *et al.*, 2010; Moura *et al.*, 2017; Perrin & Van Waerebeek, 2012; Robineau & Vely, 1998; Silva *et al.*, 2014; Tobeña *et al.*, 2016; Weir & Pierce, 2013), and show that, although sightings rates were lower beyond the continental shelf, high cetacean diversity extends into the high seas of this region.

Spatially, peaks of species richness were mostly found along the Madeira and Azores routes. However, number of sightings and number of species seen both depend on survey effort. The number of species seen reached a plateau at a high amount of effort (approximately 3000 km per 100 km^2) while, as expected, number of sightings tends to increase with effort. Results also showed that while number of sightings tends to be higher in areas closer to the coast, species richness peaks in deeper areas.

Overall, the CETUS Project has involved a sufficient amount of survey effort to provide a reliable picture of cetacean species diversity. However, number of grid cells with a high amount of effort was very low, hence a higher amount of effort across the surveyed grid cells was needed for estimated with narrower confidence intervals. In less surveyed areas around the globe, such as offshore waters, cetacean abundance and species richness is very likely to be underestimated. Higher encounter rates were registered in coastal areas, both on the continental shelves and around the islands, with the Azores EEZ being the sub-region with the highest encounter rate, followed by the EEZs of the northwestern Africa where more species were identified. These results are consistent with previous findings: the Azores archipelago is known for its high cetacean abundance and diversity (Silva *et al.*, 2014) and the coast of northwestern Africa is also a very rich area in terms of cetacean species diversity (Djiba *et al.*, 2015; Perrin & Van Waerebeek, 2012; Robineau & Vely, 1998; Weir & Pierce, 2013). It is generally recognised that cetacean abundance tends to be higher in inshore waters, for example, related to strong coastal upwelling phenomena, as in the case of the Iberian and African Atlantic coasts, and the island mass effect, whereby the topographic disturbance of oceanic flow by an island, leads to increased phytoplankton biomass proximate (Alves *et al.*, 2018; Correia *et al.*, 2015; Gove *et al.*, 2016; Moura *et al.*, 2012; Tobeña *et al.*, 2016; Viddi *et al.*, 2010). Fundamentally, the lower availability of nutrients may limit pelagic community productivity and biodiversity further offshore, while the increasing separation of seabed and photic zone limits the productivity of demersal and benthic communities in deeper waters (Mason, 2009).

During the CETUS campaigns, the transects passed through a large variety of habitat and, notably, covered substantial distances in open ocean where cetacean abundance and diversity is least well documented. Overall. 16 cetacean species were identified (at least to genus level) in the open ocean, of which the Atlantic spotted dolphin was the most frequently encountered. The encounter rate for Cuvier's beaked was highest in the open ocean

The use of OPOs in this study permitted data collection and long-term monitoring in areas that are otherwise rarely or never surveyed, especially the high-seas (Aïssi *et al.*, 2015; Alves *et al.*, 2018; Correia *et al.*, 2015; Evans & Hammond, 2004; Kiszka *et al.*, 2007; Morgado *et al.*, 2017; Moura *et al.*, 2012; Tobeña *et al.*, 2016; Viddi *et al.*, 2010). Nevertheless, it is important to acknowledge the limitations of the work. The results presented here are mostly representative of cetacean distribution from July to October. Southern areas were less represented and, consequently, the distribution of tropical species is also less represented. The monthly encounter rates, survey effort and number of species presented a high inter-annual variation, which for the last two is most likely a reflection of the effort heterogeneity both in space and time. In fact, by using OPOs, monitoring was limited by the company's schedule and routines. Surveyed routes are thin lines crossing a very wide area, with survey effort covering only a subset of the habitats in the region. Moreover, as in all marine campaigns, survey effort was also conditioned by the weather.

Only eight of the 26 species seen were recorded more than 30 times, together representing arout 45% of the total sightings. In general, cetacean species are expected to occupy distinct ecological niches, often with definable nursery areas, feeding areas and migration corridors. These niches reflect (for example) their life-history strategies, behaviour and diet, and are limited by (for example) their thermal tolerance range (MacLeod, 2009), the distribution of their preferred prey and competition with other species. Thus it is important to identify suitable habitats for each cetacean species, notably in international open ocean waters where sightings data are relatively scarce. Such habitats may be important to support a substantial part of their life cycle, not only as routes of travel and providing stopovers for feeding or resting. Knowledge of movements is important to understand connectivity and hence gene flow between, and genetic diversity of, the various populations.

Common dolphin was the most frequently encountered species as well as the species with the biggest groups recorded, group size reaching a maximum of approximately 2500 animals in an encounter off Dakar, Senegal. The species is usually found in large pods in several areas of the ENA (e.g., Alves *et al.*, 2018a; Baines & Reichelt, 2014; Camphuysen, 2000; Djiba *et al.*, 2015). Common dolphin was also the most frequently sighted species in the EEZs of the Iberian Peninsula and Azores, where it has been often reported as the most abundant cetacean species (Correia *et al.*, 2015; Paradell *et al.*, 2019; Moura *et al.*, 2012, 2017; Silva *et al.*, 2014). Although most commonly seen close to the coast in relatively shallow waters, common dolphins were also recorded offshore and in very deep waters. Previous analysis using this dataset (CETUS dataset from 2012 to 2016) showed that the species presents clear core areas of occurrence. These hotspots are apparently related with specific environmental conditions (for example, coastal colder waters related with strong coastal upwelling systems), with the species being more of an ecological specialist than a generalist (Correia *et al.*, 2019b). Further studies should focus on the presence of common dolphins in the north of continental Portugal, a poorly studied area within the range of the species along the Iberian coastline.

Bottlenose dolphins preferred shallower waters in areas closer to the coast, but they were also frequently recorded in high seas. It is generally recognised that there are both coastal and oceanic populations of this species. Preliminary analysis of data from CETUS already pointed to the presence of an oceanic population in the area between the Iberian Peninsula and Madeira (Correia *et al.*, 2015). Genetic studies have shown that resident populations in Galicia and the Sado estuary are likely to have a strong degree of genetic isolation from the populations in the archipelagos and non-resident individuals. On the other hand, a high gene flow among the Iberian archipelagos was identified (Fernández *et al.*, 2011). Transient individuals have been identified in the archipelagos of Madeira and Azores (Dinis *et al.*, 2016a; Dinis *et al.*, 2016b; Silva *et al.*, 2014); and some individuals from resident populations in Iberia Peninsula were found to undertake long distance movements (Fernández *et al.*, 2011). Bottlenose dolphins are listed under Annex II of the EU Habitats Directive (Directive 92/43/CEE), so Member States are required to designate Special Areas of Conservation for the protection of the species. Besides efforts in coastal waters to identify such areas, we need to understand the wider movements of bottlenose dolphins and assess their habitat use in high seas to identify important areas beyond the continental platform, and potentially beyond national jurisdiction. Fernández *et al.* (2011) highlighted that different populations inhabiting different areas (coastal VS oceanic) may present distinct habitat use, which should be considered when designing and implementing effective conservation measures.

In Madeira, the Canary Islands and international waters, the Atlantic spotted dolphin was the most frequently sighted species, also occurring in big groups. For Madeira, where spotted dolphin has a peak of occurrence in the summer and where average group size is highest, these results are in accordance with a previous study by Alves *et al.* (2018a). However, in the literature for the Canary Islands, although it is amongst the most frequently sighted species, the spotted dolphin is not reported as the most frequently encountered (Carrillo *et al.*, 2010; Pérez-Vallazza *et al.*, 2008). In the present study, the Atlantic spotted dolphin had a very similar niche to the one occupied by striped dolphin, both presenting characteristics of oceanic dolphin species (i.e. occurring in deeper areas further from the coast when comparing to common and bottlenose dolphins). Given the importance of both species in the cetacean community composition within the EEZs, their occurrence in international waters may indicate a high gene flow among sub-regions. The fact that these species are less frequent during winter in the Portuguese archipelagos (Alves *et al.*, 2018a; Silva *et al.*, 2014) may indicate that international waters are even more important for the species during this season.

Sperm whales have a year-round presence both in the Azores and Madeira islands (Alves *et al.*, 2018a; Silva *et al.*, 2014). During the CETUS campaigns, they were encountered from July to October in both sub-regions. Sperm whale was the most abundant cetacean in the EEZs of northwestern Africa, which is consistent with previous surveys in Mauritania (Baines & Reichelt, 2014; Camphuysen *et al.*, 2012). The IUCN global assessment determined that sperm whales are Vulnerable (www.iucnredlist.org). Northwest Africa is a hotspot area for the species, where it has an important role in ecosystem functioning (Morissette *et al.*, 2010). Several marine management issues, mostly related with inefficient management of fisheries, exist in the EEZs of northwestern Africa (Nagel & Gray, 2012). Since sperm whales seem to occupy areas closer to the coast, it is likely that their area of occupancy overlaps with areas of intensive fishing. Interaction between sperm whales and fishing activity can have negative consequences for both the animals and the economic activity (Karpouzli & Leaper, 2004; Richard *et al.*, 2017; Tixier *et al.*, 2019). Thus, there is a need to evaluate anthropogenic impacts on sperm whales in the northwest coast of Africa.

Pilot whales were the most frequently encountered cetacean taxon in the Cape Verde EEZ, where there are several reports of mass strandings of these animals (Hazevoet *et al.*, 2010). Alves *et al.* (2018b) reported on the ecological connectivity of short-finned pilot whales in Macaronesia (between Azores and Madeira and vice-versa, and between Madeira and Canaries and vice-versa). The study did not include samples from the Cape Verde archipelago. However, if the species is capable of moving between Madeira and the Azores, connectivity between the Cape Verde stock and the remaining Macaronesian archipelagos is plausible and should be investigated. In the present study, no pilot whales were seen in international waters which potentially indicates that their occurrence is mostly restricted to the EEZs (although evidently crossing areas beyond national jurisdiction for short periods of time to move between sub-regions, potentially travelling during the winter).

In general, minke whales showed a preference for very deep waters distant from the coast. The oceanic distribution of this species in the ENA, most likely related with migratory routes, is poorly known and evidently needs further investigation (Valente *et al.*, 2019; Van Waerebeek *et al.*,1999; Vikingsson & Heide-Jørgenssen, 2014). Like the minke whale, the Cuvier's beaked whale was also found in areas of deep water distant from the shore. Geographically, both were mostly distributed in western and northern areas, and rarely seen in more tropical waters.

Beaked whales are known to be sensitive to human activities such as seismic surveys, navy exercises, navigation. Indeed, evidence of mass strandings related with naval exercises has been documented (Cox *et al.*, 2006; Dolman *et al.*, 2011; MacLeod & D'Amico, 2006). Hence, data on their distribution in relation to the distribution of human activities is valuable for the conservation of the species (Cañadas *et al.*,2018; Dinis *et al.*, 2017). However, these are challenging species to study. In general, beaked whales spend only short periods of time at surface, are difficult to identify to species, and occur mostly in oceanic waters where fewer research campaigns have been undertaken (Correia *et al.*, 2015; Heyning, 1989; de Soto *et al.*, 2017; Tepsich *et al.*, 2014). The two beaked whale species which are easiest to identify are the northern bottlenose whale (*Hyperoodon ampullatus*) due to its distinct characteristics (e.g., shape of the head, size) and Cuvier's beaked whale as the adults have distinctive and easily detected colour patches (de Soto *et al.*, 2017). In fact, Cuvier's beaked whale was the species of Ziphiidae most frequently sighted during the CETUS campaigns. The status of the species worldwide, and specifically in European waters, is "Data Deficient" (www.iucnredlist.org), revealing the need for further baseline information on occurrence and distribution. The species was encountered most extensively (from June to November) in the Iberian Peninsula EEZ. However, the encounter rate was highest in international waters, highlighting the need to expand sampling efforts into areas beyond national jurisdiction in order to monitor (and ultimately conserve) Cuvier's beaked whale populations in the ENA.

Worldwide, conservation of cetacean species is mostly focused in areas of national jurisdiction and there is a huge difference in the relative extent of protected areas between waters beyond national jurisdiction and those within the EEZs (1.2% against 16.8%, UNEP-WCMC *et al.*, 2018). However, it is evident that international waters play a fundamental role in the cetacean community of the ENA, and further investigation of distribution patterns, preferred habitats and habitat use, movements, population structuring and gene flow across the entire area is needed. International agreements are fundamental to ensure cetacean conservation in international waters, to obtain baseline data, assess population status and, where necessary, design and implement management measures. The development of cost-effective monitoring programmes in high seas areas would help ensure continuity of monitoring to underpin long-term management. Solutions may rely on programmes such as CETUS Project, sophisticated multi-scale ecological niche modelling techniques, tagging programmes, new technological approaches such as monitoring through automated vehicles and cheap non-invasive techniques such as photo-ID and environmental DNA (Bohorquez *et al.*, 2019). Another challenge is dealing with the dynamism of cetacean distribution related to their life history, migration and movements, which may call for dynamic marine protected areas. This in turn requires adaptive marine management (Hooker *et al.*, 2011) and is probably not yet feasible in EU waters. Ultimately, to ensure the conservation of species, following the precautionary principle, we need to define year-round protected areas for all the core habitats of those species (even if they are only used / preferred during a specific season). Moreover, besides knowledge on occurrence, the assessment of threats (i.e., by-catch, entanglement, collision), at least in core areas of occurrence, is also essential to design specific conservation measures for effective marine management (Díaz López *et al.*, 2019).

We have to recognize the gap between monitoring and mitigation, and specifically that we cannot solve or provide solutions for all the challenges of marine management and conservation in the high seas. Effective measures in offshore waters, and specifically in areas beyond national jurisdiction, are limited by logistic and political factors (Bohorquez *et al.*, 2019). Nevertheless, the present work may be useful for the design of future dedicated campaigns, to efficiently construct a monitoring programme including both areas within the EEZs and in international waters and to support conservation and management efforts in the area. The CETUS Project is ongoing and aims to continue providing updated and reliable data, such as effort-based relative abundances, that could be used as indicators for management purposes (e.g. Marine Strategy Framework Directive), and to construct a long-term dataset. Moreover, this effort-related dataset is key to develop ecologic niche models and provide meaningful distribution maps both for the present and under future climate change scenarios in support of the European agenda for the conservation of marine ecosystems.

3.6 Acknowledgements

We thank the contribution and dedication of all the volunteers during the monitoring campaigns. We are extremely grateful to TRANSINSULAR, the cargo ship company that provided all the logistic support, and to the ships crews for their hospitality. This study was conducted within a PhD program from the Faculty of Sciences of the University of Porto, Portugal, hosted by the Centre of Marine and Environmental Research (CIIMAR – Porto, Portugal) and funded by the Portuguese national funding agency for science, research and technology (FCT) under the grant SFRH/BD/100606/2014. The CETUS Project is led by CIIMAR | University of Porto in partnership with the cargo ship company TRANSINSULAR | ETE Group.

3.7 References

Abollo E., López A., Gestal C., Benavente P. and Pascual S. (1998). Macroparasites in cetaceans stranded on the northwestern Spanish Atlantic coast. *Diseases of Aquatic Organisms*, 32: 227-231.

Aïssi M., Arcangeli A., Crosti R., Yahia M.N.D., Loussaief B., Moulins A., *et al.* (2015). Cetacean Occurrence and Spatial Distribution in the Central Mediterranean Sea Using Ferries as Platform of Observation. *Russian Journal of Marine Biology*, 5: 343-350.

Alves F., Ferreira R., Fernandes M., Halicka Z., Dias L. and Dinis A. (2018a). Analysis of occurrence patterns and biological factors of cetaceans based on data from platforms of opportunity: Madeira Island as a case study. *Marine Ecology*, e12499.

Alves F., Alessandrini A., Servidio A., Mendonça A.S., Hartma K.L., Prieto R, *et al.* (2018b). Complex biogeographical patterns support an ecological connectivity network of a large marine predator in the north-east Atlantic. *Diversity and Distributions*, 25(2): 269-284.

Baines M.E. and Reichelt M. (2014). Upwellings, canyons and whales: An important winter habitat for balaenopterid whales off Mauritania, northwest Africa. *Journal of Cetacean Research and Management*, 14: 57-67.

Beck J.R. and Shultz E.K. (1986). The use of relative operating characteristic (ROC) curves in test performance evaluation. *Archives of Pathology & Laboratory Medicine*, 110: 13-20.

Bohorquez J.J., Dvarskas A. and Pikitch E.K. (2019). Filling the Data Gap – A Pressing Need for Advancing MPA Sustainable Finance. *Frontiers in Marine Science*, 6: 45.

Boyd I.L., Wanless S. and Camphuysen C.J. (2006). *Top Predators in Marine Ecosystems. Their Role in Monitoring and Management* (Series: Conservation Biology). Cambridge, UK: Cambridge University Press.

Brito C. and Sousa A. (2011). The Environmental History of cetaceans in Portugal: Ten Centuries of Whale and Dolphin Records. *PLoS ONE*, 6(9): e23951.

Burnham K.P., Anderson D.R. (2002). *Model selection and multimodel inference: a practical information theoretic approach.* Springer Verlag, New York, USA.

Camphuysen C.J., van Spanje T.M. and Verdaat H. (2012). *Ship based seabird and marine mammal surveys off Mauritania, Nov-Dez 2012 – cruise report.* Mauritanian Institute for oceanographic research and fisheries - IMROP, 73pp.

Camphuysen C.J. (2000). *Seabirds and marine mammals off West Africa – Responses 2000 cruise report.* Netherlands Institute for Sea Research, 50pp.

Cañadas A., de Soto N.A., Aissie M., Arcangeli A., Azzoling M. and B-Nagyh A. (2018). The challenge of habitat modelling for threatened low density species using heterogeneous data: The case of Cuvier's beaked whales in the Mediterranean. *Ecological Indicators*, 85: 128-136.

Carrillo M., Pérez-Vallazza C. and Álvarez-Vázquez R. (2010). Cetacean diversity and distribution off Tenerife (Canary Islands). *Marine Biodiversity Records*, 3: 1-9.

Caldeira R.M.A. and Sangrà P. (2012). Complex geophysical wake flows. Madeira Archipelago case study. *Ocean Dynamics*, 62: 683-700.

Correia A.M., Tepsich P., Rosso M., Caldeira R. and Sousa-Pinto I. (2015). Cetacean occurrence and spatial distribution: Habitat modelling for offshore waters in the Portuguese EEZ (NE Atlantic). *Journal of Marine Systems*, 143: 73-85.

Correia A.M., Gandra M., Liberal M., Valente R., Gil A., Rosso M., *et al.* (2019a). A dataset of cetacean occurrences in the Eastern North Atlantic. *Scientific Data*, 6: 177.

Correia A.M., Gil A., Valente R., Rosso M., Pierce G.J. and Sousa-Pinto I. (2019b). Distribution and habitat modelling for short-beaked common dolphins (*Delphinus delphis*) in Eastern North Atlantic Ocean. *Journal of the Marine Biological Association UK*, 1-15.

Covelo P., Martínez-Cedeira J.A., Llavona Á., Díaz J.I. and López, A. (2016) Strandings of beaked whales (Ziphiidae) in Galicia (NW Spain) between 1990 and 2013. *Journal of the Marine Biological Association of the United Kingdom,* 96: 925-931.

Cox T.M., Ragen T.J., Read A.J., Vos E., Baird R.W., Balcomb K., *et al.* (2006). Understanding the impacts of anthropogenic sound on beaked whales. *Journal of Cetacean Research and Management*, 7(3): 177-187.

Díaz López B. and Methion S. (2019). Habitat drivers of endangered rorqual whales in a highly impacted upwelling region. *Ecological Indicators*, 103: 610-616.

Díaz López B., Methion S. and Giralt Paradell O. (2019). Living on the edge: Overlap between a marine predator's habitat use and fisheries in the Northeast Atlantic waters (NW Spain). *Progress in Oceanography*, 175: 115-123.

Dinis A., Alves F., Nicolau C., Ribeiro C., Kaufmann M., Cañadas A. and Freitas L. (2016a). Bottlenose dolphin *Tursiops truncatus* group dynamics, site fidelity, residency and movement patterns in the Madeira Archipelago (North-East Atlantic). *African Journal of Marine Science*, 38: 151-160.

Dinis A., Carvalho A., Alves F., Nicola C., Ribeiro C., Kaufmann M., *et al.* (2016b). Spatial and temporal distribution of bottlenose dolphins, *Tursiops truncatus*, in the Madeira archipelago, NE Atlantic. Arquipelago. *Life and Marine Sciences*, 33_ 45-54.

Dinis A., Marques R., Dias L., Sousa D., Gomes C., Abreu N. and Alves F. (2017). Site Fidelity of Blainville's Beaked Whale (*Mesoplodon densirostris*) off Madeira Island (Northeast Atlantic). *Aquatic Mammals*, 43(4): 387-390.

Djiba A., Bamy I.L., Bilal A.S.O. and Van Waerebee K. (2015). Biodiversity of cetaceans in coastal waters of northwest Africa: new insights through platform-of-opportunity visual surveying in 2011-2013. *IOC Technical Series*, 115: 283-297.

Dolman, S.J., Evans P.G.H., Nortarbartolo-di-Sciara G. and Frisch H. (2011). Active sonar, beaked whales and European regional policy. *Marine Pollution Bulletin*, 63: 27-34.

Fernández R., García-Tiscar S., Santos M.B., López A., Martínez-Cedeira J.A., Newton J. and Pierce G.J. (2011). Stable isotope analysis in two sympatric populations of bottlenose dolphins *Tursiops truncatus*: evidence of resource partitioning? *Marine Biology*, 158: 104-1055.

Fernández R., Santos M.B., Pierce G.J., Llavona Á., López A., Silva M.A., *et al.* (2011). Fine-scale genetic structure of bottlenose dolphins, *Tursiops truncatus*, in Atlantic coastal waters of the Iberian Peninsula. *Hydrobiologia*, 670: 11-125.

Freitas L., Dinis A., Nicolau C., Ribeiro C. and Alves F. (2012). New records of cetacean species for Madeira Archipelago with an updated checklist. *Boletim do Museu Municipal do Funchal*, 62: 25-43.

GEBCO (2017). Gridded bathymetry data. http://www.gebco.net/data_and_products/gridded_bathymetry_data/ (accessed September 2017).

Paradell O., Díaz López B. and Methion, S. (2019). Modelling common dolphin (*Delphinus delphis*) coastal distribution and habitat use: insights for conservation. *Ocean and Coastal Management*, 179:1.

Goetz S., Read F.L., Santos M.B., Pita C. and Pierce G.J. (2013). Cetacean–fishery interactions in Galicia (NW Spain): results and management implications of a face-to-face interview survey of local fishers. *ICES Journal of Marine Sciences*, 71 (3): 604-617.

Gove J.M., McManus M.A., Neuheimer A.B., Polovina J.J., Drazen J.C., Smith C.R., *et al.* (2016). Near-island biological hotspots in barren ocean basins. *Nature Communications*, 7: 10581.

Hastie T. and Tibshirani R. (1990) *Generalised Additive Models. Monographs on Statistics and Applied Probability 43.* London, UK: Chapman and Hall.

Hazen E.L., Abrahms B., Brodie S., Carroll G., Jacox M.G., Savoca M.S., *et al.* (2019). Marine top predators as climate and ecosystem sentinels. *Frontiers in Ecology and the Environment*, 17(10): 565-574.

Hazevoet C.J. and Wenzel W. (2000). Whales and dolphins (Mammalia, Cetacea) of the Cape Verde Islands with special reference to the Humpback Whale *Megaptera novaeangliae* (Borowski, 1871). *Contributions to Zoology*, 69(3): 197-211.

Hazevoet C.J., Monteiro V., López P., Varo N., Torda G., Berrow S. and Gravanita B. (2010). Recent data on whales and dolphins (Mammalia: Cetacea) from the Cape Verde Islands, including records of four taxa new to the archipelago. *Zoologia Caboverdiana*, 1: 75-99.

Heithaus M.R., Frid A., Wirsing A.J. and Worm B. (2008). Predicting ecological consequences of marine top predator declines. *Trends in Ecology and Evolution*, 23(4): 202-210.

Heyning J.E. (1989). Cuvier’s beaked whale *Ziphius cavirostris* g. Cuvier, 1823. In: Ridgway S.H., & Harrison R.S. (eds) *Handbook of marine mammals*, Vol 4. River dolphins and the larger toothed whales. Academic Press, London, pp 289−308.

Hooker S.K., Cañadas A., Hyrenbach D., Corrigan C., Polovina J.J. and Reeves R.R (2011). Making protected area networks effective for marine top predators. *Endangered Species Research*, 13: 203-218.

Karpouzli E. and Leaper R. (2004). Opportunistic observations of interactions between sperm whales and deep-water trawlers based on sightings from fisheries observers in the northwest Atlantic. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 14: 95-103.

Kaschner K., Quick N.J., Jewell R., Williams R. and Harris C.M. (2012). Global Coverage of Cetacean Line-Transect Surveys: Status Quo, Data Gaps and Future Challenges. *PLoS ONE*, 7(9): e44075.

Kiszka J., Macleod K., Van Canneyt O., Walker D. and Ridoux V. (2007). Distribution, encounter rates, and habitat characteristics of toothed cetaceans in the Bay of Biscay and adjacent waters from platform-of-opportunity data. *ICES Journal of Marine Science*, 64: 103-1043.

Liu C., Berry P.M., Dawson T.P. and Pearson R.G. (2005). Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, 28: 385-393.

MacLeod C.D. (2009). Global climate change, range changes and potential implications for the conservation of marine cetaceans: a review and synthesis. *Endangered Species Research*, 7, 125-136.

MacLeod C.D. and D'Amico A. (2006). A review of beaked whale behaviour and ecology in relation to assessing and mitigating impacts of anthropogenic noise. *Journal of Cetacean Research and Management*, 7(3), 211-221.

Mannocci L., Roberts J.J., Halpin P.N., Authier M., Boisseau O., Bradai M.N., *et al.* (2018). Assessing cetacean surveys throughout the Mediterranean Sea: a gap analysis in environmental space. *Scientific Reports*, 8: 3126.

Marubini F., Gimona A., Evans P.G.H., Wright P.J. and Pierce G.J. (2009). Habitat preferences and interannual variability in occurrence of the harbour porpoise *Phocoena phocoena* off northwest Scotland. *Marine Ecology Progress Series*, 381: 297-310.

Mason E. (2009). High-resolution Modelling of the Canary Basin Oceanic Circulation. Doctorate thesis (PhD) in Oceanography, Universidad de Las Palmas de Gran Canaria, Las Palmas de Gran Canaria. 245 pp.

Morgado C., Martins A., Rosso M., Moulins A. and Tepsich P. (2017). Fin Whale Presence and Distribution in the Pelagos Sanctuary: Temporal and Spatial Variability Along 2 Fixed-Line Transects Monitored in 2009-2013. *International Journal of Marine and Environmental Sciences*, 1: 1-14.

Morissette L., Kaschner K. and Gerber L.R. (2010). Ecosystem models clarify the trophic role of whales off Northwest Africa. *Marine Ecology Progress Series*, 404: 289-302.

Moura A.E., Sillero N. and Rodrigues A. (2012). Common dolphin (*Delphinus delphis*) habitat preferences using data from two platforms of opportunity. *Acta Oecologica*, 38: 24-32.

Moura A.E., Silva S.E., SPEA, Correia A.M., Sousa-Pinto I., Gil A., *et al.* (2017). In: Bencatel J., Álvares F., Moura A. E. and Barbosa, A. M. (eds.) (2017). *Atlas de Mamíferos de Portugal.* Universidade de Évora, Portugal.

Mullié W.C., Wagne M.M., Elmamy C.A.A., Yahya F.M., Veen J. and Van Waerebeek K. (2013). *Large number of stranded harbour porpoises Phocoena phocoena as by-catch victims in Mauritania.* Scientific Committee Document of International Whaling Commission, Jeju, Korea, June 2013. SC/65a/HIM03, 5pp.

Nagel P. and Gray T. (2012). Is the EU's Fisheries Partnership Agreement (FPA) with Mauritania a genuine partnership or exploitation by the EU? *Ocean & Coastal Management*, 56: 26-34.

Parsons E.C.M. (2016). Why IUCN should replace 'Data Deficient' conservation status with a precautionary 'Assume Threatened' status - A cetacean case study. *Frontiers in Marine Science*, 3: 193.

Pérez-Vallazza C., Álvarez-Vázquez R., Cardona L., Pintado C. and Hernández-Brito J. (2008). Cetacean diversity at the west coast of La Palma Island (Canary Islands). *Journal of the Marine Biological Association of the United Kingdom*, 88(6): 1289-1296.

Perrin W.F. and Van Waerebeek K. (2012). *The small-cetacean fauna of the west coast of Africa and Macaronesia: diversity and distribution*. Convention on the Conservation of Migratory Species of Wild Animals, Bonn, Germany, CMS Technical Series, 26, 7-17.

Richard G., Guinet C., Bonnel J., Gasco N. and Tixier P. (2017). Do commercial fisheries display optimal foraging? The case of longline fisheries in competition with odontocetes. *Canadian Journal of Fisheries and Aquatic Sciences*, 75(6): 964-976.

Robineau D. and Vely M. (1998). Ces cétacés des côtes de Mauritanie (Afrique du Nordouest). Particularltés et variations spatio-temporelles de repartition: role des facteurs oceanographiques. *Revue d Ecologie (Terre Vie)*, 53.

Quian S.S. (2009). *Environmental and Ecological Statistics with R*. Chapman & Hall, Canada.

Sala I., Caldeira R.M.A., Estrada-Allis S.N., Froufe E. and Couvelard X. (2013). Lagrangian transport pathways in the northeast Atlantic and their environmental impact. *Limnology and Oceanography: Fluids and Environments*, 3: 40-60.

Sergio F., Newton I., Marchesi L. and Pedrini P. (2006). Ecologically justified charisma: preservation of top predators delivers biodiversity conservation. *Journal of Applied Ecology*, 43: 1049-1055.

Sergio F., Caro T., Brown D., Clucas B., Hunter J., Ketchum J., *et al.* (2008). Top Predators as Conservation Tools: Ecological Rationale, Assumptions, and Efficacy. *Annual Review of Ecology, Evolution, and Systematics*, 39: 1-19.

Silva M.A., Prieto R., Cascγo I., Seabra M.I., Machete M., Baumgartner M.F. and Santos R.S. (2014). Spatial and temporal distribution of cetaceans in the mid-Atlantic waters around the Azores. *Marine Biology Research*, 10: 123-137.

de Soto N.A., Martín V., Silva M., Edler R., Reyes C., Carrillo M., *et al.* (2017). True's beaked whale (*Mesoplodon mirus*) in Macaronesia. *PeerJ*, 5: e3059.

Tepsich P., Rosso M., Halpin P.N. and Moulins A. (2014). Habitat preferences of two deep-diving cetacean species in the northern Ligurian Sea. *Marine Ecology Progress Series*, 508: 247-260.

Tixier P., Burch P., Richard G., Olsson K., Welsford D., Lea M.A., *et al*. (2019). Commercial fishing patterns influence odontocete whale longline interactions in the Southern Ocean. *Scientific Reports*, 9: 1904.

Tobeña M., Prieto R., Machete M. and Silva M.A. (2016). Modeling the Potential Distribution and Richness of Cetaceans in the Azores from Fisheries Observer Program Data. *Frontiers in Marine Science,* 3: 202.

Tulp I., and Leopold M.F. (2004). *Marine mammals and seabirds in Mauritanian waters. Pilot study April 2004 - Internal Report*. RIVO-Netherlands Institute for Fisheries Research - Animal Sciences Group. Wageningen UR, 42 pp.

UNEP-WCMC, IUCN and NGS (2018). *Protected Planet Report 2018.* UNEP-WCMC, IUCN and NGS: Cambridge UK; Gland, Switzerland; and Washington, D.C., USA.

Valente R., Correia A.M., Gil A., Gonzalez-Garcia L. and Sousa-Pinto I. (2019). Baleen whales in Macaronesia: insights on occurrence patterns through a bibliographic review. *Mammal Review*, 49(2): 129-151.

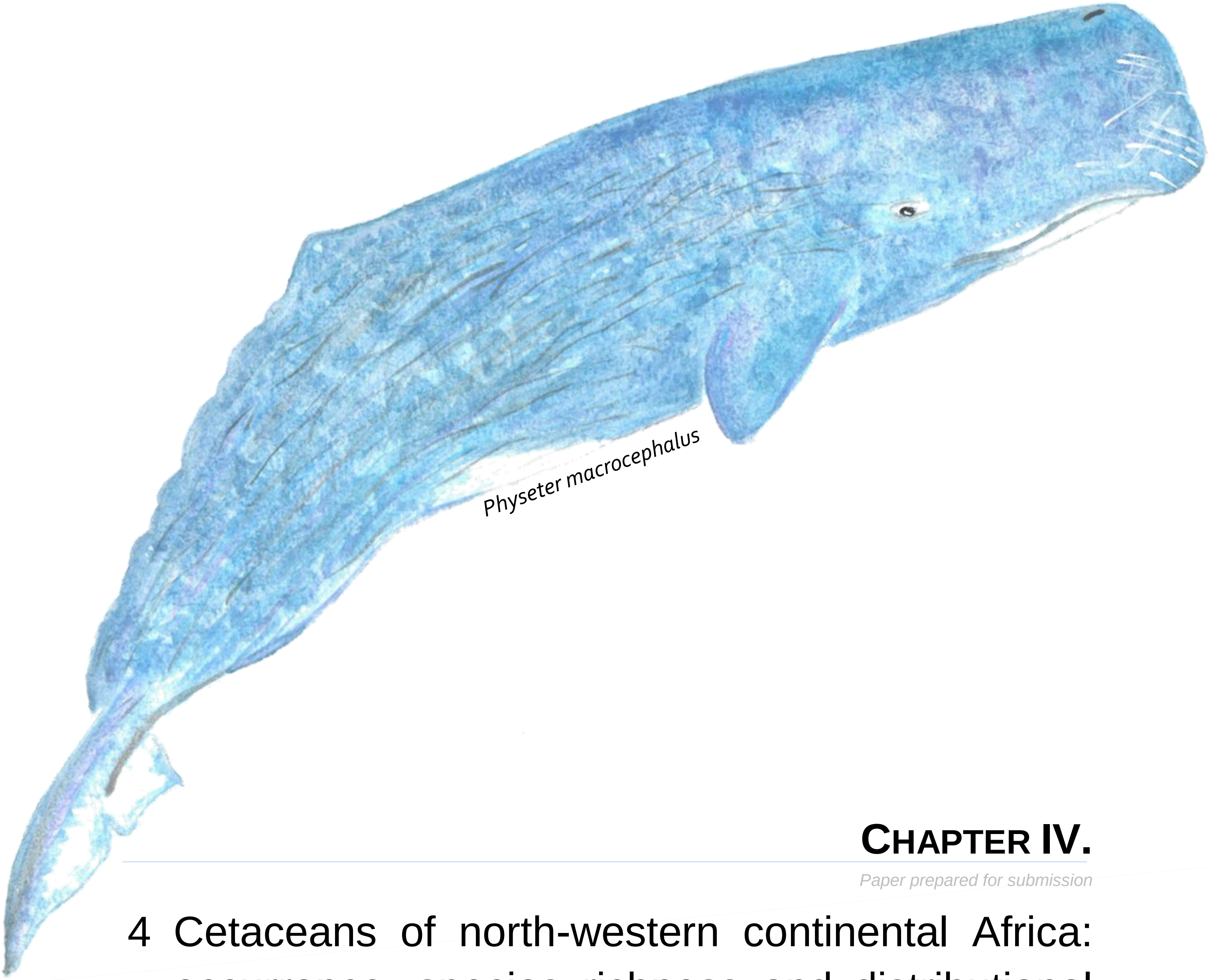
Van Waerebeek K., André M., Sequeira M., Martín V., Robineau D., Collet A., *et al.* (1999). Spatial and temporal distribution of the minke whale, *Balaenoptera acutorostrata* (Lacepede, 1804), in the Mediterranean Sea, with reference to stock identity. *Journal of Cetacean Research and Management*, 1(3): 223-237.

Vikingsson G.A. and Heide-Jørgensen M.P. (2015). First indications of autumn migration routes and destination of common minke whales tracked by satellite in the North Atlantic during 2001–2011. *Marine Mammal Science*, 31(1): 376-385.

Viddi F.A., Hucke-Gaete R., Torres-Florez J.P. and Ribeiro S. (2010). Spatial and seasonal variability in cetacean distribution in the fjords of northern Patagonia, Chile. *ICES Journal of Marine Science*, 67: 959-970.

Weir C.R. and Pierce G.J. (2013). A review of the human activities impacting cetaceans in the eastern tropical Atlantic. *Mammal Review*, 43: 258-274.

Zenk W., Klein B. and Schroder M. (1991). Cape Verde frontal zone. *Deep-Sea Research*, 38, 505–530.Zuur, A.F., Ieno, E. & Elphick, C. (2010). A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution*, 1: 3-14.



CHAPTER IV.

Paper prepared for submission

4 Cetaceans of north-western continental Africa: occurrence, species richness and distributional range from Morocco to Liberia

Ana M. Correia[1,2], Ágatha Gil[1,2], Raul Fonseca Valente[1,2], Massimiliano Rosso[1,3], Graham J. Pierce[4,5,6], Isabel Sousa-Pinto[1,2]

[1]Interdisciplinary Centre of Marine and Environmental Research (CIIMAR). 4450-208 Matosinhos, Portugal.
[2]Department of Biology, Faculty of Sciences, University of Porto (FCUP). 4169-007 Porto, Portugal.
[3]CIMA Research Foundation. 17100 Savona, Italy.
[4]Instituto de Investigacións Mariñas (CSIC). 36208 Vigo, Pontevedra, Spain.
[5]Oceanlab, University of Aberdeen. AB41 6AA, UK.
[6]CESAM and Department of Biology, University of Aveiro. 3810-193 Aveiro, Portugal.

4.1 Abstract

The north-western coast of Africa (NWA) is one of the most productive marine systems in the world and a high diversity of cetacean species has been recorded. However, management of human activities which impact cetaceans in the area is inadequate, in part due to poor knowledge of the occurrence of cetaceans and the lack of systematic surveys. We review existing knowledge on cetacean distribution in the NWA and provide new records (along with information on encounter rates) collected within the Exclusive Economic Zones of Western Sahara, Mauritania and Senegal during CETUS Project (cetacean monitoring programme). The literature review yielded 3873 records of cetaceans belonging to a total of 37 species. The species reported as occurring most frequently by most sources was common dolphin but, overall, bottlenose dolphin was the species with most records. During the CETUS surveys, approximately 9832 kilometres of survey track were scanned for cetacean presence, resulting in 332 sightings of 17 species, the sperm whale being the most frequently sighted. While Mauritanian waters were best covered by the literature, in the CETUS campaigns, most survey effort was in the Western Saharan. Information on cetacean distribution and abundance in the NWA is fundamental to ensure their conservation.

Keywords: cetaceans; distribution; Mauritania; Senegal; Western Sahara

4.2 Introduction

The African Large Marine Ecosystem (ALME) is among the most productive in the world, with the Canary Current System (CCS) resulting in a very powerful coastal upwelling and highly dynamic oceanographic processes along the north-western coast of Africa (NWA) (Caldeira *et al.,* 2002; Mason, 2009; Sala *et al.,* 2013; Satia, 2015; UNEP-CMS, 2008; UNEP-CMS, 2012). In the region, the Islamic Republic of Mauritania's Banc d'Arguin National Park, the largest Marine Protected Area (MPA) in the western Africa, has some of the richest fish resources in the world (FAO, 2013). This MPA is one of the legally protected areas established between Mauritania and Guinea, where a significant percentage of the west African marine biodiversity can be found, including several species of marine mammals (Subregional Fisheries Commission, 2003). A high number of cetacean species has been recorded in the NWA, some of them passing through the area during their annual migrations (see Jefferson *et al.,* 1997; Perrin & Van Waerebeek, 2012; Robineau & Vely, 1998; Weir & Pierce, 2013). The conservation of cetacean populations in the NWA depends on the management of the CCS area (UNEP-CMS, 2012).

Under the Convention on Migratory Species of Wild Animals (CMS), the Memorandum of Understanding (MOU) Concerning the Conservation of the West African Manatee and Small Cetaceans of Western Africa and Macaronesia was concluded in 2008 and it was the first regional initiative to address the conservation of cetaceans within these waters. However, although discussed, large cetaceans were not included in the memorandum, to avoid overpal with the remit of the International Whaling Commission. Moreover, the memorandum is not legally binding, it was still not signed by all west African states, and some of them, e.g. Western Sahara, are not even range states. Within the MOU, particular emphasis is given to the need for data collection and the urgency of adequate management of fishery resources (Caddell, 2009; CMS, https://www.cms.int). The CMS (United Nations Environment Programme (UNEP)) (2008) listed numerous threats to small cetaceans on the African coast due to human activities, among them those related with fisheries: by-catch, direct take, trophic interaction with fisheries, overfishing and ship strikes.

Since 1990, there has been an overcapacity in the industrial fishing sector, mostly due to the high number of trawlers in the west African coastal zone (FAO, 2013), and over-exploitation of marine resources is the major threat to marine mammals (UNEP-CMS, 2008; Trouillet *et al.,* 2011; UNEP-CMS, 2012; Weir and Pierce, 2013; Diedhiou and Yang, 2018). Although efforts have been made to manage over-exploitation of resources in west African waters through partnerships between the European Union (EU) and west African states (namely, the Fisheries Partnership Agreement introduced in 2003), these have so far failed to accomplish their objective to achieve sustainable use of fishery resources through shared decision-making between EU and African countries (Corten, 2014; Nagel & Gray, 2011). Diedhiou & Yang (2018) concluded that, although some policies have been effective and efforts are being made to improve management of fisheries, these are not sufficient and both illegal fishing and overfishing remain issues to be addressed.

Weir & Pierce (2013) reviewed the impact of the human activities on cetaceans in the eastern tropical Atlantic but the scarcity of data, due to the absence of routine cetacean monitoring programmes, makes it difficult to estimate the scale of the impacts. Baseline knowledge on cetacean distribution in the NWA is lacking and existing records come from occasional and isolated boat-based surveys (Djiba *et al.,* 2015; Expósito & Qnimba, 2010), surveys covering small parts of the area (Baines & Reichelt, 2014; Bamy *et al.,* 2010; Boisseau *et al.,* 2010; Leeney *et al.,* 2016; Massky & de Stéphanis, 2015; Moore *et al.,* 2010; Mullié *et al.,* 2013; Rojo-Nieto *et al.,* 2011; Russel *et al.,* 2018; Tudela *et al.,* 2005) and/or dedicated to certain species (Benchoucha *et al.,* 2018; Boisseau *et al.,* 2007; Fertl *et al.,* 2003; Hammond & Lockyer, 1988; Jung *et al.,* 2016; Notarbartolo di Sciara *et al.,* 1997; Robineau & Vely, 1993; Tsai & Mead, 2018; Van Waerebeek *et al.,* 1999; 2013; 2017; Weir *et al.,* 2014; Weir & Collins, 2015; Weir, 2016), and “grey literature” (Bowman Bishaw Gorham, 2003; Camphuysen, 2000; Camphuysen *et al.,* 2012; Tulp & Leopold, 2004; UNEP-CMS 2012; Van Waerebeek *et al.,* 2000). The available information has previously been summarised in four reviews (Jefferson *et al.,* 1997; Perrin & Van Waerebeek, 2012; Robards & Reeves, 2011; Robineau & Vely 1998; Weir & Pierce 2013). The creation and maintenance of an updated cetacean species inventory in the area, through non-lethal research, is essential to allow reliable assessments of impacts of anthropogenic threats and to support efficient marine management (UNEP-CMS 2008; 2012; Weir and Pierce 2013).

During 2015 and 2016, 11 surveys crossed three of the African Exclusive Economic Zones (EEZ) in the NWA (Western Sahara, Mauritania and Senegal) and recorded cetacean sightings. These surveys were part of the CETUS Project, a cetacean monitoring program in Macaronesia that started in 2012 and has been collecting cetacean occurrence data along routes between Continental Portugal, Macaronesian islands and the NWA, with highest survey effort in oceanic waters. Monitoring is performed from cargo ships belonging to the Portuguese maritime company, TRANSINSULAR, which are used as Observation Platforms of Opportunity (OPOs). The surveys, although from OPOs, are performed systematically and by dedicated observers allowing effort-based data collection in poorly surveyed areas and seasons. The monitoring program is cost-effective and can potentially provide the long-term data needed to estimate cetacean abundance trends.

In order to provide a complete and quantitative inventory of cetacean species and identify data gaps, we review existing cetacean records for the NWA, from Morocco to Liberia, by species and by range state; and summarise the new findings from CETUS, from Western Sahara to Senegal, including information on spatial distribution and relative abundance.

4.3 Material and Methods

4.3.1 Literature review

Published data on cetacean occurrence in the NWA, from Morocco to Liberia, with the southern limit as defined by Weir (*et al.,* 2014), were compiled by source and by species/taxa. For each source, we accessed information on the type of data, the range of the study area, the period covered by the surveys, number of species reported, number of records and most frequently reported species (table in Supplementary file 1). Number of species reported was determined considering only positive identifications within the range of study (between Morocco and Liberia). In the case of sightings reported for the Canary Islands, as presented in Djiba *et al.* (2015) and Camphuysen (2000), the range state was considered as Morocco and/or Western

Sahara because the surveys took place between the west coast of Africa and the Canary Islands, crossing both EEZs (Spanish and African) and, therefore, it was safe to assume that sightings were within (or at the limit) of the study area. Records were categorized into sightings and non-sightings: sightings correspond to occurrences reported from visual records, from which group size could often be assessed (each record is a single sighting regardless of the number of individuals) in order to provide an insight into dispersion and abundance of each species in the NWA; while non-sightings records correspond to other sources of data such as strandings, by-catches, direct takes, cetacean collection and acoustic data, and consist of information on the number of individuals. Classification of non-sightings is not presented as in many cases the origin of a record is ambiguous or it falls under more than one category. For the number of records, both confirmed and unconfirmed identifications from either sightings or non-sightings records were included as long as the locations were definitely within the range of study. To compile data by species, both quantitative and qualitative information were considered, in order to provide number of records, group size and distributional range. Records potentially within the study area but not confirmed are also presented by species; as well as data for records with identifications not confirmed to the species level (in Supplementary file 2). In the latest case, both sightings with doubts among species or categorized as non-identified (NI) dolphins, beaked and baleen whales or cetaceans were included.

All data were cross-checked between sources and all efforts were made to avoid repeated records. Nevertheless, when insufficient data was provided (date and region), although this rarely happened, cross-checking may have failed. If repeated among the sources compiled, records were maintained for the first source presenting the data, while counting all the sources when indicating the number of species by source and number of sources by species. An exception was made for the humpback dolphin (*Sousa teuszii*), for which a very complete and recent review was made by Weir & Collins (2015), with new unpublished data also included. Since it was not possible to trace all these records back to the original sources, we treated Weir & Collins (2015) as the primary source, provided that the other sources were cited by them. Since several secondary records were thus used in the present study, the original sources are cited in supplementary material (Supplementary file 3) along with other information for each record collected (date, region).

4.3.2 New data collection

Surveys crossing the EEZs of the Western Sahara, Mauritania and Senegal were performed in 2015 (7 surveys from May to October) and in 2016 (4 surveys from August to November) aboard a cargo ship belonging to a Portuguese maritime transport company, TRANSINSULAR. The vessel was used as an OPO for cetacean monitoring. Surveys were carried out along a route between Continental Portugal and Cape Verde, with stop-overs in the Canary Islands and Mauritania, and were part of a cetacean monitoring program for Macaronesia that started in 2012, the CETUS Project (Correia *et al.,* 2015). During every survey, two trained Marine Mammal Observers (MMOs), one on each side of the ship, collected the data, monitoring from the front of the vessel with a field of view covering 180º (each MMO covering 90º, except when the other MMO was resting (as detailed below), in which case the lone MMO covered 180º), with the naked eye and occasionally scanning (or confirming species identification) with binoculars (7 x 50 mm).

MMOs stood on the navigation bridge and wings of the bridge, searching for cetacean presence continuously from sunrise to sunset, whenever weather conditions were favourable (with sea state and wind state up to 4, on the Douglas and Beaufort scales respectively, and visibility over 1 km) and the ship was sailing outside the ports. MMOs rested in turns, for an hour each at mealtimes (lunch and dinner), and optionally for additional periods of about 40 minutes (one in the morning and one in the afternoon). During resting periods, the MMO standing on the observation deck would cover 180º. Survey effort stopped occasionally when observers could not access the observation stands, i.e., during safety drills, cleaning of the deck or manoeuvres.

When cetaceans were sighted, the species was identified and number of individuals recorded. Identification was always attempted to the species level, but on several occasions, the identification was only possible to the genus level or registered as NI dolphin, baleen or beaked whale, or cetacean. When it was not possible to determine the exact number of individuals, a minimum and maximum number of animals was registered, as well as the most probable number of individuals according to the observer´s perception (best estimate). The behaviour of the animals towards the ship and their direction of travel was also registered. Occurrence of other conspicuous megafauna (e.g. sharks, turtles) was noted. Information on weather conditions was assessed at the beginning and end of survey and whenever there were significant changes. Marine traffic information was recorded at the beginning and end of survey, every hour and during cetacean sightings.

4.3.3 Data analysis

For data collected during CETUS surveys, encounter rates (ERs) were calculated as the number of sighting records per 100 kilometres and computed for each EEZ, each year and in total. For the calculation of ERs, only sightings recorded during survey effort were considered (i.e., “on effort” sightings). The total numbers of sightings reported include off effort sightings recorded opportunistically (e.g., during unfavourable weather conditions, safety drills, cleaning of the deck or manoeuvres). Number of sightings and group sizes were compiled per year of survey and in total. The best estimate for number of animals was used for the group size. Distribution maps were created in ArcGIS Pro by mapping occurrences and effort tracks within the limits of the EEZs of Western Sahara, Mauritania and Senegal. The coordinates of the position of the ship at the moment of the sightings were used for the location of the occurrences. Maps were georeferenced with the coordinate system GCS_WGS_1984 (EPSG: 4326) and projected to World_Mercator (EPSG: 3395). New data collected were analysed together and in comparison with data compiled from the literature review for each species encountered.

4.4 Results and Discussion

4.4.1 Literature review

A total of 36 published sources yielded 3873 records of cetaceans, 3830 positively confirmed as occurring within the study range, and a total of 37 identified species reported, with quantitative information available for 36 of them (*Eubalaena glacialis* with qualitative information only) (Table 1).

Table 1. Table summarizing the number of sightings records with confirmed identifications (to species level) found in the literature for each cetacean species in north-western Africa.

Species	Sightings	Non-sightings	Group size (min-max)	Unconfirmed location	No of sources
Phocoena phocoena	22	180[a]	1-11	1ns	15 (1,2,4,5,8,13,16,18,20,21,22,23,28,30,34)
Delphinus capensis	3	80	500	2ns	4 (6,21,25,30)
Delphinus delphis	142	303	1-400	-	15 (1,2,5,6,7,8,9,16,21,23,24,25,27,28,30)
Feresa attenuata	-	1	-	-	5 (13,21,22,23,30)
Globicephala macrorhynchus	72	192	10-30	1ns	15 (1,2,7,9,13,15,16,18,20,21,22,23,25,30,34)
Globicephala melas	3	27	8-14	-	8 (7,13,16,18,20,21,23,24)
Grampus griseus	20	9	3-15	-	14 (1,6,8,9,13,15,16,20,21,22,23,24,25,30)
Lagenodelphis hosei	-	1	-	-	4 (9,16,21,30)
Lagenorhynchus acutus	-	-	-	1s (-)	1 (13)
Lagenorhynchus albirostris	-	-	-	6s (-)	1 (13)
Sousa teuszii	341	48	1-40	-	15 (2,6,13,15,16,17,20,21,22,23,30,32,33,34, 36)
Stenella attenuata	19	-	1-300	2s (-)	9 (7,9,11,13,21,22,23,25,35)
Stenella clymene	27[b]	8	2-150	-	12 (1,6,7,9,13,15,21,23,25,30,34,35)
Stenella coeruleoalba	7	274	12-150	-	15 (1,5,6,7,9,13,16,20,21,23,24,27,30,34)
Stenella frontalis	38	43	2-50	-	13 (1,2,8,9,10,13,16,18,20,21,25,30,34)
Stenella longirostris	11	10	9-40	1s (-)	10 (6,7,9,13,16,21,23,25,30,34)
Steno bredanensis	7	42	12-?	1ns	11 (1,2,13,16,17,21,22,23,25,30,34)
Orcinus orca	68[c]	21	1-15	-	12 (1,9,12,13,15,16,20,21,23,28,30,34)
Peponocephala electra	-	3	-	-	6 (9,13,15,21,23,30)
Pseudorca crassidens	3	-	1-40	-	5 (13,16,19, 21,23)
Tursiops truncatus	337	246 [d,e]	1-50	2ns	22 (1,2,6,7,8,9,10,13,15,16,17,18,20,21,22,23,24, 25,28,30,32,34)
Hyperoodon ampullatus	3	-	1-3	-	4 (7,16,21,23)
Mesoplodon densirostris	-	1	-	-	3 (16,21,23)
Mesoplodon europaeus	-	2	-	-	5 (15,16,21,23,30)
Ziphius cavirostris	3	5	2-2	-	7 (7,10,16,21,23,24,30)
Kogia breviceps	-	7	-	-	7 (2,16,21,22,23,30,34)
Kogia sima	2	3	-	-	5 (3,21,23,25,30)
Physeter macrocephalus	54	18	1-10	1ns	13 (1,5,7,8,9,15,16,20,23,24,25,30,34)
Balaenoptera acutorostrata	5[f]	17	1-2	2ns	10 (2,10,16,20,22,23,24,29,30,34)
Balaenoptera borealis	18	2	1-18	1ns	8 (1,5,6,9,16,23,25,30)
Balaenoptera edeni	6	4	1-2	-	8 (2,8,9,16,23,25,30,34)
Balaenoptera musculus	19	-	1-3	1s (12)	6 (1,8,9,16,23,25)
Balaenoptera omurai	-	1	-	-	2 (9,14)
Balaenoptera physalus	7	17	1-6	-	9 (1,5,7,8,16,23,24,28,34)
Caperea marginata	-	1	-	-	1 (26)
Eubalaena glacialis	Qualitative data only				2 (16,23)
Megaptera novaeangliae	29	5	1-6	-	12 (1,2,8,9,15,16,23,24,25,28,31,34)

No – number; s – sightings records; ns – non-sightings records. Records compiled as with unconfirmed location are those probably within the study range (Morocco to Liberia) but without sufficient data for a positive confirmation, the group size (min-max) of these sightings records is within brackets. The sources reporting on the species are within brackets and identified by the numbers: 1. Baines & Reichelt, 2014; 2. Bamy *et al.,* 2010; 3. Benchoucha *et al.,* 2018; 4. Boisseau *et al.,* 2007; 5. Boisseau *et al.,* 2010; 6. Bowman Bishaw Goham, 2003; 7. Camphuysen, 2000; 8. Camphuysen *et al.,* 2012; 9. Djiba *et al.,* 2015; 10. Expósito and Qnimba, 2010; 11. Gray & Van Waerebeek, 2011; 12. Hammond & Lockyer, 1988; 13. Jefferson *et al.,* 1997; 14. Jung *et al.,* 2016; 15. Leeney *et al.,* 2016; 16. Massky & de Stéphanis, 2015; 17. Moore *et al.,* 2010; 18. Mullié *et al.,* 2013; 19. Notarbartolo di Sciara *et al.,* 1997; 20. Pinela *et al.,* 2010; 21. Perrin & Van Waerebeek, 2012; 22. Robards & Reeves, 2011; 23. Robineau & Vely, 1998; 24. Rojo-Nieto *et al.,* 2011; 25. Russel *et al.,* 2018; 26. Tsai & Mead, 2018; 27. Tudela *et al.,* 2005; 28. Tulp & Leopold, 2004; 29. Van Waerebeek *et al.,* 1999; 30. Van Waerebeek *et al.,* 2000; 31. Van Waerebeek *et al.,* 2013; 32. Van Waerebeek *et al.,* 2017; 33. Weir & Collins, 2015; 34. Weir & Pierce, 2013; 35. Weir *et al.,* 2014; 36. Weir, 2016.

[a] There was one more non-sighting report: 19 captures + strandings. Captures are mostly/all cited in Weir & Pierce (2013). There is no confirmation on the exact number of strandings, hence, this record was not considered from the source Robineau & Vely (1998).

[b] One sighting corresponds to one individual, the “Senegal dolphin”, sighted several times in the area.

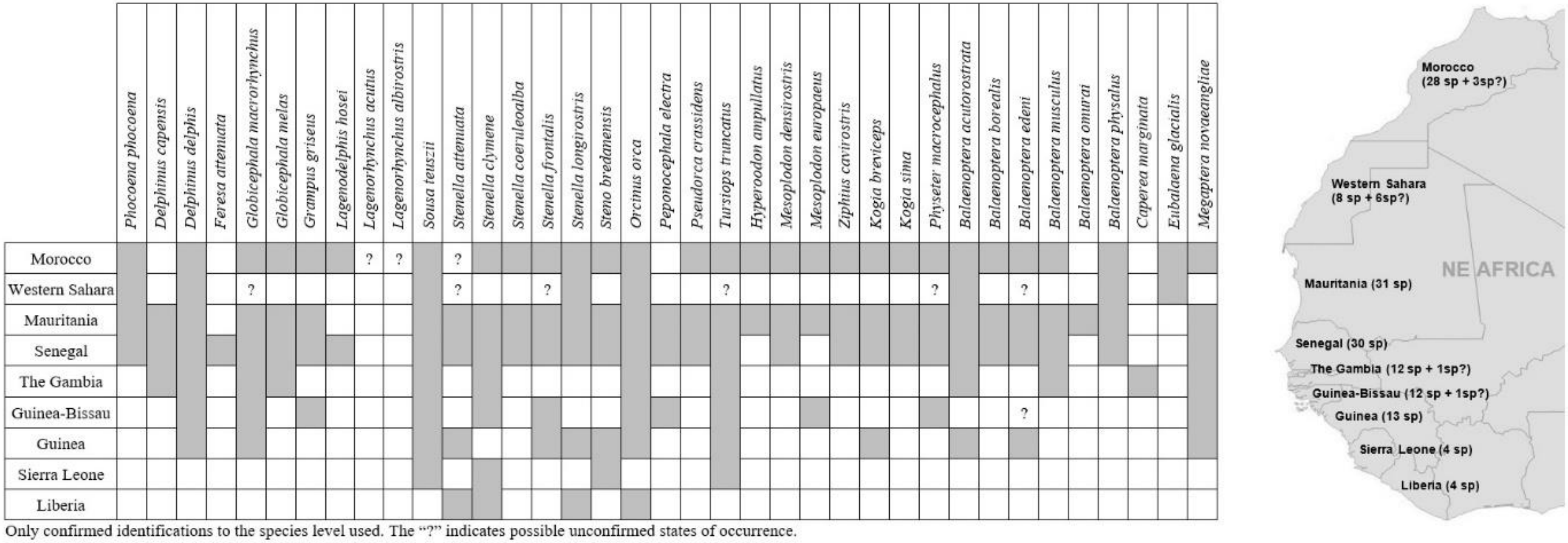
[c] Three sightings correspond to groups that were sighted several times in the area.

[d] There was two more non-sightings reports: 15 captures + strandings; 17 specimens. Captures are mostly/all cited in Weir and Pierce 2013 and the 17 specimens are under the IFAN collection (presented by Van Waerebeek *et al.,* 2000). There is no confirmation on the exact number of strandings, hence, these records were not considered from the source Robineau & Vely (1998).

[e] There was one more by-catch event of 10 animals of *Tursiops truncatus* + *Delphinus* sp., but since there was no specification of numbers of each species, these were not considered.

[f] One sighting corresponds to one individual sighted several times in the area.

The distributional range is provided for all species, but two of them (from the Lagenorhynchus genus) only have unconfirmed occurrence in the area (Morocco). Mauritania and Senegal were the countries with the highest number of confirmed species (31 and 30 species, respectively) (Figure 1).



	Phocoena phocoena	*Delphinus capensis*	*Delphinus delphis*	*Feresa attenuata*	*Globicephala macrorhynchus*	*Globicephala melas*	*Grampus griseus*	*Lagenodelphis hosei*	*Lagenorhynchus acutus*	*Lagenorhynchus albirostris*	*Sousa teuszii*	*Stenella attenuata*	*Stenella clymene*	*Stenella coeruleoalba*	*Stenella frontalis*	*Stenella longirostris*	*Steno bredanensis*	*Orcinus orca*	*Peponocephala electra*	*Pseudorca crassidens*	*Tursiops truncatus*	*Hyperoodon ampullatus*	*Mesoplodon densirostris*	*Mesoplodon europaeus*	*Ziphius cavirostris*	*Kogia breviceps*	*Kogia sima*	*Physeter macrocephalus*	*Balaenoptera acutorostrata*	*Balaenoptera borealis*	*Balaenoptera edeni*	*Balaenoptera musculus*	*Balaenoptera omurai*	*Balaenoptera physalus*	*Caperea marginata*	*Eubalaena glacialis*	*Megaptera novaeangliae*
Morocco	▨		▨		▨	▨	▨	▨	?	?	▨	?	▨	▨	▨	▨	▨	▨		▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨		▨		▨	▨
Western Sahara	▨		▨		?						▨	?			?	▨		▨			?							?	▨		?			▨		▨	
Mauritania	▨	▨	▨		▨	▨	▨				▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨			▨
Senegal	▨	▨	▨	▨	▨	▨	▨	▨			▨	▨	▨	▨	▨	▨	▨	▨	▨	▨	▨		▨		▨	▨	▨	▨	▨	▨	▨	▨		▨			▨
The Gambia		▨	▨		▨	▨					▨		▨					▨			▨								▨			▨			▨		▨
Guinea-Bissau			▨		▨		▨				▨		▨		▨			▨	▨		▨			▨				▨			?						▨
Guinea			▨		▨						▨	▨			▨	▨	▨	▨			▨					▨			▨		▨						▨
Sierra Leone											▨		▨				▨				▨																
Liberia												▨	▨			▨		▨																			

Only confirmed identifications to the species level used. The "?" indicates possible unconfirmed states of occurrence.

Figure 1. Distributional range of cetacean species reported in literature for the northwest Africa (from Morocco to Liberia).

Mauritania was the most frequently surveyed area within the NWA (21 sources), followed by Morocco (16 sources), Senegal (13 sources), The Gambia, Guinea-Bissau and Guinea (12 sources), Western Sahara and Sierra Leone (nine sources), and Liberia (eight).

Djiba *et al.* (2015) presented the highest number of sightings records (286), while highest number of non-sightings records (551) was compiled from Van Waerebeek *et al.* (2000) (Supplementary file 1). Van Waerebeek *et al.* (2000) reported on the Project WAFCET I (CMS-UNEP) that aimed to assess the conservation status of cetaceans in Senegal, The Gambia and Guinea-Bissau, and reviewed cetacean occurrence within those waters. Moreover, this work presents the historical registers of cetacean specimens (most of the records from skulls) resulting from a review and rehabilitation of the cetacean collection from the Institut Fondamental d’Afrique Noir (IFAN-Dakar). In Liberia and Sierra Leone, there were no non-sightings records collected, while in Senegal and Mauritania, several sources provide data on strandings, captures and by-catches (e.g., Robineau & Vely, 1998, Van Waerebeek *et al.,* 2000, Weir & Pierce, 2013). Robards & Reeves (2011) present worldwide information on marine mammal consumption by humans and within the area, Senegal is the country where more cetacean species are consumed. Consentino & Fisher (2016) suggest that consumption of aquatic mammals from incidental and direct takes in many west African is still common. Among collected non-sightings are also included 50 acoustic records of harbour porpoise (*Phocoena phocoena*) collected from Morocco to Mauritania (Boisseau *et al.,* 2007).

4.4.2 New data collected

Between 2015 (May to October) and 2016 (August to November), 11 surveys were carried out, covering a total of 9832 km in the NWA. Survey effort was heterogeneous between years (from 6006 km in 2015 to 3826 km in 2016) and along the route (around 68% of effort coverage (6562 km) was within the Western Saharan EEZ) (Table 2). Additionally, CETUS surveys were mostly during summer months, which are poorly represented in published literature.

Table 2. Summary table of the CETUS project cetacean monitoring campaigns, with total survey effort, sightings, encounter rate and number of species sighted by Exclusive Economic Zones (EEZ) covered and by year of survey.

Surveyed EEZ	**2015**				**2016**				**Total**			
	Effort	**Sight. Tot. / Ef.**	**ER**	**Sp.**	**Effort**	**Sight. Tot. / Ef.**	**ER**	**Sp.**	**Effort**	**Sight. Tot. / Ef.**	**ER**	**Sp.**
Western Saharan EEZ	3671	94 / 78	2.12	10	2891	144 / 64	2.21	13	6562	238 / 142	2.16	13
Mauritanian EEZ	2226	32 / 30	1.35	10	935	55 / 24	2.57	7	3161	87 / 54	1.71	11
Senegalese EEZ	109	7 / 6	5.50	5	0	-	-	0	109	7 / 6	5.50	5
Total	6006	133 / 114	1.90	15	3826	199 / 88	2.30	14	9832	332 / 202	2.05	17

Effort is presented in kilometres; Sight. Tot / Ef. – Number of sightings recorded in total (Tot.) and on effort (Ef.); ER – Encounter rate presented in sightings per 100 kilometres; Sp. – total number of species identified (at least, to the genus level).

The new sightings records confirm or extend the documented distributional range to the Western Sahara for 10 taxa: bottlenose dolphin (*Tursiops truncatus*), Atlantic spotted dolphin (*Stenella frontalis*), Clymene dolphin (*Stenella clymene*), striped dolphin (*Stenella coeruleoalba*), rough-toothed dolphin (*Steno bredanensis*), sperm whale (*Physeter macrocephalus*), pilot whales (*Globicephala* sp.), Risso's dolphin (*Grampus griseus*), humpback whale (*Megaptera novaeangliae*) and beaked whales (Ziphiidae). Details on these encounters are presented in supplementary material (Supplementary file 4).

A total of 332 new sightings of 15 identified cetacean species (17 including those identified only to the genus level) was recorded in the present study. There were 202 sightings collected on effort, resulting in an overall encounter rate of 2.05 sightings / 100 km. The highest encounter rate was registered in 2015 in the Senegalese EEZ but this is most likely not representative of the region and rather a consequence of the low effort undertaken. The Western Saharan EEZ had a higher overall encounter rate compared to the Mauritanian EEZ (2.16 sightings per 100 km against 1.71 sightings per 100 km), and more identified species (13 against 11 species) (Table 3).

Table 3. Table with cetacean sightings per species collected during the CETUS project in the Exclusive Economic Zones of Western Sahara, Mauritania and Senegal, for each year of cetacean monitoring surveys, and in total, as well as the group sizes.

Species	2015		2016		TOTAL	
	Total sightings / Sightings on effort	Group size min-max (mean±SD)	Total sightings / Sightings on effort	Group size min-max (mean±SD)	Total sightings / Sightings on effort	Group size min-max (mean±SD)
Stenella frontalis	10 / 6	8-150 (45.60±46.59)	18 / 9	8-100 (31.72±27.28)	28 / 15	8-150 (36.68±35.18)
Delphinus sp.	8 / 8	3-2500 (438±851.84)	7 / 3	5-200 (49.29±68.64)	15 / 11	3-2500 (256.60±636.50)
Stenella coeruleoalba	2 / 2	7-50	12 / 7	9-120 (38±35.64)	14 / 9	7-120 (36.64±34.03)
Tursiops truncatus	6 / 5	4-30 (12.50±11.78)	5 / 3	10-130 (62.4±58.47)	11 / 8	4-130 (35.18±46.00)
Stenella clymene	8 / 7	30-700 (190.63±258.85)	1 / 0	25	9 / 7	25-700 (172.22±248.35)
Lagenodelphis hosei	1 / 1	450	-	-	1 / 1	450
Steno bredanensis	-	-	1 / 1	30	1 / 1	30
NI dolphin	32 / 27	1-100 (17.44±22.22)	42 / 25	1-120 (26.31±32.27)	74 / 52	1-120 (22.47±28.53)
Physeter macrocephalus	19 / 18	1-7 (1.84±1.71)	23 / 8	1-4 (1.35±0.71)	42 / 26	1-7 (1.57±1.27)
Globicephala sp.	5 / 4	7-100 (38.4±36.75)	10 / 8	4-50 (24.10±15.44)	15 / 12	4-100 (28.87±24.25)
Orcinus orca	1 / 1	2	3 / 2	2-5 (4±1.73)	4 / 3	2-5 (3.50±1.73)
Peponocephala electra	3 / 1	40-200 (93.33±92.38)	1 / 1	3	4 / 2	3-200 (70.75±87.91)
Grampus griseus	-	-	1 / 0	15	1 / 0	15
NI beaked whale	7 / 6	1-3 (1.43±0.79)	4 / 2	2-5 (3.25±1.50)	11 / 8	1-5 (2.09±1.38)
Megaptera novaeangliae	3 / 3	1-3 (2.33±1.15)	2 / 1	1-1	5 / 4	1-3 (1.80±1.10)
Balaenoptera acutorostrata	2 / 2	1-1	2 / 0	1-3	4 / 2	1-3 (1.50±1.00)
Balaenoptera physalus	3 / 3	1-4 (2±1.73)	1 / 1	1	4 / 4	1-4 (1.75±1.50)
Balaenoptera edeni	1 / 1	2	-	-	1 / 1	2
Balaenoptera musculus	1 / 1	2	-	-	1 / 1	2
NI baleen whale	3 / 3	1-2 (1.33±0.58)	21 / 4	1-8 (1.62±1.72)	24 / 7	1-8 (1.58±1.61)
NI cetacean	12 / 10	1-5 (1.67±1.23)	42 / 13	1-50 (5.12±10.55)	54 / 23	1-50 (4.35±9.41)
Globicephala sp. / *Tursiops truncatus*	2 / 1	45-110	2 / 0	50-100	4 / 1	45-110 (76.25±29.02)
Delphinus sp. / *Stenella frontalis*	2 / 2	20-200	-	-	2 / 2	20-200
Globicephala sp. /*Stenella clymene*	1 / 1	32	-	-	1 / 1	32
Physeter macrocephalus / NI dolphin	1 / 1	3	-	-	1 / 1	3
Stenella frontalis / NI dolphin	-	-	1 / 0	100	1 / 0	100
TOTAL	133 / 114	-	199 /88	-	332 / 202	-

min – minimum number of individuals recorded; max – maximum number of individuals recorded; mean – average group size; SD – standard deviation of the group size; NI – Non-identified

It was also where most taxa were recorded, with the Fraser's dolphin (*Lagenodelphis hosei*), the melon-headed whale (*Peponocephala eletra*), Bryde's whale (*Balaenoptera edeni*) and the blue whale (*Balaenoptera musculus*) being the only species not encountered. In Senegalese waters, only five taxa were registered. Differences in species appear to be related with the amount of research effort undertaken in each EEZ (Supplementary file 5).

With 239 sightings (including the sightings with associated species) (Table 3), odontocetes comprised around 72% of the total sightings (Figure 2 and 3), while mysticetes represented only 12% (Figure 4). There were also sightings of non-identified species: 54 non-identified cetaceans (the 16% not included in any of the odontocete and mysticete categories), 74 non-identified dolphins, 11 non-identified beaked whales and 24 non-identified baleen whales (Figure 5).

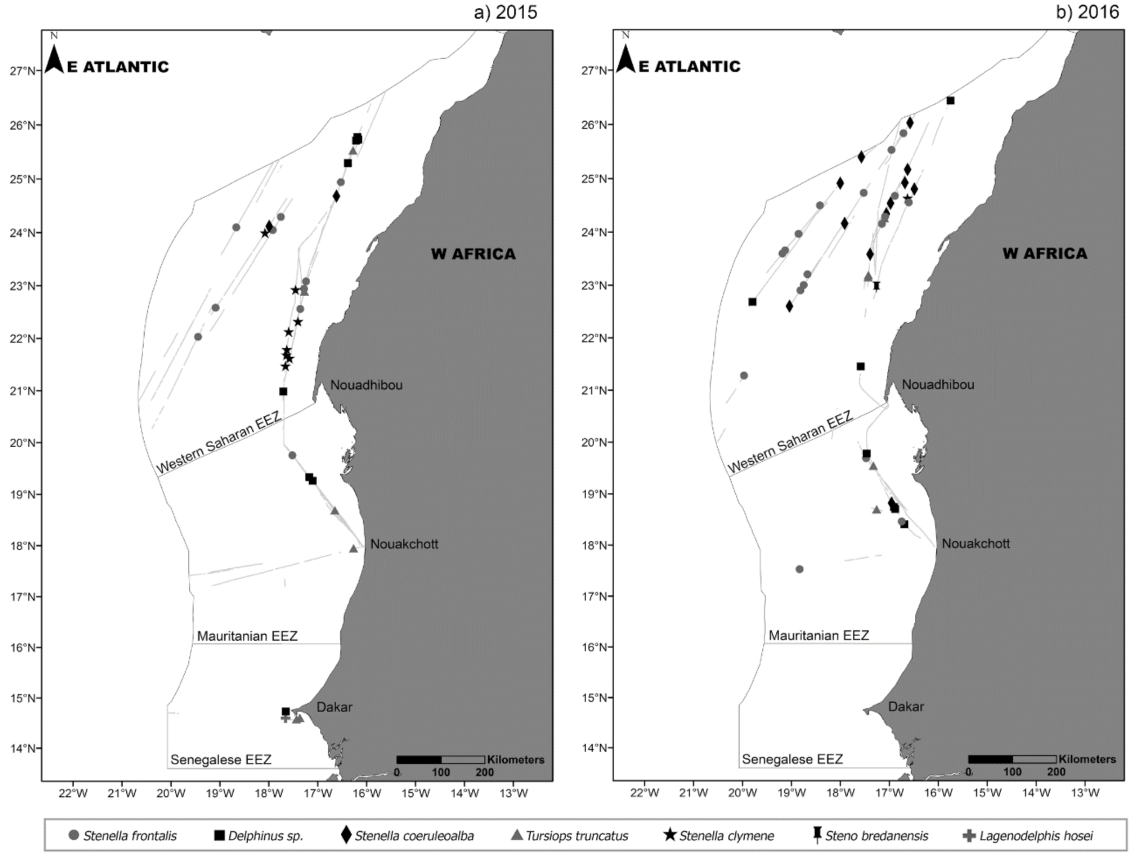


Figure 2. Distribution of dolphin sightings and survey effort from CETUS Project is presented for the Exclusive Economic Zones of Western Sahara, Mauritania and Senegal for the years of 2015 (a) and 2016 (b). Effort tracks are presented in light grey lines.

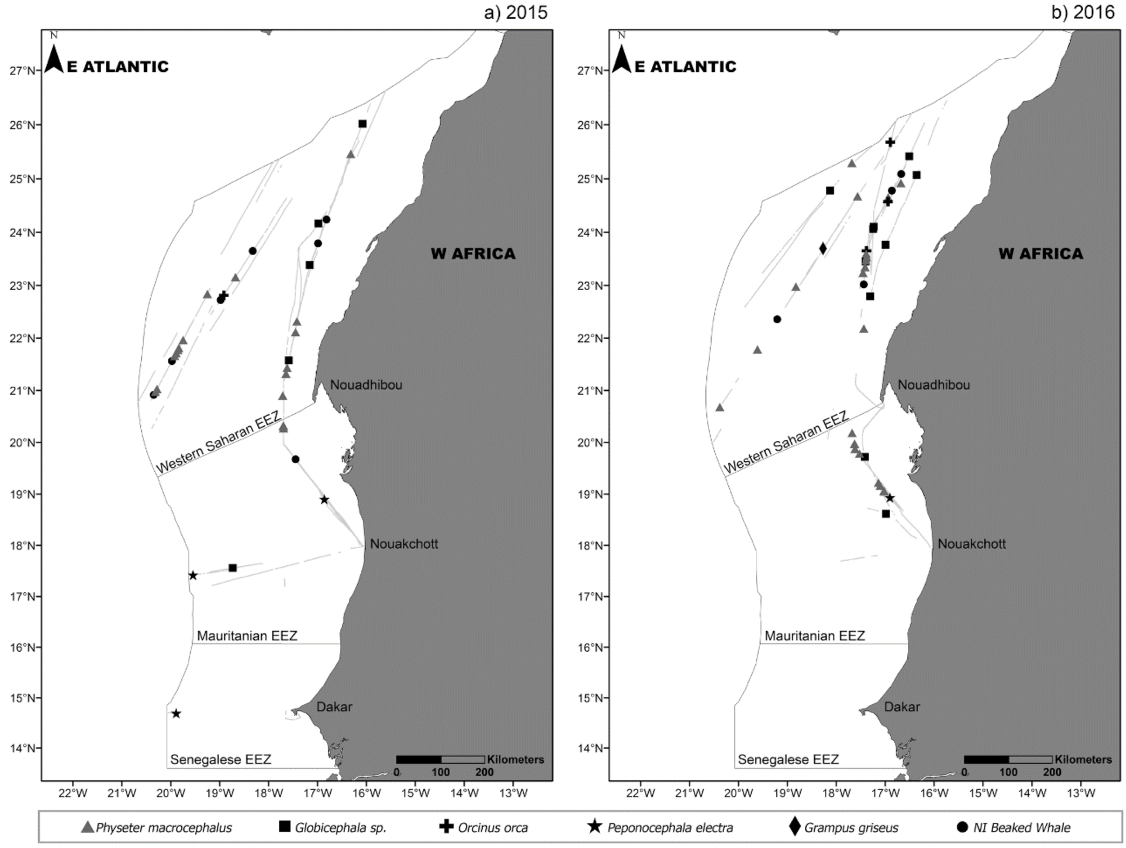


Figure 3. Distribution of toothed whale sightings and survey effort from CETUS Project is presented for the Exclusive Economic Zones of Western Sahara, Mauritania and Senegal for the years of 2015 (a) and 2016 (b). NI – Non-identified. Effort tracks are presented in light grey lines.

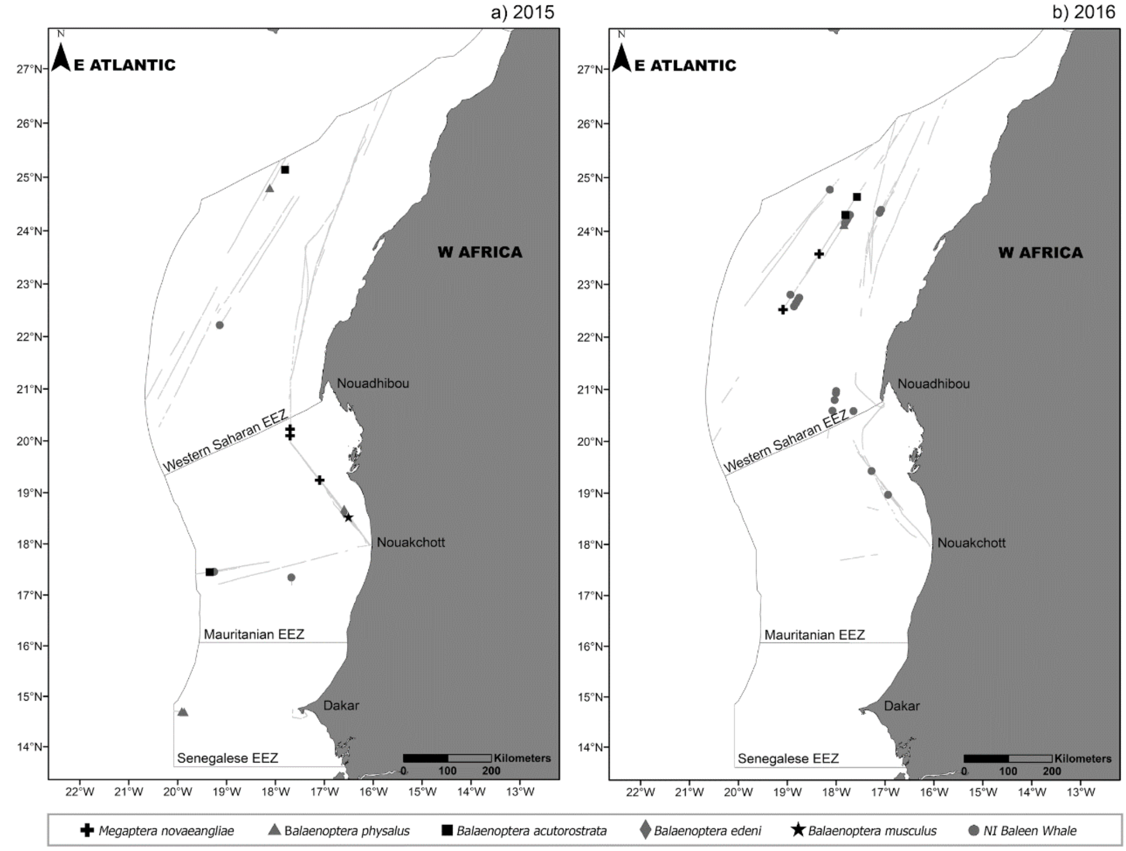


Figure 4. Distribution of baleen whale sightings and survey effort from CETUS Project is presented for the Exclusive Economic Zones of Western Sahara, Mauritania and Senegal for the years of 2015 (a) and 2016 (b). NI – Non-identified. Effort tracks are presented in light grey lines.

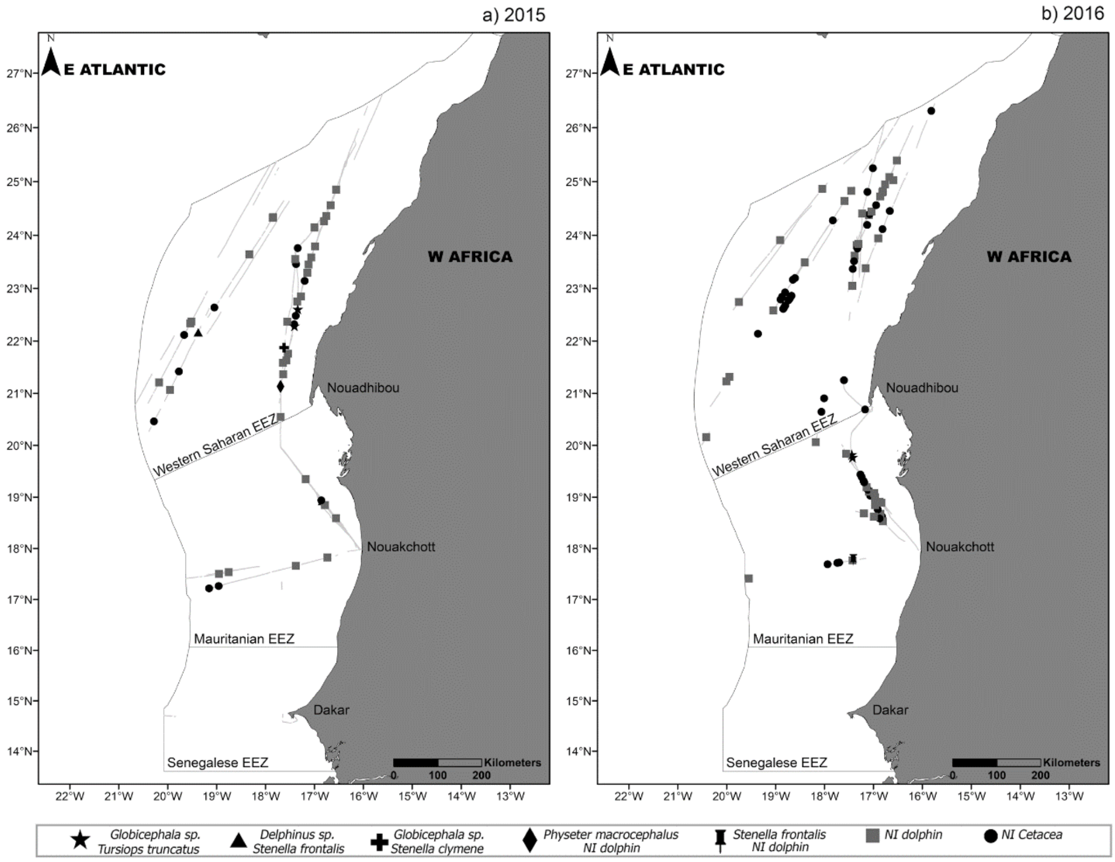


Figure 5. Distribution of non-identified cetaceans, non-identified dolphins, sightings with associated species, and survey effort from CETUS Project is presented for the Exclusive Economic Zones of Western Sahara, Mauritania and Senegal for the years of 2015 (a) and 2016 (b). NI – Non-identified. Effort tracks are presented in light grey lines.

4.4.3 Species accounts

Harbour porpoise (*Phocoena phocoena*)

The presence of harbour porpoises in the area is reported by 15 publications, which yielded 180 non-sightings records and 22 sightings of groups of between one to 11 animals (Table 1). Aside from the acoustic records (50 from Boisseau *et al.,* 2007), most non-sightings were strandings but there are also several records of direct takes and by-catches (Mullié *et al.,* 2013; Van Waerebeek *et al.,* 2000; Weir & Pierce, 2013). Fontaine (2016) points to the need for an assessment of the status and threats faced by the southern ecotype (and putative subspecies) of harbour porpoise in the eastern North Atlantic, *Phocoena phocoena meridionalis*. This is a unique ecotype with small populations and low genetic diversity, inhabiting the coasts of Iberian Peninsula (Iberia) and north-western Africa. There is a need for further research on the impacts of exploitation and adaptation in the face of climate change (Fontaine, 2016). The present work shows that the distributional range of the species in the study area extends, at least, from Morocco to Senegal (Figure 1).

This is the smallest cetacean species in the North Atlantic and generally presents a coastal distribution. Nevertheless, excursions beyond the continental shelf are also documented, for example off Greenland (Nielsen *et al.,* 2018). No encounters of harbour porpoises in the area were recorded during the CETUS campaigns. However, this species is difficult to spot from the platform of observation (at a considerable height) and most of the effort was in oceanic waters.

Long-beaked common dolphin (*Delphinus capensis*) and short-beaked common dolphin (*Delphinus delphis*)

It has been previously suggested and recently recognized by the Society for Marine Mammology that only one species of common dolphins exist, the *Delphinus delphis* (Cunha *et al.,* 2015; List of Marine Mammal Species and Subspecies, www.marinemammalscience.org). However, in the literature the two common previously recognized common dolphin species were distinguished and separately reported. The short-beaked common dolphin was reported as the most frequent species in five sources (Supplementary file 1). It was often sighted in big pods of up to 400 animals (Table 1). There were 142 sightings and 303 non-sightings records of short-beaked common dolphins collected from 15 literature sources (Table 1), with the distributional rage extending from Morocco to Guinea (Figure 1).

The long-beaked common dolphin was reported by Bowman Bishaw Gorham (2003) as a single sighting of 500 animals off Mauritania; a further two sightings records were collected from Russel *et al.* (2018) and 80 non-sightings from Van Waerebeek *et al.* (2000). The distributional range extends from Mauritania to The Gambia and is thus less extensive than the distribution of the short-beaked dolphin (Figure 1).

Considering records of animals identified only to the genus level, i.e. for *Delphinus* sp., there were 5 more sightings and 252 non-sightings records encountered in literature.

During CETUS surveys, 15 new records of *Delphinus* sp. were collected, and this was the species with the largest group size observed: an estimated 2500 animals in a group off Dakar in 2015 (Figure 2 and Table 3). Common dolphins were distributed across the three surveyed EEZs and were encountered associated with spotted dolphins on two occasions (Figure 2 and 5).

Pygmy killer whale (*Feresa attenuata*)

There was one single capture (that could be either from a directed take or a by-catch) of a pygmy killer whale in Senegal (Van Waerebeek *et al.,* 2000), its distribution in the region being restricted to this country (Table 1 and Figure 1).

Short-finned pilot whale (*Globicephala macrorhynchus*) and long-finned pilot whale (*Globicephala melas*)

Both species of pilot whales have been previously encountered in the area. Short-finned pilot whale was reported 264 times (192 non-sightings records and 72 sightings) from 15 sources in groups of between 10 and 30 animals (Table 1). Short-finned pilot whale was encountered from Morocco to Guinea, with unconfirmed occurrence in Western Sahara (Figure 1). Van Waerebeek *et al.* (2000) mention a mass-stranding event of 151 short-finned pilot whales in Senegal (20th May 1943) (Table 1). The species was the most frequently encountered by Russel *et al.* (2018) when monitoring Mauritanean waters during a geophysical survey.

The long-finned pilot whale was documented less frequently (30 records) from eight sources and was seen in smaller groups (eight to 14 animals) (Table 1). Long-finned pilot whale was distributed from Morocco to The Gambia (with no records in the Western Sahara) (Figure 1).

In addition, there were 15 records of *Globicephala* sp., with groups of between three to 20 animals, derived from seven sources (Supplementary file 2).

In the campaigns of the CETUS Project, it was not possible to distinguish between the two pilot whale species, and sightings of these animals were therefore registered as *Globicephala* sp. (Table 3). There were 15 sightings, distributed mostly in the Western Saharan EEZ but with 3 of the sightings in the Mauritanian EEZ (Table 3 and Figure 3). Groups varied in size between four to 100 animals, and the species was seen associated with dolphins on five occasions: four times with bottlenose dolphins and once with Clymene dolphins (Table 3 and Figure 5).

Rissos's dolphin (*Grampus griseus*)

Nine non-sightings records and 20 sightings of Risso's dolphin, in groups of between three to 15 animals, were collected from 14 literature sources. Occurrences were mostly documented in the area of Mauritania, but Djiba *et al.* (2015) and Rojo-Nieto *et al.* (2011) present two confirmed sightings and one stranding (respectively) in Morocco; and Van Waerebeek *et al.* (2000) document a sighting of six individuals in Guinea-Bissau (Table 1). According to literature, the species distributes from Morocco to Guinea-Bissau, but was not encountered in Western Sahara and The Gambia (Figure 1).

Rissos's dolphin was sighted only once during the CETUS sea-surveys, as a group of around 15 animals in the Western Saharan EEZ in 2016 (Table 3 and Figure 3).

Fraser's dolphin (*Lagenodelphis hosei*)

Quantitative records of Fraser's dolphin were reported in the NWA only by Van Waerebeek *et al.* (2000) and Djiba *et al.* (2015), who document the same stranding in Senegalese waters (Table 1). Jefferson *et al.* (1997) and Perrin and Van Waerebeek (2012) confirmed the occurrence of the species in Morocco (Figure 1). The CETUS surveys yielded a single occurrence of a group with about 450 animals, registered in 2015 in Senegal, off Dakar (Table 3 and Figure 2).

Atlantic white-sided dolphin (*Lagenorhynchus acutus*) and white-beaked dolphin (*Lagenorhynchus albirostris*)

Sightings of both species of the *Lagenorhynchus* genus were restricted to the Strait of Gibraltar, hence likely occurring, but without confirmation, in Moroccan waters (Table 1 and Figure 1). Additionally, there is an unconfirmed record of a skull of *Lagenorhynchus* sp. in Senegal but Jefferson *et al.* (1997) admits it may have been from a melon-headed whale (Supplementary file 2).

Atlantic humpback dolphin (*Sousa teuszii*)

Although the species was never encountered during the CETUS surveys, the Atlantic humpback dolphin was reported 389 times in literature (341 sightings and 48 non-sightings records), from 15 sources (Table 1). The distributional range extends from Morocco to Sierra Leone (Figure 1). Weir & Collins (2015) present an extensive and recent review of the occurrence of this species from the Strait of Gibraltar to Angola.

It is an endemic species and there are concerns about its status due to its restricted distribution and threats from human activities, such as climate change, over-fishing, marine bushmeat and coastal development (e.g, over-exploitation of mangroves, coastal construction, aquaculture, oil and gas exploration and extraction, accidental spills, increased shipping, tourism, and effluents), a consequence of the strong human demographic growth in the region (UNEP-CMS, 2007a; Van Waerebeek *et al.,* 2003; Weir *et al.,* 2010).

Pantropical spotted dolphin (*Stenella attenuata*)

The pantropical spotted dolphin is mentioned in eight literature sources but was not encountered during the CETUS surveys in the area (Table 1). In fact, within the area surveyed in the CETUS project, occurrence of the species is confirmed only for Senegal, where little survey effort was undertaken. Djiba *et al.* (2015) documented an unconfirmed sighting of a mixed group of pantropical spotted dolphins and common dolphins off Senegal (categorized as NI dolphin) and a confirmed group in Guinea, while Perrin & Van Waerebeek (2012) and Robards & Reeves (2011) report the occurrence of the species in Senegal and Guinea without mentioning specific sightings records (Table 1). A single source documents the presence of the species in the waters of Liberia (Gray & Van Waerebeek 2011).

Clymene dolphin (Stenella clymene)

In the literature, Clymene dolphins were reported 36 times, with 27 sightings of groups of between two and 150 animals, from 12 sources (Table 1). The species was reported from Morocco to Liberia, although it was never encountered in the Western Sahara and Guinea (Figure 1). Note that the review of Weir *et al.* (2014) presents 92 occurrences of the species from Mauritania to Angola, including the offshore islands, but here we considered only the 17 sightings from Mauritania to Liberia.

During the CETUS surveys, the Clymene dolphin was sighted nine times in groups of between 25 and 700 animals. All sightings were made in the Western Saharan EEZ, eight of them in 2015 (Figure 2). On one occasion, the species was found in a mixed group of about 32 animals with pilot whales (Table 3 and Figure 5).

Striped dolphin (*Stenella coeruleoalba*)

The literature review yielded 274 non-sightings records and seven sightings of groups of between 12 and 150 animals, from 15 sources. The species was documented in Morocco, Mauritania and Senegal (Table 1 and Figure 1). Most non-sightings records (266) were described for the Mediterranean Moroccan coast, with 178 strandings (Rojo-Nieto *et al.,* 2011) and 88 by-catches (Tudela *et al.,* 2005). Striped dolphin was the most frequently reported species in both these sources.

Striped dolphin was sighted 14 times during the CETUS surveys, 11 of which were in 2016 in the Western Saharan EEZ. Two other sightings were in the Western Saharan EEZ in 2015, with a single occurrence in the Mauritanian EEZ in 2016. Groups ranged in size from seven to 120 animals (Figure 2 and Table 3). Based on the literature review and new data collected, the species apparently occurs more frequently in northern waters.

Atlantic spotted dolphin (*Stenella frontalis*)

There were 43 non-sightings records, plus 38 occurrences of groups of between two and 50 Atlantic spotted dolphins, collected from 13 literature sources (Table 1). The species was the most frequently reported by Expósito & Qnimba (2010) but with only 2 occurrences in Moroccan waters. The distributional range of the species extended from Morocco to Guinea with unconfirmed occurrence in the Western Sahara and it was never encountered The Gambia (Figure 1).

In the CETUS surveys, the Atlantic spotted dolphin was the most frequently sighted dolphin species, with 28 sightings of groups of between eight and 150 animals, distributed between the Western Sahara and Mauritanian EEZs (Table 3 and Figure 2). Moreover, the species was found associated with other dolphin species on three occasions: twice in mixed groups with common dolphins and once with non-identified dolphins (Table 3 and Figure 5).

Spinner dolphin (*Stenella longirostris*)

Spinner dolphin was reported in 10 of the literature sources, totalling 10 non-sightings and 11 sightings of groups of between nine and 40 animals (Table 1). In the CETUS Project, the species was never encountered in the study area. Although spinner dolphins do not seem to occur very frequently in the NWA, their distributional range potentially includes the entire area as the species was seen from Morocco to Liberia (although no records were collected in The Gambia, Guinea-Bissau or Sierra Leone) (Figure 1).

Rough-toothed dolphin (*Steno bredanensis*)

In the literature, rough-toothed dolphin was reported 49 times (seven sightings plus 42 non-sightings records) from 11 sources. Group size was documented only once, by Baines & Reichelt (2014) who reported a group of 12 animals in Mauritanian waters (Table 1). The species had a scattered occurrence in the area: from Morocco to Sierra Leone, but absent in Western Sahara, The Gambia and Guinea-Bissau (Figure 1). During the CETUS surveys, a single sighting was recorded: a group of about 30 animals in the Western Saharan EEZ in 2016 (Table 3 and Figure 1).

Killer whale (*Orcinus orca*)

The literature review yielded 21 non-sightings records and 68 sightings of killer whales from 12 sources, with groups ranging in size from one to 15 animals (Table 1). The species is documented from Morocco to Liberia, although never encountered in Sierra Leone (Figure 1). Hammond & Lockyer (1988) present a review of killer whale occurrence from the British Islands to the Ecuador (Supplementary file 1).

During the CETUS surveys, there were four sightings of groups of two to five killer whales, all in the Western Saharan EEZ (Table 3 and Figure 3).

Melon-headed whale (*Peponocephala electra*)

There were three non-sightings records of melon-headed whale in the literature. The distribution range of the species includes Mauritania, Senegal and Guinea-Bissau.

During the CETUS surveys, the species was encountered on four occasions, three times in the Mauritanian EEZ and once in the Senegalese EEZ (in 2015). Two of the sightings occurred in offshore areas, at the limit of the EEZs (Figure 3). Groups ranged in size from three to 200 animals (Table 3).

False killer whale (*Pseudorca crassidens*)

Only three sightings off Morocco were collected for the false killer whale (Jefferson *et al.,* 1997; Notarbartolo di Sciara *et al.,* 1997), although the species is documented to occur from Morocco to Senegal (not encountered in the Western Sahara) (Table 1 and Figure 1).

Bottlenose dolphin (*Tursiops truncatus*)

In the literature, bottlenose dolphin was the most frequently reported species, with 22 sources yielding 246 non-sightings records and 337 sightings of groups of between one and 150 animals (Table 1). The species is stated to be one of the most impacted by human activities in the area, which is reflected by the high number of non-sightings records, mainly of strandings, but also including several reports of incidental and direct takes (Robineau & Vely, 1998; Van

Waerebeek *et al.,* 2000; Weir & Pierce, 2013). In four of the sources, this species was reported as being the most frequently recorded cetacean (Supplementary file 1). The species distribution was documented as extending from Morocco to Sierra Leone, with unconfirmed occurrence in the Western Sahara (Figure 1).

In CETUS surveys, bottlenose dolphin was encountered 11 times in groups of between four and 130 animals and was found in all three EEZs (Table 3 and Figure 2), thus confirming its occurrence in the Western Sahara. It was also found in mixed groups with pilot whales on four occasions, an association commonly reported in literature (e.g., Djiba *et al.,* 2015; Tulp & Leopold, 2004) (Table 3 and Figure 5). There were few sightings in the Senegalese EEZ, probably related with the low effort (Figure 2).

Northern bottlenose whale (*Hyperoodon ampullatus*), Cuvier’s beaked whale (*Ziphius cavirostris*), Gervais' beaked whale (*Mesoplodon europeaus*), Blainville's beaked whale (*Mesoplodon densirostris*)

Four species of beaked whales were previously documented in the area: northern bottlenose whale, Cuvier’s, Gervais' and Blainville's beaked whales. However, there are few records for each (from one to six records), Cuvier’s beaked whale being the most frequently reported (Table 1). These species are mainly distributed in northern waters although the occurrence of Gervais’ beaked whale in Guinea-Bissau is confirmed by the finding of a skull at an unspecified location - originally reported by Reiner (1980) and cited in Robineau & Vely (1998); and also mentioned by Perrin & Van Waerebeek (2012). According to literature, these species were not encountered in Western Sahara (Figure 1).

During the CETUS surveys, beaked whales were never identified to the species level, but 11 sightings of non-identified beaked whales were recorded, 10 in the Western Saharan EEZ and one in the Mauritanian EEZ (Table 3 and Figure 3).

Beaked whales are elusive and difficult to spot at sea. However, they are extremely susceptible to human threats (Kaschner *et al.,* 2012; MacLeod *et al.,* 2008), hence their occurrence and distribution in the area need to be further studied.

Pygmy sperm whale (*Kogia breviceps*) and Dwarf sperm whale (*Kogia sima*)

Pygmy sperm whale was reported seven times in the literature by seven sources, and there were five records of dwarf sperm whale from five sources (both in Senegal) (Table 1). Both species present a patchy distribution in the NWA (Morocco and Mauritania to Senegal, with the *K. breviceps* reported also in Guinea (Figure 1). The genus *Kogia* was described as occurring in Sierra Leone by Moore *et al.* (2010), however no quantitative information or identification of the species were provided (Supplementary file 2). The genus was never encountered during the CETUS surveys.

Sperm whale (*Physeter macrocephalus*)

In the literature, sperm whale was the most frequently reported species in two of the sources (Supplementary file 1). In total, there were 18 non-sightings records and 54 sightings of groups varying from one to 10 animals, documented in 13 sources (Table 1). The species is distributed from Morocco to Guinea-Bissau, with unconfirmed occurrence in the Western Sahara and never encountered in The Gambia (Figure 1).

During the CETUS surveys, there were 42 sightings of sperm whale and it was the most frequently sighted cetacean species, overall and in both years of the survey. Groups varied in size between one and seven animals (Table 3). The species was encountered in the Western Saharan and Mauritanian EEZs (Figure 3). In 2015, a sperm whale was sighted associated with two non-identified dolphins in the Western Saharan EEZ (Figure 5).

Minke whale (*Balaenoptera acutorostrata*)

Other than the 17 non-sightings records, the minke whales were sighted only five times in the area and in small groups (one to two animals). They are reported by 10 sources (Table 1), which document a distribution from Morocco to Guinea (although not encountered in Guinea-Bissau) (Figure 1). Van Waerebeek *et al.* (1999) presents 11 of the records in a review of the species occurrence from Galicia to Gabon, including offshore waters and oceanic islands (Supplementary file 1).

The species was encountered four times during the CETUS surveys in groups ranging in size from one to three animals (Table 3). They were mostly distributed in the Western Saharan EEZ, with only one occurrence in the Mauritanian EEZ (Figure 4).

Sei whale (*Balaenoptera borealis*) and Bryde’s whale (*Balaenoptera edeni*)

In the literature, there were 20 confirmed records of sei whales, 19 in Mauritanian waters and one in Morocco, compiled from eight sources, with maximum group size of eight individuals (Table 1).

Ten records of Bryde's whale were collected from eight sources, from which six sightings with group size varying from one to two animals (Table 1).

Both species are distributed from Morocco to Senegal, although the occurrence of Bryde's whale in Western Sahara is unconfirmed and the sei whale was never encountered in the Western Sahara. Moreover, Bryde's whale has unconfirmed occurrence in Guinea-Bissau and was reported to occur in Guinea (Figure 1).

There were 62 more sightings of either sei or Bryde's whale and five of sei, Bryde's or blue whale. If these unconfirmed sightings are included, both sei and Bryde´s whale are potentially distributed from Morocco to Guinea (Supplementary file 2).

Although sei whale was not identified during the CETUS surveys and only one group of two Bryde's whales was sighted (Table 3 and Figure 4), these species were probably present among the non-identified baleen whales, hence their occurrence is likely underestimated. As is evident from the probable sightings mentioned in the literature (see above), these two species are often indistinguishable at sea (Baines & Reichelt, 2014; Camphuysen *et al.,* 2012; Djiba *et al.,* 2015).

Blue whale (*Balaenoptera musculus*)

For blue whales, 19 sightings were compiled from six sources, with groups ranging in size from one to three animals (Table 1). Occurrences were reported mainly in the Mauritanian EEZ, with a single sighting of a group of three individuals off The Gambia (Djiba *et al.,* 2015). The distributional range was described as extending from Morocco to The Gambia (but not encountered in Western Sahara) (Figure 1).

Data collected during the CETUS surveys included a single sighting of two blue whales in 2015 in the Mauritanian EEZ (Figure 4 and Table 3).

Omura's whale (*Balaenoptera omurai*)

In 2013, one specimen of Omura's whale stranded on a beach in the south of Mauritania. This was an isolated event and the origin of this individual was uncertain (Jung *et al.,* 2016) (Table 1).

Fin whale (*Balaenoptera physalus*)

In the literature, fin whales were reported by nine sources, totalling 24 records, including seven sightings of groups ranging in size from one to six animals (Table 1). The southern limit of the species occurrence is Senegal (Figure 1).

Fin whale sightings occurred 4 times during the CETUS surveys, with groups ranging in size from one to four animals, and three of the encounters were in 2015 (Table 3). They were distributed across the three surveyed EEZs (Figure 4).

Pygmy right whale (*Caperea marginata*)

A stranded pygmy right whale in a beach of The Gambia, in January 1995, was reported by Tsai & Mead (2018). This location of the stranding delimits the northern limit of the range distribution for the species. No other records of the whale distribution in the study area were found in literature.

North atlantic right whale (*Eubalaena glacialis*)

No records were collected for the North Atlantic right whale, but qualitative information confirms the occurrence of the species in Morocco and Western Sahara (Table 1 and Figure 1).

Humpback whale (*Megaptera novaeangliae*)

Humpback whales were reported 34 times, including 29 sightings of groups of between one to six animals, by 12 sources – it was the most frequently registered baleen whale, with a distribution range extending from Morocco to Guinea (but not reported in the Western Sahara) (Table 1 and Figure 1).

During the CETUS surveys, humpback whale was also the most commonly identified baleen whale species, occurring in groups of between one to three animals. It was sighted on five occasions: three in 2015 in the Mauritanian EEZ and two in 2016 in the Western Saharan EEZ (Table 3 and Figure 4).

The Cape Verde islands are a well-known wintering ground for this species in the northeast Atlantic (e.g., Hazevoet & Wenzel, 2000; Hazevoet *et al.,* 2010; 2011) but the distribution of the species in the coastal areas of the NWA is poorly studied. The occurrences compiled here, both from the literature and the CETUS surveys, may represent migrant animals. Nonetheless, Van Waerebeek (2013) presents strong evidence that the area between Senegal and Guinea may also be a wintering and nursing ground for a south Atlantic stock. Although the species seems to occur mainly to the south within the study area, during the CETUS campaigns occurrences were also recorded in the northern states (Western Sahara and Mauritania).

4.5 Conclusions

This work reviewed existing cetacean occurrence data for the NWA, from Morocco to Liberia, and provides new records on cetaceans within the EEZs of Western Sahara, Mauritania and Senegal. At least 35 cetacean species have been reported in the NWA (plus a further two with unconfirmed occurrence in Morocco). While most research effort reported in the literature has been in Mauritania, new data collected were based on substantially higher effort in the Western Sahara. The new sightings data either document for the first time or confirm the occurrence of 10 taxa in Western Sahara. Moreover, summer months are poorly represented in the literature but were the most frequently surveyed months during the CETUS project. Hence, these recent surveys offer a new and important contribution to the limited knowledge of cetacean occurrence and distribution in the NWA, filling data gaps, mainly in the Western Saharan EEZ. Furthermore, in published works, records come mainly from coastal surveys and specimens found on the beach, while in the CETUS project, the survey effort was mostly in offshore areas of the EEZs.

The most frequently sighted species during CETUS campaigns, the sperm whale, is categorized as “Vulnerable” and, prior to the moratorium of commercial whaling, was a target for commercial whaling off NWA (Weir & Pierce, 2013). The region is probably a hotspot for sperm whale populations in the eastern north Atlantic. In terms of biomass, sperm whales are probably the most important toothed whale species in the region, and may have an important influence on ecosystem functioning (Morissette *et al.,* 2010).

Regular dedicated surveys are needed to monitor abundance and fine-scale distribution of cetaceans, and there remains a need to assess the level of impact of threats to cetacean populations in the area. Southern areas of the NWA (i.e. Sierra Leone and Liberia) are clearly less surveyed and still represent data gaps in terms of knowledge of cetacean occurrence. In fact, while this work contributes with new data and more information on cetacean distribution in the area, the lack of records of a species in a country does not constitutes evidence of its absence, and, for conservation purposes, it is important to keep an updated cetacean inventory.

Van Waerebeek *et al.* (2000; 2003), also UNEP-CMS (2007a; 2007b; 2008; 2012), identified various conservation issues that need to be addressed and provided several recommendations. One species of concern is the endemic Atlantic humpback dolphin, which has a restricted distribution and may be severely impacted by the rapid coastal development occurring in the NWA (UNEP-CMS 2007a; Van Waerebeek *et al.,* 2003; 2008; Weir *et al.*, 2010).

Overall, there is evidently a need to reconcile the rapid socio-economic development occurring on the coast - and the imperative to provide fish for human consumption - with the need to ensure sustainable utilization of the area´s living marine resources and to reduce or eliminate the increasing threats to cetaceans and their habitats (UNEP-CMS, 2008; Weir and Pierce, 2013).

4.6 Acknowledgements

A special acknowledgement to all CETUS volunteers for their contribution and dedication during the monitoring campaigns. We are extremely grateful to TRANSINSULAR that provided all the logistic support, and to the ship crews for their hospitality. We also thank Caroline Weir for assisting in cross-checking the literature records; and the two anonymous reviewers whose suggestions were extremely helpful in improving the manuscript.

This study was conducted within a PhD program funded by the Portuguese national funding agency for science, research and technology (FCT) under the grant SFRH/BD/100606/2014.

4.7 References

Baines M.E. and Reichelt M. (2014). Upwellings, canyons and whales: An important winter habitat for balaenopterid whales off Mauritania, northwest Africa. *Journal of Cetacean Research and Management*, 14: 57-67.

Bamy I.L., Van Waerebeek K., Bah S.S., Dia M., Kaba B., Keita N. and Konate S. (2010). Species occurrence of cetaceans in Guinea, including humpback whales with southern hemisphere seasonality. *Marine Biodiversity Records*, 3: e48.

Benchoucha S., Tamsouri M.N. and El Aamri F. (2018). First stranding record of *Kogia sima* (Owen, 1866) in Morocco (Strait of Gibraltar). *Journal of Black Sea*, 24(2): 175-179.

Boisseau O., Matthews J., Gillespie D., Lacey C., Moscrop A. and El Ouamari N. (2007). A visual and acoustic survey for harbour porpoises off North-West Africa: further evidence of a discrete population. *African Journal of Marine Science*, 29(3): 403-410.

Boisseau O., Lacey C., Lewis T., Moscrop A., Danbolt M. and McLanaghan R. (2010). Encounter rates of cetaceans in the Mediterranean Sea and contiguous Atlantic area. *Journal of the Marine Biological Association of the United Kingdom*, 90(8): 1589-1599.

Bowman Bishaw Gorham. Chinguetti Development Project. (2003). *Seabird and Cetacean Surveys in the Vicinity of the Chinguetti Oil Field, Offshore Mauritania*. Woodside Australian Energy, 66pp.

Caddell J.R. (2009). International Law and the Regulation of Cetaceans: An Analysis of the Role of Global and Regional Instruments and Institutions in the Conservation of Marine Living Resources. Doctorate thesis (PhD). Cardiff University, Cardiff, Wales. 373 pp.

Caldeira R.M.A., Groom S., Miller P., Pilgrim D. and Nezlin N.P. (2002). Sea-surface signatures of the Island mass effect phenomena around Madeira Island, Northeast Atlantic. *Remote Sensing of Environment,* 80: 336-360.

Camphuysen C.J. (2000). *Seabirds and marine mammals off West Africa – Responses 2000 cruise report*. Netherlands Institute for Sea Research, 50pp.

Camphuysen C.J., van Spanje T.M. and Verdaat H. (2012). *Ship based seabird and marine mammal surveys off Mauritania, Nov-Dez 2012 – cruise report*. Mauritanian Institute for oceanographic research and fisheries - IMROP, 73pp.

Correia A.M., Tepsich P., Rosso M., Caldeira R. and Sousa-Pinto I. (2015). Cetacean occurrence and spatial distribution: Habitat modelling for offshore waters in the Portuguese EEZ (NE Atlantic). *Journal of Marine Systems,* 143: 73-85.

Consentino A.M. and Fisher S. (2016). The Utilization of Aquatic Bushmeat from Small Cetaceans and Manatees in South America and West Africa. *Frontiers in Marine Science,* 3: 163.

Corten A. (2014). EU–Mauritania fisheries partnership in need of more transparency. *Marine Policy*, 49: 1-11.

Cunha H.A., de Castro R.L., Secchi E.R., Crespo E.A., Lailson-Brito J., Azevedo A.F., *et al.* (2015). Molecular and Morphological Differentiation of Common Dolphins (*Delphinus* sp.) in the Southwestern Atlantic: Testing the Two Species Hypothesis in Sympatry. *PLoS ONE,* 10(12): e0145354.

Djiba A., Bamy I.L., Bilal A.S.O. and Van Waerebeek K. (2015). Biodiversity of cetaceans in coastal waters of northwest Africa: new insights through platform-of-opportunity visual surveying in 2011-2013. *IOC Technical Series,* 115: 283-297.

Expósito C.G. and Qninba A. (2010). Observations de Tortues, d'Oiseaux et de Mammifères marins au large des côtes atlantiques marocaines durant l'automne 2010. *Bulletin de l'Institut Scientifique, Rabat, section Sciences de la Vie,* 32 (2): 129-132.

FAO (2013) *Marine protected areas - Country case studies on policy, governance and institutional issues*. In: FAO Fisheries and Aquaculture Technical Paper. Food and Agriculture Organization of the United Nations. Available at http://www.fao.org/publications/card/en/c/452ece29-e667-5f47-92a5-8a0e8eb5cbab/ (accessed November 2018).

Fertl D., Jefferson T.A., Moreno I.B., Zerbini A.N. and Mullin K.D. (2003). Distribution of the Clymene dolphin (*Stenella clymene*). *Mammal Review,* 33: 253-271.

Fontain M.C. (2016). Harbour porpoises, *Phocoena phocoena*, in the Mediterranean Sea and adjacent regions: biogeographic relicts of the last glacial period. *Advances in Marine Biology,* 75: 333-358.

Gray H. and Van Waerebeek K.V. (2011). Postural instability and akinesia in a pantropical spotted dolphin, *Stenella attenuata*, in proximity to operating airguns of a geophysical seismic vessel. *Journal for Nature Conservation*, 19: 363-367.

Hammond P.S. and Lockyer C. (1988). Distribution of killer whales in the eastern North Atlantic. *Rit Fiskideildar*, 11: 24-41.

Hazevoet C.J. and Wenzel W. (2000). Whales and dolphins (Mammalia, Cetacea) of the Cape Verde Islands with special reference to the Humpback Whale *Megaptera novaeangliae* (Borowski, 1871). *Contributions to Zoology,* 69(3): 197-211.

Hazevoet C.J., Monteiro V., López P., Varo N., Torda G., Berrow S. and Gravanita B. (2010). Recent data on whales and dolphins (Mammalia: Cetacea) from the Cape Verde Islands, including records of four taxa new to the archipelago. *Zoologia Caboverdiana,* 1: 75-99.

Hazevoet C.J., Gravanita B., López Suárez P. and Wenzel F.W. (2011). Seasonality of humpback whale *Megaptera novaeangliae* (Borowski, 1781) records in Cape Verde seas: evidence for the occurrence of stocks from both hemispheres. *Zoologia Caboverdiana,* 2(1): 25-29.

Jefferson T.A., Curry B.E., Leatherwood S. and Powell J.A. (1997). Dolphins and porpoises of West Africa: a review of records (Cetacea: Delphinidae, Phocoenidae). *Mammalia,* 61: 87-108.

Jung J., Mullié W.C., Van Waerebeek K., Wagne M.M., Bilal A.S.O., Sidaty Z.E.O., *et al.* (2016). Omura's whale off West Africa: autochthonous population or inter-oceanic vagrant in the Atlantic Ocean? *Marine Biology Research,* 12: 66-75.

Kaschner K., Quick N.J., Jewell R., Williams R. and Harris C.M. (2012). Global coverage of cetacean line-transect surveys: status quo, data gaps and future challenges. *PLoS One*, 7(9): e44075.

Leeney R.H., Weir C.R., Campredon P., Regalla A. and Foster J. (2016). Occurrence of Atlantic humpback (*Sousa teuszii*) and bottlenose (*Tursiops truncatus*) dolphins in the coastal waters of Guinea-Bissau, with an updated cetacean species checklist. *Journal of the Marine Biological Association of the United Kingdom*, 96(4): 933-941.

MacLeod C.D., Mandleberg L., Schweder C., Bannon S.M. and Pierce G.J. (2008). A comparison of approaches for modelling the occurrence of marine animals. *Hydrobiologia*, 612: 21-32.

Mason E. (2009). High-resolution Modelling of the Canary Basin Oceanic Circulation. Doctorate thesis (PhD) in Oceanography, Universidad de Las Palmas de Gran Canaria, Las Palmas de Gran Canaria. 245 pp.

Moore J.E., Cox T.M., Lewison R.L., Read A.J., Bjorkland R., McDonald S.L., *et al.* (2010). An interview-based approach to assess marine mammal ans sea turtle captures in artisanal fisheries. *Biological Conservation,* 143: 795-805.

Masski H. and de Stéfanis R. (2015). Cetaceans of the Moroccan coast: information from areconstructed strandings database*. Journal of the Marine Biological Association of the United Kingdom,* 1-9.

Morissette L., Kaschner K. and Gerber L.R. (2010). Ecosystem models clarify the trophic role of whales off Northwest Africa. *Marine Ecology Progress Series,* 404: 289-302.

Mullié W.C., Wagne M.M., Elmamy C.A.A., Yahya F.M., Veen J. and Van Waerebeek K. (2013). *Large number of stranded harbour porpoises Phocoena phocoena as by-catch victims in Mauritania.* Scientific Committee Document of International Whaling Commission, Jeju, Korea, June 2013. SC/65a/HIM03, 5pp.

Nagel P. and Gray T. (2012). Is the EU's Fisheries Partnership Agreement (FPA) with Mauritania a genuine partnership or exploitation by the EU? *Ocean & Coastal Management,* 56: 26-34.

Nielsen N.H., Teilmann J., Sveegaard S., Hansen R.G., Sinding M.H.S, Dietz R. and Heide-Jørgensen M.P. (2018). Oceanic movements, site fidelity and deep diving in harbour porpoises from Greenland show limited similarities to animals from the North Sea. *Marine Ecology Progress Series,* 597:259-272.

Notarbartolo di Sciara G., Barbaccia G. and Azzellino A. (1997). Birth at sea of a false killer whale, Pseudorca crassidens. *Marine Mammal Science*, 13(3):508-511.

Pinela A.M., Borrel A., Cardona L. and Aguilar A. (2010). Stable isotope analysis reveals habitat partitioning among marine mammals off the NW African coast and unique trophic niches for two globally threatened species. *Marine Ecology Progress Series,* 416: 295-306.

Perrin W.F. and Van Waerebeek K. (2012). *The small-cetacean fauna of the west coast of Africa and Macaronesia: diversity and distribution*. Convention on the Conservation of Migratory Species of Wild Animals, Bonn, Germany, CMS Technical Series, 26, 7-17.

Reiner F. (1980). First record of an Antillean beaked whale, *Mesoplodon europaeus* Gervais 1855, from República Popular da Guiné-Bissau. *Memórias do Museu do Mar, Série Zoológica,* 1(8): 1-8.

Reiner F., dos Santos M.E. and Wenzel F.W. (1996). Cetaceans of the Cape Verde Archipelago. *Marine Mammal Science,* 12(3): 434-443.

Robards M.D. and Reeves R.R. (2011). The global extentand character of marine mammal consumption by humans (1970-2009). *Biological conservation,* 144: 2770-2786.

Robineau D. and Vely M. (1993). Stranding of a specimen of Gervais' Beaked whale (*Mesoplodon europaeus*) on the coast of west Africa (Mauritania). *Marine Mammal Science,* 9(4): 438-440.

Robineau D. and Vely M. (1998). Ces cétacés des côtes de Mauritanie (Afrique du Nordouest). Particularltés et variations spatio-temporelles de repartition: role des facteurs oceanographiques. *Revue d Ecologie (Terre Vie),* 53.

Rojo–Nieto E., Álvarez–Díaz P.D., Morote E., Burgos–Martín M., Montoto–Martínez T., Sáez–Jiménez J. and Toledano F. (2011). Strandings of cetaceans and sea turtles in the Alboran Sea and Strait of Gibraltar: a long–term glimpse at the north coast (Spain) and the south coast (Morocco). *Animal Biodiversity and Conservation,* 34(1): 151-163.

Russell G., Sánchez-Cabanes A. and Nimak-Wood M. (2018). The autumn occurrence and spatial distribution of cetaceans in the waters of Mauritania during a geophysical survey in 2012. *African Journal of Marine Science,* 40(4): 371-381.

Sala I., Caldeira R.M.A., Estrada-Allis S.N., Froufe E. and Couvelard X. (2013). Lagrangian transport pathways in the northeast Atlantic and their environmental impact. *Limnology and Oceanography – Fluids & Environments,* 3: 40-60.

Satia B.P. (2016). An overview of the large marine ecosystem programs at work in Africa today. *Environmental Development,* 17: 11-19.

Subregional Fisheries Commission (2003). *Regional Strategy for Marine Protected Areas in West Africa.* Available at http://www.fibarguin.org/var/plain/storage/original/application/b5cbded23840211ccdf9d60838afb302.pdf (accessed November 2018).

Trouillet B., Guineberteau T., Bernardon M. and Le Roux S. (2011). Key challenges for maritime governance in West Africa: Fishery-based lessons from Guinea and Mauritania. *Marine Policy,* 35: 155-162.

Tsai C.H. and Mead J.G. (2018). Crossing the equator: a northern occurrence of the pygmy right whale. *Zoological Letters*, 4: 30.

Tudela S., Kai Kai A., Maynou F., El Andalossi M. and Guglielmi P. (2005). Driftnet fishing and biodiversity conservation: the case study of the large-scale Moroccan driftnet fleet operating in the Alboran Sea (SW Mediterranean). *Biological Conservation,* 121: 65-78.

Tulp I. and Leopold M.F. (2004). *Marine mammals and seabirds in Mauritanian waters. Pilot study April 2004 - Internal Report*. RIVO-Netherlands Institute for Fisheries Research - Animal Sciences Group. Wageningen UR, 42 pp.

UNEP-CMS (2007a). *Conservation Status of the Atlantic Humpback Dolphin, a Compromised Future?* Convention on the Conservation of Migratory Species of Wild Animals, Bonn, Germany, 10pp.

UNEP-CMS (2007b). *Conservation Status of the Clymene Dolphin in West Africa*. Convention on the Conservation of Migratory Species of Wild Animals, Bonn, Germany, 9pp.

UNEP-CMS (2008). *Action Plan for the Conservation of Small Cetaceans of Western Africa and Macaronesia. In: Annex II to the Memorandum of Understanding Concerning the Conservation of the Manatee and Small Cetaceans of Western Africa and Macaronesia*. CMS - Convention on the Conservation of Migratory Species of Wild Animals. Available at http://www.cms.int/en/document/action-plan-conservation-small-cetaceans-western-africa-and-macaronesia (accessed November 2018).

UNEP-CMS (2012). *Conserving cetaceans and manatees in the western African region.* In: CMS Technical Series No. 26. CMS - Convention on the Conservation of Migratory Species of

Wild Animals. Available at http://www.cms.int/en/publication/conserving-cetaceans-and-manatees-western-african-region-ts-no-26 (accessed November 2018).

Van Waerebeek K., André M., Sequeira M., Martín V., Robineau D., Collet A., *et al.* (1999). Spatial and temporal distribution of the minke whale, Balaenoptera acutorostrata (Lacépède, 1804), in the southern northeast Atlantic Ocean and the Mediterranean Sea, with reference to stock identity. *Journal of Cetacean Research and Management,* 1: 223-237.

Van Waerebeek K., Ndiaye E., Djiba A., Diallo M., Murphy P., Jallow A., *et al.* (2000*). A survey of the conservation status of cetaceans in Senegal, The Gambia and Guinea-Bissau.* UNEP/CMS Secretariat, Bonn, Germany, 80pp.

Van Waerebeek K., Barnett L., Camara A., Cham A., Diallo M., Djiba A., *et al.* (2003). *Conservation of Cetaceans in The Gambia and Senegal 1999-2001, and Status of the Atlantic humpback dolphin.* UNEP/CMS Secretariat, Bonn, Germany. 56 pp.

Van Waerebeek K., Bamy I.L., Jiddou A.M., Sequeira M., Diop M. and Ofori-Danson P.K., *et al.* (2008). *Indeterminate status of West African populations of inshore common bottlenose dolphins Tursiops truncatus cautions against opportunistic live-capture schemes.* Foundation du Banc d'Arguin, 9 pp.

Van Waerebeek K., Djiba A., Krastad J.O., Bilal A.S.O., Almeida A. and Mbye E.M. (2013). New evidence for a South Atlantic stock of humpback whales wintering on the Northwest African continental shelf. African Zoology 48(1): 177-186.

Van Waerebeek K., Uwagbae M., Segniagbeto G., Bamy I.L. and Ayissi I. (2017). New records of Atlantic humpback dolphin (*Sousa teuszii*) in Guinea, Nigeria, Cameroon and Togo underscore pressure from fisheries and marine bushmeat demand. *Revue d'Ecologie (Terre et Vie),* 72(2):192-205.

Weir C.R. (2010). A review of cetacean occurrence in West African waters from the Gulf of Guinea to Angola. *Mammal Review,* 40(1): 2-39.

Weir C.R. (2016). Atlantic humpback dolphins *Sousa teuszii* in the Saloum Delta (Senegal): distribution, relative abundance and photo-identification. *African Journal of Marine Science,* 38(3): 385-394.

Weir C. and Collins T. (2015). A review of the geographical distribution and habitat of the Atlantic humpback dolphin (*Sousa teuszii*). *Advances in Marine Biology,* 72: 79-117.

Weir C.R. and Pierce G.J. (2013). A review of the human activities impacting cetaceans in the eastern tropical Atlantic. *Mammal Review*, 43: 258-274.

Weir C.R., Coles P., Ferguson A., May D., Baines M., Figueirdo I., *et al.* (2014). Clymene dolphins (*Stenella clymene*) in the eastern tropical Atlantic: distribution, group size, and pigmentation pattern. *Journal of Mammalogy,* 95(6): 1289-1298.

CHAPTER V.

Paper published in the Journal of Marine Systems, 2015. doi: 10.1016/j.jmarsys.2014.10.016

5 Cetacean occurrence and spatial distribution: Habitat modelling for offshore waters in the Portuguese EEZ (NE Atlantic)

Ana M. Correia[1,2], Paola Tepsich[3,4], Massimiliano Rosso[3], Rui Caldeira[1], Isabel Sousa-Pinto[1,2]

[1]Interdisciplinary Centre of Marine and Environmental Research (CIIMAR). 4450-208 Matosinhos, Portugal.
[2]Department of Biology, Faculty of Sciences, University of Porto (FCUP). 4169-007 Porto, Portugal.
[3]CIMA Research Foundation. 17100 Savona, Italy.
[4]DIBRIS, University of Genova. 16145 Genova, Italy.

5.1 Abstract

In the Portuguese Exclusive Economic Zone (EEZ) (NE Atlantic), little survey effort dedicated to cetacean species has been carried out in offshore waters. As a consequence, data on their occurrence, distribution and habitat preferences is scarce. In this area, 48 sea surveys along fixed transects within Continental Portugal and Madeira Island were performed in 2012 and 2013, from July to October, using platforms of opportunity. We used an environmental envelope approach and GAM habitat models to identify the role of oceanographic, topographic and geographical variables in shaping cetacean distribution. Results demonstrate the richness of offshore waters in this area as in 10,668 nmi sampled, we recorded 218 sightings from at least nine cetacean species, resulting in an overall ER of 2.04 sightings / 100 nmi. The interaction of topographic and oceanographic features was shown to influence the distribution of the species/groups along the routes. Among the sighted species, only common dolphin showed a preference for coastal waters, while for all the other species high seas proved to be determinant. This result reinforces the need to address conservation issues in open ocean. This preliminary assessment showed the importance of the entire area for the distribution of different cetacean species and allowed the identification of several species/group specific potential suitable habitats.

Considering the Habitats Directive resolutions, ACCOBAMS priorities, EEZ extension for the area and Maritime Spatial Planning Directive, and the urgent need for management plans, we suggest that the sampling strategy here presented is a cost-effective method to gather valuable data, to be used to improve cetacean habitat models in the area.

Keywords: Management, Portuguese EEZ, high seas, cetacean distribution, habitat modelling

5.2 Introduction

The Portuguese Exclusive Economic Zone (PEEZ) is located in the Northeast of the Canary Basin (NE Atlantic Ocean). The occurrence of eddies often associated with numerous topographic structures (canyons and seamounts), confirms the complex and dynamic topographic–oceanographic system of this region (Mason, 2009). In the PEEZ, the main topographic features are the Madeira-Tore Rise, where the Josephine Seamount, which is the only offshore Marine Protected Area (MPA) in the PEEZ, is found, the Horseshoe Seamount Chain, which includes the Gorringe Bank, and the Mid-Atlantic Ridge. This last is known to particularly affect the Azorean ecosystems (Mason, 2009). From an oceanographic point of view, the predominant northeasterly trade winds (NTW), linked with the formation of upwelling systems occurring along the African and Portuguese coasts, interact with atmospheric high-pressure (Azores High): the latitudinal changes in these features are responsible for seasonal variation in mesoscale oceanic eddies (Mason, 2009). Dynamic environments, resulting from the interaction of topographic and oceanographic structures, are known to favour ecosystem richness and consequently, cetacean occurrence (Ballance *et al.*, 2006; Hoyt, 2011; Redfern *et al.*, 2006). However, to date, knowledge on cetacean occurrence, spatial distribution and habitat preferences in the PEEZ is scarce, especially for offshore waters. Published data are mainly restricted to coastal areas and to a few regional hotspots (e.g. the Azores and Madeira archipelagos), where data collection is limited to a few miles from the coast and usually encompassing only certain species (e.g. *Tursiops truncatus*) (Alves *et al.,* 2013; 2014; Augusto, 2007; Brito & Sousa, 2011; Moura *et al.*, 2012; Silva *et al.*, 2003; Visser *et al.*, 2011). As a consequence, most of the MPAs for cetaceans are also limited to coastal areas (Hoyt, 2011).

Whereas conservation efforts focus on coastal regions, anthropogenic impacts and the non-sustainable use of marine resources is increasingly affecting offshore areas (Evans *et al*., 2012; Hooker & Gerber, 2014). Therefore, there is a growing urgency for the management of high seas, a challenging task for both decision-makers and the scientific community. On one hand, marine environments are very complex and dynamic and change in time and space (Evans *et al*., 2012) and lack administrative boundaries; on the other hand, the logistic requirements for monitoring high seas are very demanding (Kiszka *et al.*, 2007; Viddi *et al.,* 2010). To date, one of the main approaches to study and manage such complex habitats focuses on the ecology and conservation status of top predators, as they often act as *flagship species*, being politically used to attract funding due to their high public profile, *umbrella species*, as due to their position in the trophic ladder and their wide-range distribution, protecting their habitat will directly lead to the protection of several other species, *keystone species*, by significantly impacting their ecosystems and chains of energy in the trophic ladder, and *indicator species,* being highly sensitive and responsive to ecosystem health (Hooker and Gerber, 2014; Sergio *et al.*, 2006; 2008). Though controversial and context-dependent, it has been proved that the management of protected areas based on top predators distribution is highly efficient, leading to higher biodiversity levels and more ecosystem benefits (Sergio *et al.*, 2006; 2008). Consequently, protecting cetaceans and their habitats is a priority issue in marine management plans, as measures directed to their protection tend to act as measures for the management of seas in general (Hooker *et al.*, 2011; Hooker & Gerber, 2014; Sergio *et al.*, 2006; 2008). Thus, there is an emerging need for baseline knowledge on cetacean occurrence and habitat preferences in high seas. In particular, for the PEEZ, this knowledge is demanding, considering that the area is under the scope of the Habitats Directive (Directive 92/43/CEE) and has been recently included in the “Agreement on the Conservation of

Cetaceans of the Black Sea, Mediterranean Sea and contiguous Atlantic Area” (ACCOBAMS, 2013). Moreover, considering the planned and accepted extension (Estrutura de Missão para a Extensão da Plataforma Continental – EMEPC, 2013) (Figure 1), the PEEZ will soon be the biggest European EEZ. The need for Portugal to develop efficient and informed management plans for its EEZ has been recently reinforced also by the recent European Union Directive (Directive 2014/89/EU) that has established a framework for maritime spatial planning for Member States to manage human activities, uses and interests in their maritime areas. This Directive followed an Impact Assessment (European Commission, 2013) that documented the expansion of commercial activities in marine areas, such as fisheries, shipping, dredging, oil exploitation, tourism, mineral extraction or recently wind energy and offshore marine aquaculture, and identified several problems in managing the use of the marine space.

With the aim of closing the existing data gap about cetacean distribution in the area, line-fixed transects within the Mainland and Madeira Island PEEZ (Figure 1) were sampled using cargo ships as Observation Platforms of Opportunity (OPOs) to obtain data on cetacean occurrence and distribution. Facing the logistic requirements of high seas monitoring, the use of OPOs is a cost-effective and recommended approach to obtain useful and preliminary data in less-surveyed areas (Kiszka *et al.*, 2007; Moura *et al.*, 2012; Redfern *et al.*, 2006; Viddi *et al.*, 2010). Occurrence data was then related to topographic, oceanographic and geographical variables, to investigate habitat preferences of the encountered species and to develop exploratory habitat models to predict where cetaceans are more likely to occur in space and time, and how this relates to the environmental conditions (McLeod *et al.*, 2008; Redfern *et al.*, 2006).

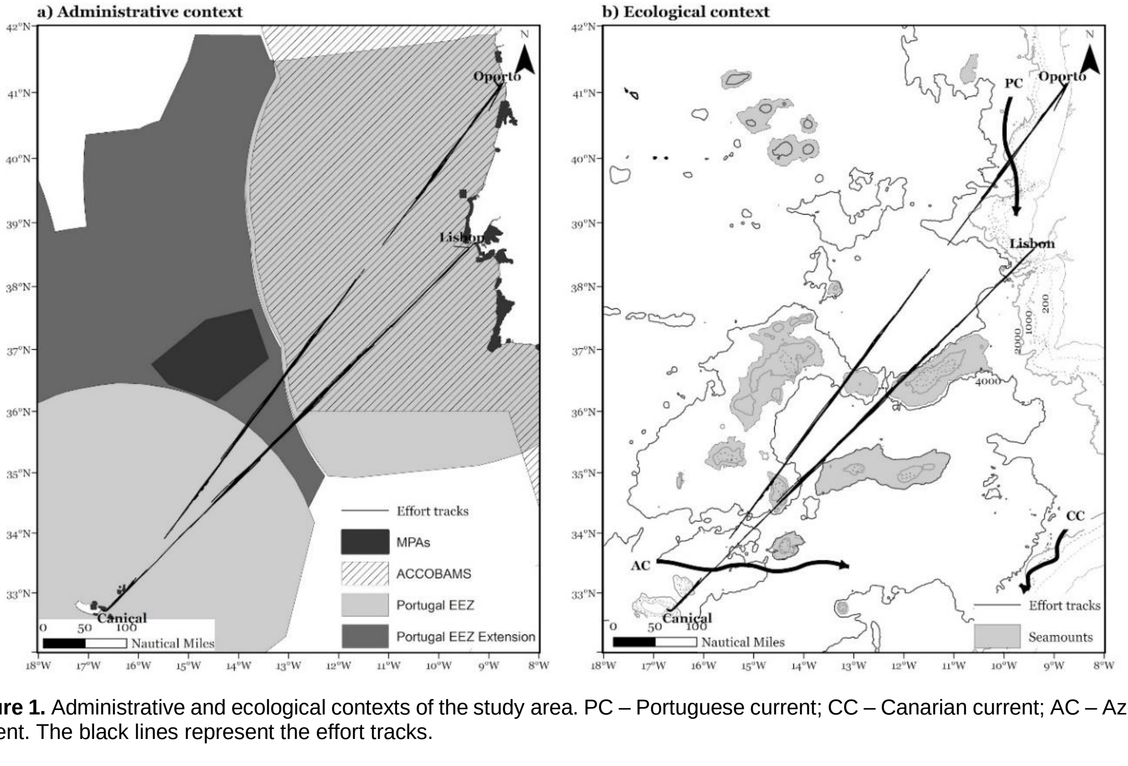


Figure 1. Administrative and ecological contexts of the study area. PC – Portuguese current; CC – Canarian current; AC – Azores current. The black lines represent the effort tracks.

5.3 Material and Methods

5.3.1 Sampled transects

From July to October in 2012 and 2013, a team of nine trained Marine Mammal Observers (MMOs) performed surveys weekly or every 15-days along fixed transects in the Mainland and Madeira Island PEEZ, using cargo ships from the Transinsular Company as OPOs (Figure 1). Considering the need of good weather conditions for performing cetacean monitoring, the summer season has been chosen as in the area, weather and sea-state condition during the rest of the year are not suitable for cetacean research. Though it was not possible to include June due to logistic constrains, the four monitored months encompass the overall summer season. Among the itineraries performed by the cargo ship company, the one sampled (Oporto – Lisbon – Caniçal – Lisbon – Caniçal – Oporto) crosses the area of highest habitat diversity (Figure 1), including different topographic systems (continental platform, abyssal plains, steep slope, seamounts and canyons) and a diversity of oceanographic features, such as several mesoscale eddies and two major currents, the Portugal and Azores currents (Mason, 2009) (Figure 1). The sampled route is thus assumed to represent the study area in terms of habitat diversity. The mean extent of the routes varied: 210 nmi (Oporto – Lisbon); 640 nmi (Lisbon – Caniçal) and 807 nmi (Caniçal – Porto). A sea survey was considered as a one-way journey from one port to another. Autocorrelation of the surveys did not present a problem for this study, as each survey was performed on non-consecutive days and the purpose was not to access the abundance of animals, but to investigate habitat preferences. The monitoring program resulted in a total of 48 sea-surveys (17 in 2012 and 29 in 2013), with 83.02% of the sampling effort within the current PEEZ and 100% in the future PEEZ (Figure 1).

5.3.2 Data collection

Sampling was performed aboard two tween cargo ships (“n/m Monte da Guia” and “n/m Monte Brasil”), 127 m long and 20 m wide and the cruise speed varied from 14 knots to 16 knots. Observers were placed at a height of 19.4 m above sea level in the navigation bridge and wings of the bridge. Since observation height affects the detectability of cetaceans, survey conditions were accessed in relation to the height of the platform used (Stockin *et al.*, 2008). Sampling effort was carried out from sunrise to sunset, and ceased with sea state or wind force above 4 (Douglas scale and Beaufort scale, respectively) or when the monitoring team could not remain in the navigation bridge. The mobile application “My Tracks” (Google, 2013) was used on a smartphone (Optimus Monte Carlo) with GPS to record the coordinates of the effort transects (recording points every five to ten seconds) and to register sightings and other sampling points.

The sampling protocol adopted was the one usually followed during line-transect sampling and was performed by two MMOs: each observer stood on one side of the vessel and monitored cetacean presence in the direction of the route to 90º seawards, with and without binoculars (Paralux Nemo deep Sea 7 × 50 mm). Observers changed sides every 60 min to avoid fatigue and to reduce observer detection capacity bias. Following detection, the species/group name, some distinct surface behaviours, as well as the distance and angle from the boat (calculated with the compass and scale of the binoculars) were recorded. For group size, a minimum, maximum and best estimate was registered when the exact number of animals could not be accessed; the best estimate was then used in the analysis. A sighting was considered to be an individual animal or a group of animals having the same activity, being in close proximity and swimming in the same direction (Kiszka *et al.*, 2007). Weather conditions (sea state on the Douglas scale, wind force on Beaufort scale, wind direction, visibility, sky cover and rainfall) were recorded at the beginning and at the end of the surveys and whenever they changed significantly. Vessel traffic was accessed visually, with and without binoculars, every hour and following every sighting. The data collected were logged into an Excel database and prepared for subsequent analysis.

Habitat variables for the study of habitat preferences were selected, considering the ecological meaning for cetacean occurrence, based on the state of the art for this topic (see referenced literature; and for reviews on the subject, see Ballance *et al.*, 2006 and Redfern *et al.*, 2006), the characteristics of the sampled area (Mason, 2009) and the data availability (satellite data) (Robinson, 2010). Table 1 presents the details of all the chosen variables.

Topographic variables were selected, obtained from GEBCO (2012) and EarthRef (2012), as well as geographical variables, as these have been shown to highly influence cetacean distribution (among others, Azzellino *et al.*, 2012; Esteban *et al.*, 2013; Kiszka *et al.*, 2007). Depth, slope and distance to coast (dist.coast) were chosen as topographic variables (Table 1). Moreover, taking into account the topographic characteristics of the region, the distance to seamounts (dist.sm) was considered, as these features are so unique and have been demonstrated to be important for several marine *taxa*, namely cetaceans (Morato *et al.*, 2008; 2010; Schlacher *et al.*, 2010). The seamounts were defined, based on the location of their centre given by EarthRef (2012) and their limits were defined at the base, when slope started to increase significantly compared to the surrounding area. Therefore, dist.sm is, in fact, the distance to the base of the seamounts. Finally, latitude (lat) was included as a non-environmental variable that accounts for the unknown ecological predictors, completing those that were considered (Pirotta *et al.*, 2011; Spyrakos *et al.*, 2011). Furthermore, given the large north–south extension of the sampled tracks, this geographical variable was important to encompass the spatial variability in the analysis (Certain *et al.*, 2008; Esteban *et al.*, 2013).

Oceanographic variables were also included, because these act as indicators of productive areas, and therefore, as good proxies for species distribution in marine habitats. Furthermore, habitat models that combine static and dynamic variables perform more efficiently (Ballance *et al.*, 2006). Therefore, chlorophyll-a (CHL) and sea-surface temperature (SST) were selected, since they are good proxies for productive areas in upwelling phenomena, characterised by cooler and nutrient-rich waters (Robinson, 2010). The area is mainly oligotrophic during the summer season, with peaks of CHL only in the coastal region due to coastal upwelling and river drainage. However, in the spring, the CHL bloom extends towards the open ocean (Correia, 2013). Thus, instead of the monthly data of CHL for the sampled months, data from the bloom month were used as a proxy for productivity. In fact, it was shown that bottom-up control may have a spatio-temporal lag, especially considering the intermediate trophic levels from phytoplankton to top predators (Frederiksen *et al.*, 2006; Grémillet *et al.*, 2008), and therefore, blooms of phytoplankton may influence future distribution of top predators, even though this relation has to be carefully interpreted (Grémillet *et al.*, 2008). To determine the exact bloom month for each year, monthly mean maps and mean values for the study area in 2012 and 2013 were checked in the Giovanni platform (NASA, 2013a), and CHL monthly data for March 2012 and April 2013 were selected for data analysis. Both SST (night at 4 µm) and CHL were derived from the satellite MODIS – Aqua Mapped data from NASA (2013b). Finally, Mean Sea Level Anomalies (MSLA) were considered. Although the study of altimetry as a proxy for cetacean distribution is in its infancy, it has already been linked with cetacean occurrence and movements, namely with the movements of the one-tagged beaked whale (Baird *et al.*, 2011; Davis *et al.*, 2002). Additionally, MSLA are especially important to understand the dynamics and patterns of circulation, which are crucial when investigating the habitat in high seas (Robinson, 2010). The MSLA data were obtained from Ssalto/Duacs Near-Real Time gridded products derived from Topex/Poseidon satellite and provided by AVISO (AVISO, 2013). For both SST and MSLA, data from the sampled months (July to October) was used. These variables allow the identification of oceanographic processes that can lead to subsuperficial blooms during the sampling season, otherwise not detected by the CHL satellite data for this period. All data were visualised, computed and processed in MatLab R2012a (MATLAB version 7.14.0, 2012) and in ArcGIS 10 (ESRI, 2011). Created rasters and shapefiles were referenced with the coordinate system GCS_WGS_1984 (EPSG: 4326) and projected with World_Mercator (EPSG: 3395).

Table 1. Characteristics of the habitat variables used in spatial analysis.

Variables	Name used	Source	Unit	Type	Spatial Resolution	Temporal Resolution
Depth	Depth	GEBCO (2012)	m	Static	30 arc-second	-
Slope	Slope	GEBCO (2012)	Degree	Static	30 arc-second	-
Distance to coast	Dist.coast	-	km	Static	30 arc-second	-
Distance to seamounts	Dist.sm	GEBCO (2012) and EarthRef (2013)	km	Static	30 arc-second	-

Latitude	Lat	GPS	UTM Northing (m)	Static	-	-
Chlorophyll - a	CHL	MODIS (NASA, 2013a)	mg/m**3	Dynamic	4km	Monthly
Sea surface temperature	SST	MODIS (NASA, 2013b)	ºC	Dynamic	4km	Monthly
Maps of sea level anomaly	MSLA	AVISO (2013)	cm	Dynamic	1/3ºx1/3º, Mercator	Monthly

5.3.3 Presence analysis

Presence analysis was conducted considering the Encounter Rates (ER). Total ER was defined as follows:

ER = (n/D) x 100,

where n is the total number of sightings on-effort and D is the distance sampled on-effort in nautical miles (nmi). This measure was computed for each species using the total distance sampled, to give an overview of the species occurrence in the area.

Considering the higher number of sightings, the analysis proceeded for common dolphins, bottlenose dolphins and sperm whales. Other sightings were grouped as either baleen or beaked whales, which have common feeding habits within the group (Cox *et al*., 2006; MacLeod *et al*., 2006; Viddi *et al.*, 2010). Besides, 25 out of 30 sightings of baleen whales were identified as minke whales, with only five non-identified sightings; and beaked whales species were demonstrated to have very similar habitat preferences and are often analysed as a group (Cox *et al.,* 2006; MacLeod *et al*., 2008). Additionally, and even though beaked whales share the same prey with sperm whales (cephalopods), and a toothed whale group was considered, they were distinguished, as segregation among them has been suggested for the Mediterranean Sea (Azzellino *et al*., 2008; Tepsich *et al.*, 2014). This division of sightings was used to compute ER, generate distribution maps in ArcGIS 10 (ESRI, 2011) and for subsequent analysis.

5.3.4 Habitat preference analysis

Habitat preference was investigated following a used versus available habitat approach (Pearce & Boyce, 2006; Elith & Leathwick, 2009). Spatial Analyst tools from ArcGIS 10 (ESRI, 2011) were used to prepare the data: rasters of the habitat variables were sampled for species / group sightings, resulting in the used habitat; a set of equidistant points (2.5 nmi) was created along the effort tracks and was considered to represent the available habitat. Sampling effort was also analysed through a line-density raster, created following the parameters used in Viddi *et al.* (2010) – the total length of the effort tracks fell within a radius of 5 km in the neighbourhood of each output raster cell of 1 × 1 km. Since there was a large heterogeneity in the sampling effort along transects that possibly affected the habitat preferences analysis, an effort threshold was set. Therefore, only sightings and available habitat points that fell in a 1 × 1 km cell with more than 1 km sampled were kept for the subsequent analysis.

Two different techniques were then performed in R 2.15.0 Software (R Development Core Team, 2012): quantiles analysis and habitat modelling (mgcv package). For quantiles analysis, boxplots were computed for each group and for the available habitat (Austin, 2007; Elith & Leathwick, 2009; Kiszka *et al.*, 2007; Pearce & Boyce, 2006) and habitat ranges were investigated.

For habitat modelling, correlation among variables was verified, and whenever they had a positive correlation (Pearson correlation ≥ 0.8), one of them was excluded from the model (Marubini *et al.*, 2009). Then, resource selection function techniques (RSF) were chosen as they are recommended to determine animal distribution as a function of a variable, allowing used and available habitats to be distinguished (Elith & Leathwick, 2009). Generalised Additive Models (GAMs) were chosen over Generalised Linear Models (GLMs), as they are more flexible and the accuracy is less affected by a low prevalence of animals (Barbet-Massin *et al.*, 2012) and they are widely used to explain cetacean distribution (Brotons *et al.*, 2004; Elith & Leathwick, 2009; MacLeod *et al.*, 2008; Pearce & Boyce, 2006; Torres *et al.*, 2008; Viddi *et al.*, 2010; among others). A quasibinomial distribution (used / available) and logit link function were used. The model-fitting process followed the one used in Viddi *et al.* (2010). All variables were considered for the first fitting, followed by a backward selection to obtain the best-fitting models based on the generalised cross-validation (GCV) scores. The process consisted in taking the variable with the least significant *p*-value from the saturated first fitted model. If this led to a model with a lower GCV, this model was retained. Otherwise, the variable was maintained and the next less-significant variable was taken. The process was repeated until the best GCV score was obtained. The maximum number of splines was set to eight to prevent an overfitted model (Dalla Rosa *et al.*, 2012). The scale parameter was set to -1.0 and gamma to 1.4 to deal better with overdispersion in the data (Viddi *et al.*, 2010). A weight parameter was included in the models, corresponding to the number of animals sighted (best estimate). Since the size of groups of dolphin species had a wide range (minimum and maximum) and a high uncertainty of the best estimate, the number of animals was divided into three categories and three weights were attributed accordingly: a small group – from one to five animals (weight = 1); a medium group – from six to 20 animals (weight = 2); a large group – more than 20 animals (weight = 3).

To understand better the relationships illustrated by GAM plots, a zero line was used to define a positive effect of the predictors on cetacean occurrence (function > 0), in a process called *GAMvelope* by Torres *et al.* (2010).

5.4 Results

5.4.1 Presence analysis

From the 48 surveys, a total of 10,668 nmi were sampled, however the effort was not equally distributed among routes, neither in space nor time (Figure 1 and Table 2).

Table 2. Total observation effort (nautical miles) in the sampled months.

	2012				2013				
Routes	**July**	**August**	**September**	**October**	**July**	**August**	**September**	**October**	**Total**
Oporto-Lisbon	43	0	0	0	0	42	0	37	122
Lisbon-Caniçal	865	531	615	535	1277	889	1135	803	6651
Caniçal-Oporto	265	1200	626	375	208	391	301	530	3896
Total	1173	1731	1241	910	1485	1322	1436	1370	10668

A total of 218 sightings was registered, resulting in a total ER of 2.04, and accounting for at least nine cetacean species: Common dolphin (*Delphinus delphis*), Bottlenose dolphin (*Tursiops truncatus*), Minke whale (*Balaenoptera acutorostrata*), Sperm whale (*Physeter catodon*), Spotted dolphin (*Stenella frontalis*), Cuvier's beaked whale (*Ziphius cavirostris*), Striped dolphin (*Stenella coeruleoalba*), Pilot whale (*Globicephala sp.*) and Pigmy sperm whale (*Kogia breviceps*) (Table 3). Considering the groups used for analysis, both dolphin species had higher ERs, with 0.35 sightings / 100 nmi for common dolphins and 0.30 sightings / 100 nmi for bottlenose dolphins; followed by baleen and beaked whales, with 0.28 and 0.26 sightings / 100 nmi, respectively; and finally sperm whales, with the lowest ER of 0.13 sightings / 100 nmi.

Table 3. Total number of sightings, group size and ER for all identified and non-identified species.

Species	Sightings	Group size mean (±σ, range)	Total ER
Common dolphin	37	14.22 (±22.90, 1-120)	0.35
Bottlenose dolphin	32	8.88 (±6.55, 2-35)	0.30
Minke whale	25	1.60 (±0.85, 1-4)	0.23
Sperm whale	14	1.29 (±0.59, 1-3)	0.13
Spotted dolphin	13	10.69 (±6.04, 2-25)	0.12
Cuvier's beaked whale	13	2.62 (±1.73, 1-6)	0.12
Striped dolphin	3	10.00 (±1.63, 8-12)	0.03
Pilot whale	2	(1,5)	0.02
Pigmy sperm whale	1	(1)	0.01
NI dolphins	58	7.28 (±8.16, 1-40)	0.54
NI beaked whales	15	2.60 (±1.36, 1-5)	0.14

NI baleen whales	5	1.20 (±0.40, 1-2)	0.05
Total	218	6.70 (±11.27, 1-120)	2.04

NI – Non-identified; σ – Standard Deviation; ER – Encounter Rate (number of sightings / 100 nautical miles).

By mapping the sightings, the distribution of the groups was clearly heterogeneous (Figure 2): dolphin species were the only species to occupy the northern latitudes, very close to the continental coast, whereas baleen and beaked whales were more prevalent in middle latitudes, contrasting with the frequent occurrences of sperm whales in the waters very close to the coast of Madeira Island.

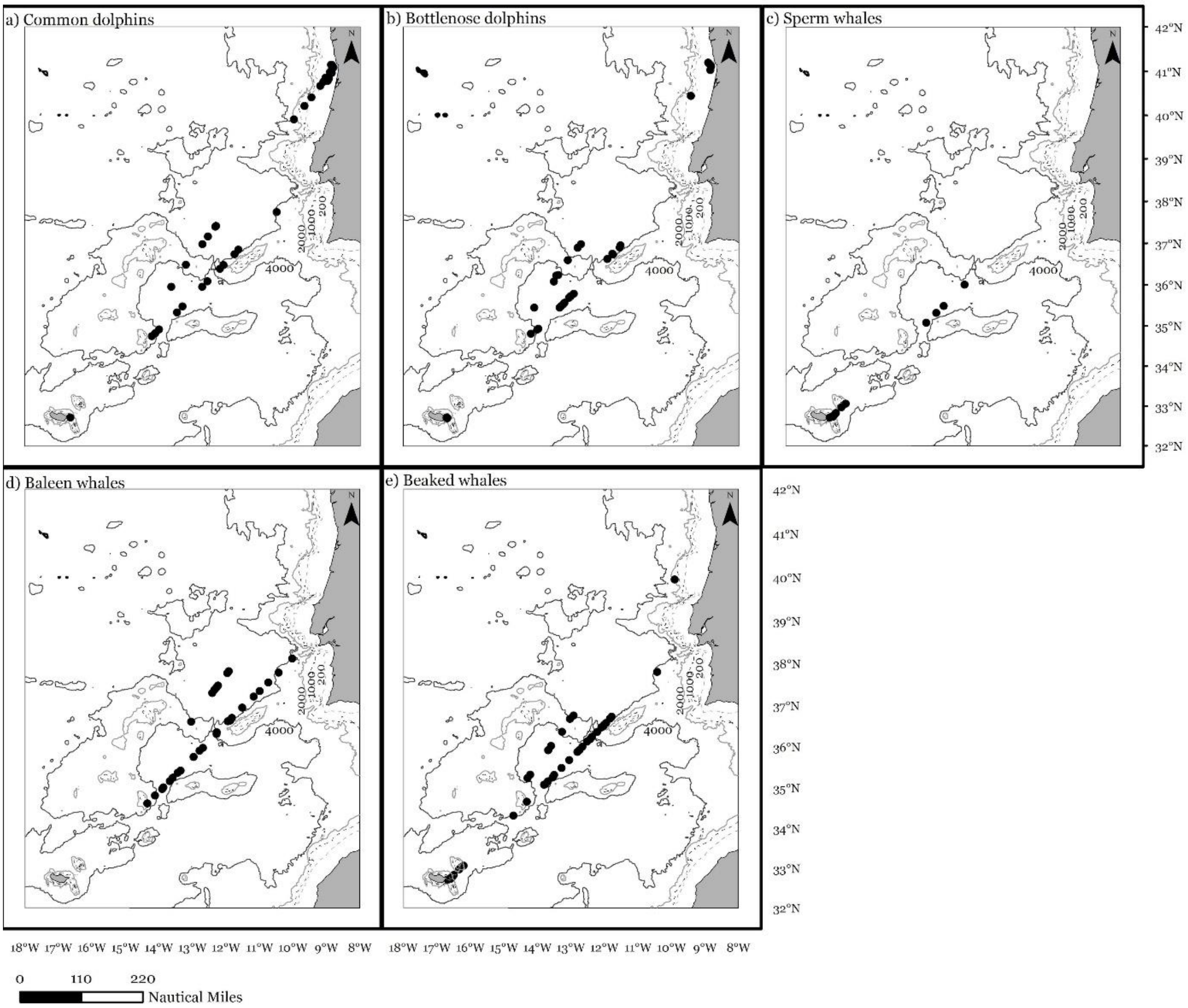


Figure 2. Cetacean sighting distributions in the study area according to the species / groups used for subsequent habitat preferences and modelling analysis.

5.4.2 Habitat preference analysis

By applying the effort threshold, there was a decrease in the number of sightings. Therefore, the sightings used for this analysis were: common dolphins – 16; bottlenose dolphins – 19; sperm whales –14; baleen whales – 15; beaked whales – 22.

The boxplots highlight the characteristics of the habitat sampled (available habitat) considering the variables analysed. Given the extent of the transects, effort was higher in: high depths (median: 4,326 m; range: 0 m to 5,147 m), low slopes (median: 1.19º, range: 0º to 24.26º), far from the coast (median: 402.02 km, range: 0.81 km to 513.81 km), close to seamounts (median: 24.59 km; range: 0 to 266.26 km), in middle to southern latitudes (median: 4,248,099 m; range: 3,831,511 m to 4,999,631 m), oligotrophic waters (CHL median: 0.19 mg/m^3; range: 0.11 mg/m^3 to 10.28 mg/m^3), warmer temperatures (median: 21.77ºC; range: 15.84ºC to 26.68ºC) and across a wide range of MSLA (median: 7.61 cm; range: -6.20 cm to 16.85 cm) (Figure 3).

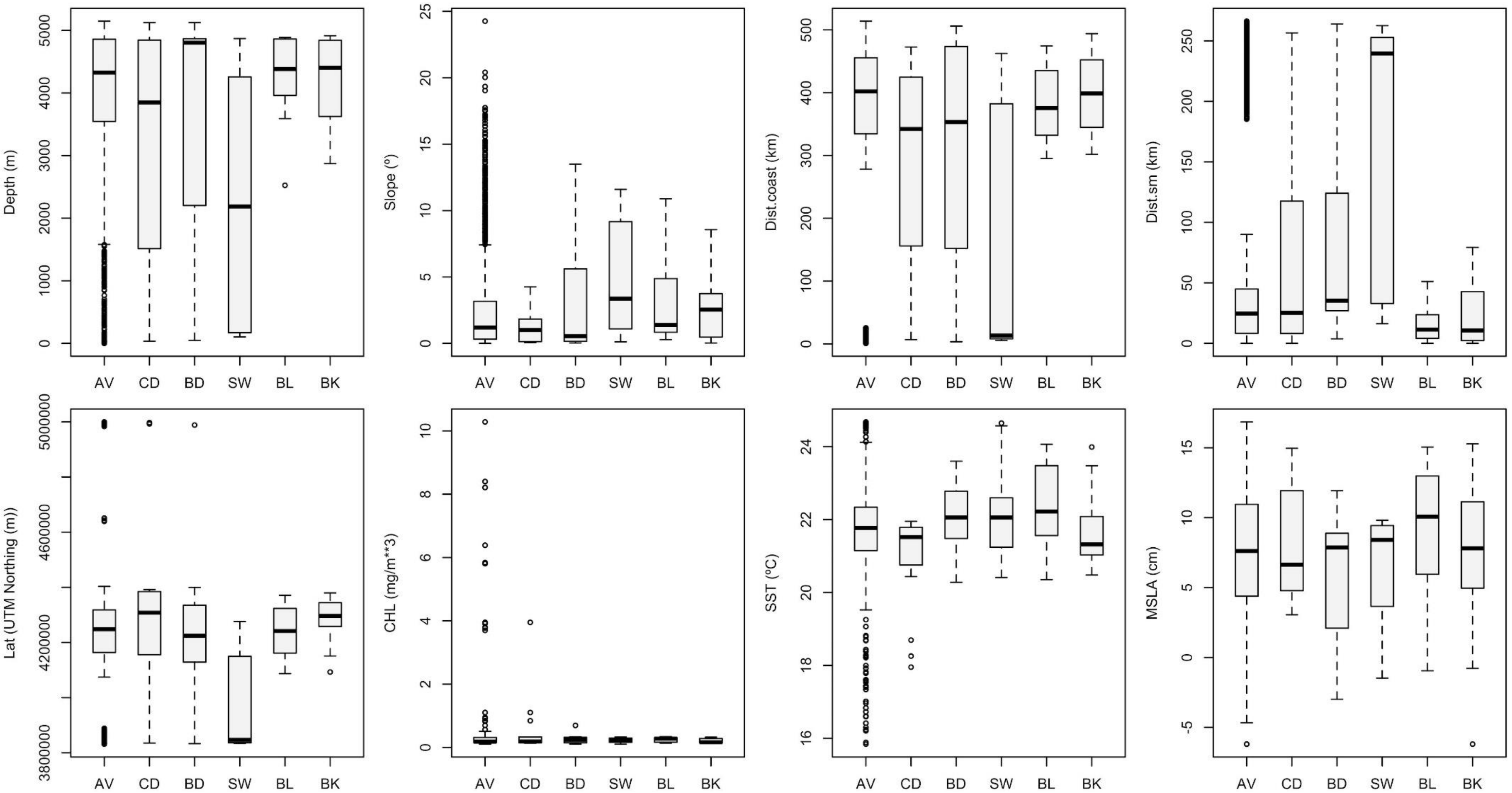


Figure 3. Habitat range of cetacean sightings concerning the variables used as proxies for distribution in the habitat modelling. AV – Available habitat; CD – Common dolphins; BD – Bottlenose dolphins; SW – Sperm whales; BL – Baleen whales; BK – Beaked whales.

This environmental envelope approach demonstrates the habitat preferences of the species/groups, when complemented with the modelling approach. In this final analysis, dist.coast was excluded from the fitting in the modelling process as it was positively correlated with both depth and dist.sm in all species/groups. Therefore, only seven variables were considered for fitting. Missing satellite data resulted in a reduced number of sightings for common dolphins, and only 12 out of 16 sightings were included in the modelling process (Table 4).

Table 4. Best GAM model results obtained.

Species/Groups Parameters	Estimate	edf	se	t-value	F-value	p-value	Deviance explained (%)	r^2	GCV score
Common dolphins									
Intercept	-2058.00		646.60	-3.18		0.001			
Smoother terms									
Depth		7.00			22.55	<0.001			
Slope		5.65			20.67	<0.001			
Dist.sm		4.78			12.21	<0.001			
Lat		7.00			32.14	<0.001			
CHL		2.11			20.75	<0.001			
SST		6.52			37.65	<0.001			
MSLA		6.95			25.51	<0.001			
Best final model (n=2659; 12 presences): CD~s(Depth)+s(Slope)+s(Dist.sm)+s(Lat)+s(CHL)+s(SST)+s(MSLA)							59.90	0.40	0.04
Bottlenose dolphins									
Intercept	-1148.36		89.73	-12.80		<0.001			
Smoother terms									
Depth		7.00			35.20	<0.001			
Slope		6.28			28.85	<0.001			
Dist.sm		6.99			59.01	<0.001			
Lat		6.89			30.47	<0.001			
SST		6.21			48.23	<0.001			
MSLA		6.62			31.32	<0.001			
Best final model (n= 2666; 19 presences): BD~s(Depth)+s(Slope)+s(Dist.sm)+s(Lat)+s(SST)+s(MSLA)							51.10	0.27	0.06
Sperm whales									
Intercept	-343.72		39.76	-8.65		<0.001			
Smoother terms									
Depth		6.96			38.63	<0.001			
Slope		6.96			28.05	<0.001			
Dist.sm		6.99			34.21	<0.001			
Lat		7.00			24.64	<0.001			
CHL		2.27			29.57				
SST		7.00			39.66	<0.001			
MSLA		5.97			36.89	<0.001			
Best final model (n= 2661; 14 presences): SW~s(Depth)+s(Slope)+s(Dist.sm)+s(Lat)+s(CHL)+s(SST)+s(MSLA)							66.70	0.43	0.03
Baleen whales									
Intercept	-3378.30		420.31	-8.04		<0.001			
CHL	6.80		1.38	4.93		<0.001			
Smoother terms									
Depth		6.54			14.36	<0.001			
Slope		6.95			35.58	<0.001			
Dist.sm		4.23			11.91	<0.001			
Lat		6.94			19.10	<0.001			
SST		6.81			17.65	<0.001			
MSLA		6.12			8.73	<0.001			
Best final model (n= 2662; 15 presences): BL~ CHL+s(Depth)+s(Slope)+s(Dist.sm)+s(Lat)+s(SST)+s(MSLA)							38.80	0.18	0.06
Beaked whales									
Intercept	-7445.90		1556.27	-4.78		<0.001			
CHL	-12.52		1.57	-8.00		<0.001			
Smoother terms									
Depth		6.44			10.41	<0.001			
Slope		5.08			18.12	<0.001			
Dist.sm		6.49			15.71	<0.001			
Lat		7.00			24.46	<0.001			

SST		6.96			31.47	<0.001			
MSLA		7.00			30.18	<0.001			
Best final model (n= 2669; 22 presences): BK~ CHL+s(Depth)+s(Slope)+s(Dist.sm)+s(Lat)+s(SST)+s(MSLA)							32.10	0.12	0.14

edf – effective degrees of freedom; se – standard error; n – total number of points (used / available) considered in the model fitting.

Common dolphins

Common dolphins were present along the routes and showed very wide ranges for several variables. However, there were clear habitat preferences: compared with the available habitat; this species preferred less deeper areas that were slightly closer to the coast (median: 3,143 m and 342.21 km, respectively). Among the species/groups analysed, they were the most prevalent in northern latitudes, with the higher median for latitude (4,308,201 m). Occurrences were distributed only in positive anomalies, from 3.05 cm to 14.96 cm (Figure 3).

The best GAM model obtained showed an explained deviance of 59.90% (Table 4). Common dolphins were positively influenced by depth, slope and dist.sm lower than 984 m, 6.96º and 84 km, respectively, and finally, by positive anomalies higher than 1.28 cm (Figure 4).

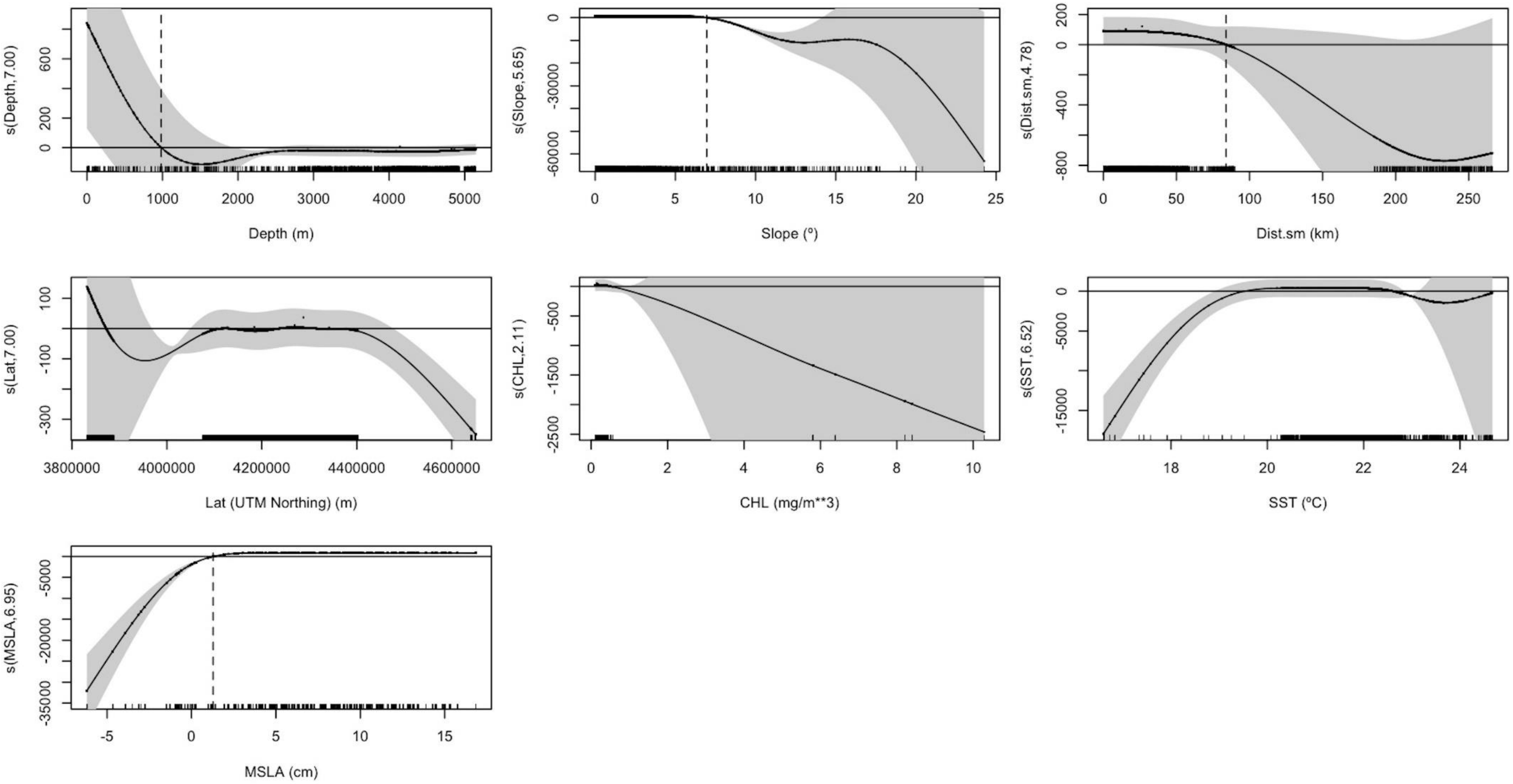


Figure 4. GAM-predicted non-linear splines of the response variable used / available habitat for common dolphins as a function of the explanatory variables. The degrees of freedom for non-linear fits are in parenthesis on the y-axis. Tick marks above the x-axis indicate the distribution of observations (with and without sightings). The shaded areas represent the 95% confidence intervals of the spline functions and dots on the graph area represent the residuals.

Bottlenose dolphins

The bottlenose habitat range overlapped with that of common dolphins in several variables. However, when analysing medians and the available habitat, there were distinct differences in habitat preferences: bottlenose dolphins were sighted in deeper waters (median: 4,804 m) with a wider habitat range concerning slope and dist.coast (from 0.03º to 13.39º and 3.74 km to 263.94 km, respectively). Geographically, this species occurred more often in southern latitudes (median: 4,224,146 m). Moreover, bottlenose dolphins were more common in warmer waters (median: 22.05ºC) and were distributed in areas with negative to positive sea-level anomalies (from -2.99 cm to 11.93 cm) (Figure 3).

The best GAM model obtained explained 51.10% of the deviance (Table 4). This species preferred areas with depths greater than 3,010 m. There were two peaks of preference for slope: from 0.90º to 7.63º and 10.00º to 15.21º; and the same was observed for dist.sm (lower than 74.50 km and higher than 234.10 km). Geographically, a positive influence was verified in southern latitudes (lower than 4,073,100 m). As for dynamic variables, there was a preference for colder waters in the area (lower than 20.31ºC) and for positive and negative anomalies from -5.00 cm to 8.93 cm (Figure 5). The CHL variable was removed from the model, leading to a better performance (Table 4).

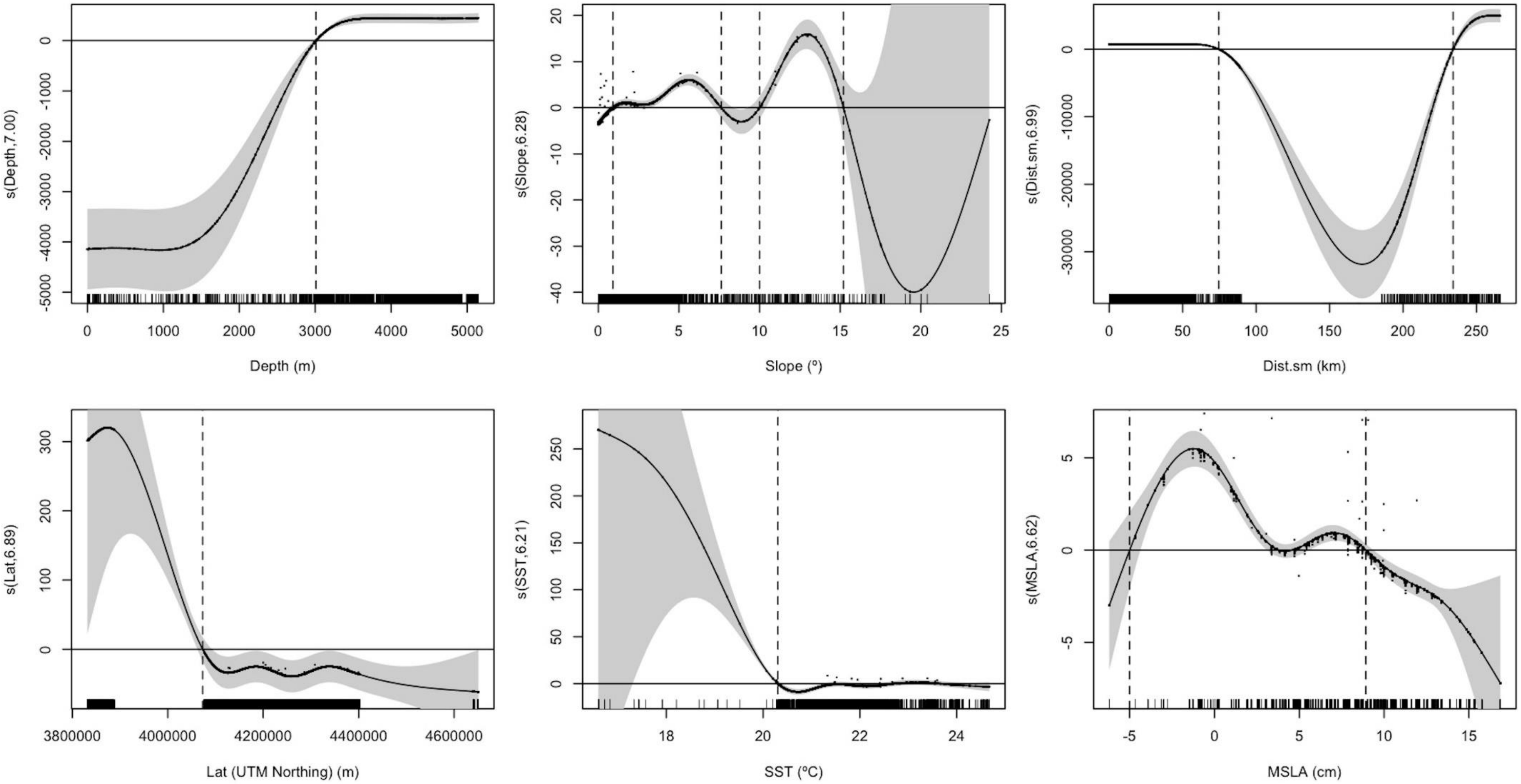


Figure 5. GAM-predicted non-linear splines of the response variable used / available habitat for bottlenose dolphins as a function of the explanatory variables. The degrees of freedom for non-linear fits are in parenthesis on the y-axis. Tick marks above the x-axis indicate the distribution of observations (with and without sightings). The shaded areas delimit the 95% confidence intervals of the spline functions and dots on the graph area represent the residuals.

Sperm whales

When compared with the rest of the species/groups, sperm whale prevalence was higher in shallow waters (median: 2,186 m), however, it was distributed along a wide range of depths (104 m to 4,870 m) and slopes (0.11º to 11.6º). The group of sperm whales had the lowest median for dist.coast (13.32 km) and the highest for dist.sm (239.59 km). The high prevalence near Madeira Island resulted in a clear preference for southern latitudes (median: 3,847,141 m). Sperm whales were distributed in oligotrophic waters in temperatures from 20.41ºC to 24.64ºC (median: 22.05ºC) and in sea-level anomalies from -1.48 cm to 9.81 cm (median: 8.41 cm) (Figure 3).

The best GAM model obtained explained the highest deviance of all the fitted models (66.70%) (Table 4). The GAM plots indicated a positive influence of depths above 4,019 m and slopes below 12.76º. Sperm whales preferred areas further from the seamounts (more than 102 km) and in latitudes between 4,022,700 m and 4,355,000 m. Concerning dynamic variables, these positively influenced the distribution at CHL, SST and MSLA lower than 0.90 mg/m^3, 20.07ºC and 11.07 cm, respectively (Figure 6).

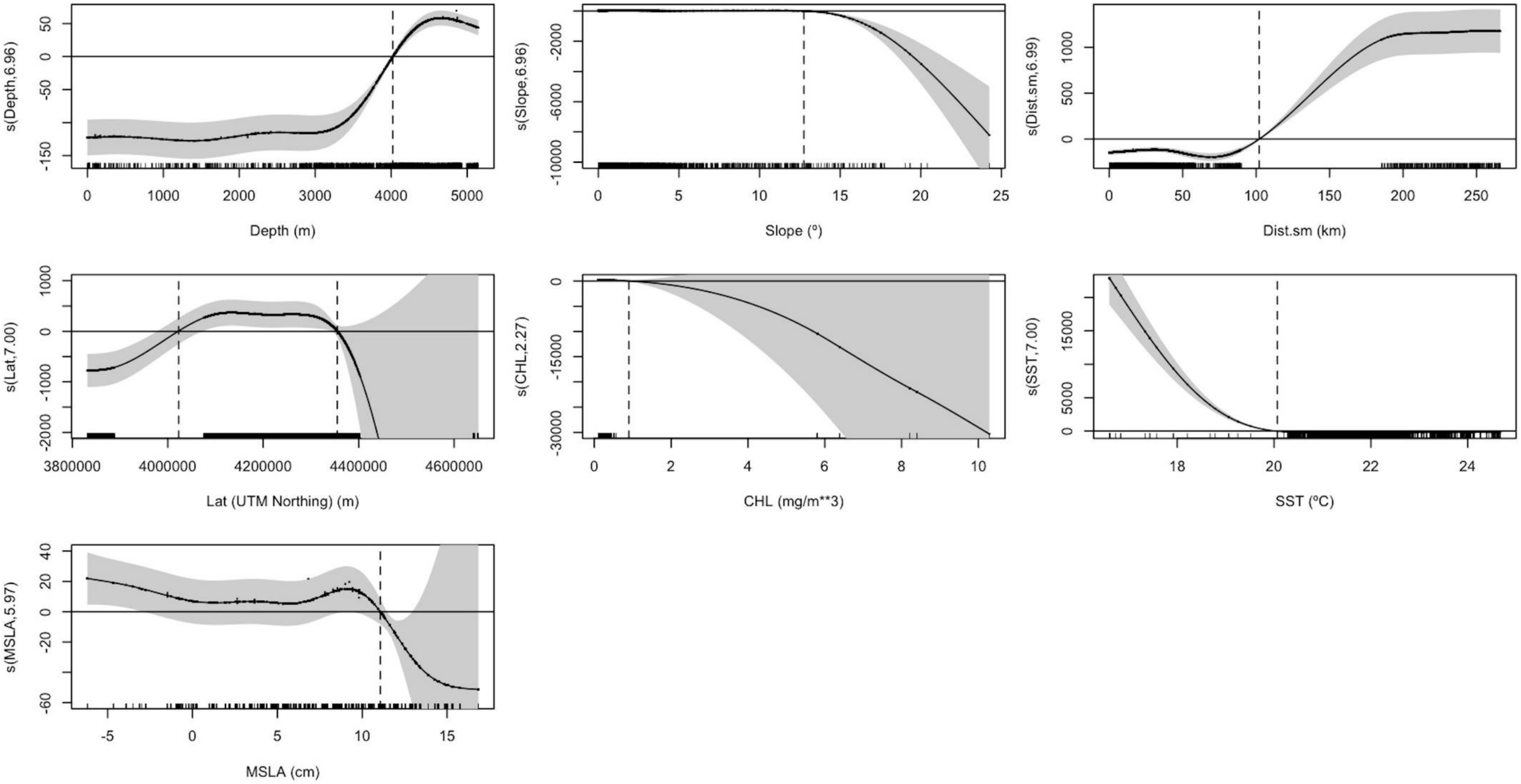


Figure 6. GAM-predicted non-linear splines of the response variable used / available habitat for sperm whales as a function of the explanatory variables. The degrees of freedom for non-linear fits are in parenthesis on the y-axis. Tick marks above the x-axis indicate the distribution of observations (with and without sightings). The shaded areas delimit the 95% confidence intervals of the spline functions and dots on the graph area represent the residuals.

Baleen whales

Baleen whales had medians similar to those for the available habitat, but contained smaller ranges. This group was mostly prevalent in deeper waters (median: 4,382 m), low slopes (median: 1.39º), further from the coast (median: 375.50 km) and close to the seamounts (median: 11.41 km). Baleen whales had the highest median for the dynamic variables SST and MSLA (22.22ºC and 100.07 cm, respectively) (Figure 3).

The best GAM model obtained explained 38.80% of the deviance (Table 4). Positive influences were verified at depths higher than 1,275 m, slopes and dist.sm lower than 11.28º and 107 km, respectively. Spatially, there was a preference towards northern latitudes (higher than 4,024,000 m). Warmer temperatures (higher than 19.87ºC) positively influenced distribution, and regarding MSLA, there were two intervals of positive influence: from -3.79 cm to 3.34 cm; and higher than 13.32 cm (Figure 7). Finally, a linear positive relationship with CHL was verified (Figure 8).

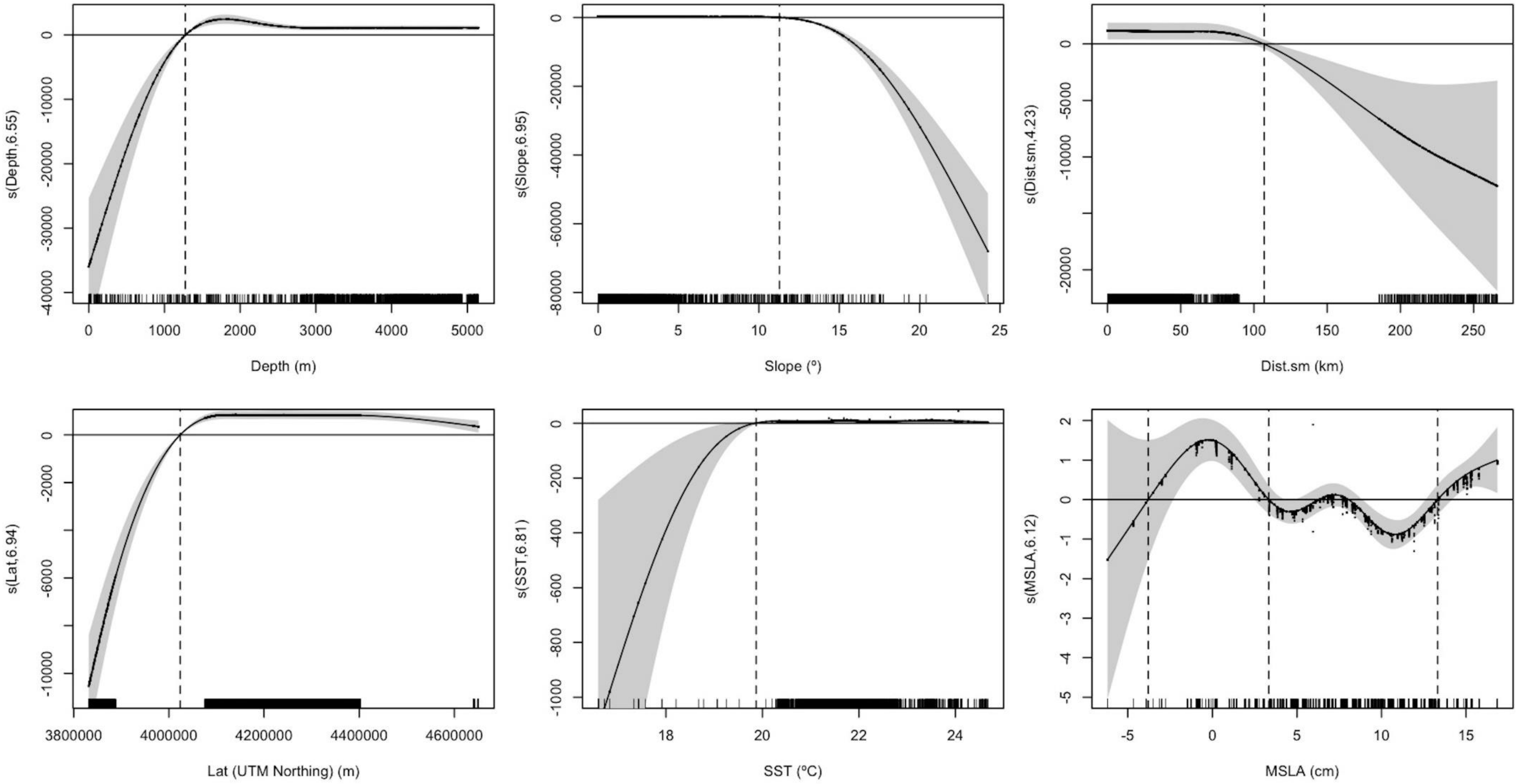


Figure 7. GAM-predicted non-linear splines of the response variable used / available habitat for baleen whales as a function of the explanatory variables. The degrees of freedom for non-linear fits are in parenthesis on the y-axis. Tick marks above the x-axis indicate the distribution of observations (with and without sightings). The shaded areas delimit the 95% confidence intervals of the spline functions and dots on the graph area represent the residuals.

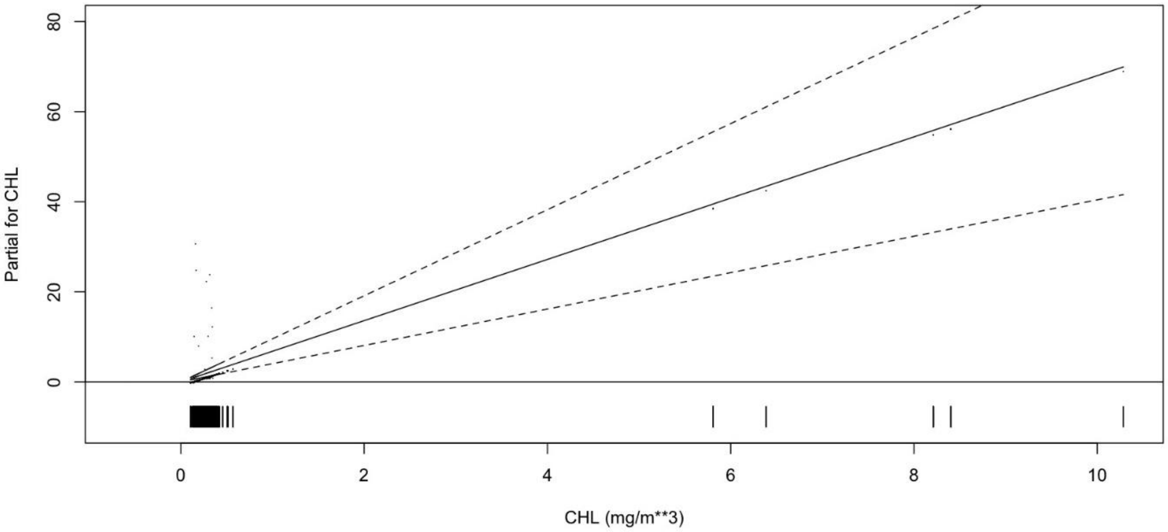


Figure 8. GAM-predicted linear spline of the response variable used / available habitat for baleen whales as a function of the CHL predictor. Dashed lines delimit the 95% confidence interval of the functions and dots on the graph area represent the residuals.

Beaked whales

The habitat range of beaked and baleen whales overlapped in several variables. However, the former had higher medians for depth (4,403 m), slope (2.54º) and dist.coast (398.60 km); and lower medians for SST and MSLA (21.32ºC and 7.81 cm, respectively) (Figure 3).

The best GAM model obtained explained 32.10% of the deviance (Table 4). This group preferred deeper areas (depth greater than 1,725 m) and slopes up to 10.06º, and was never found further than 155.50 km from the seamounts. Geographically, there was a clear preference towards southern areas (positive influence of latitude lower than 4,012,500 m). The dynamic variables positively influenced the distribution in temperatures up to 19.78ºC and in negative and positive anomalies (lower than -1.85 cm and higher than 6.25 cm, respectively) (Figure 9). A linear negative correlation with CHL was detected (Figure 10).

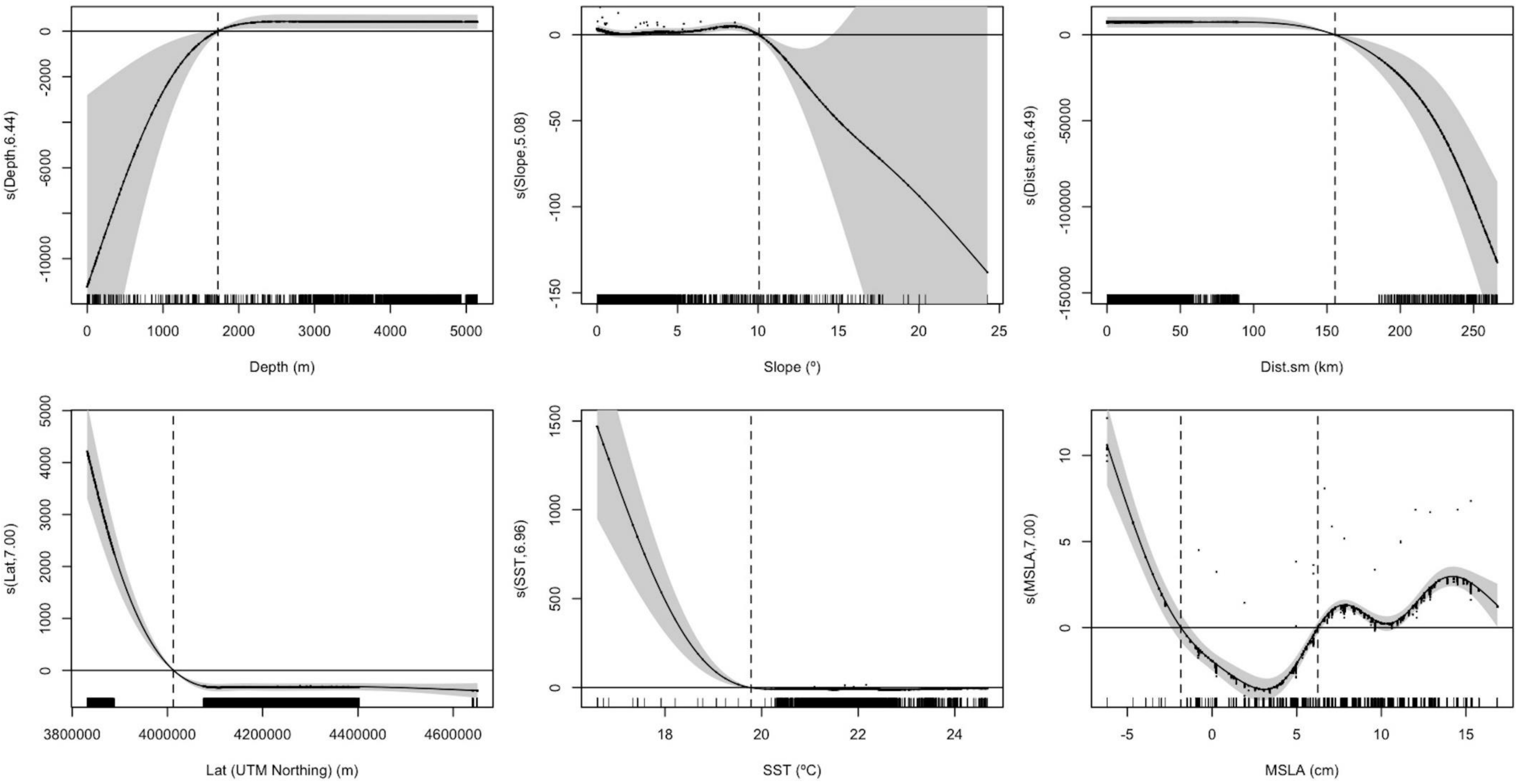


Figure 9. GAM-predicted non-linear splines of the response variable used / available habitat for beaked whales as a function of the explanatory variables. The degrees of freedom for non-linear fits are in parenthesis on the y-axis. Tick marks above the x-axis indicate the distribution of observations (with and without sightings). The shaded areas delimit the 95% confidence intervals of the spline functions and dots on the graph area represent the residuals.

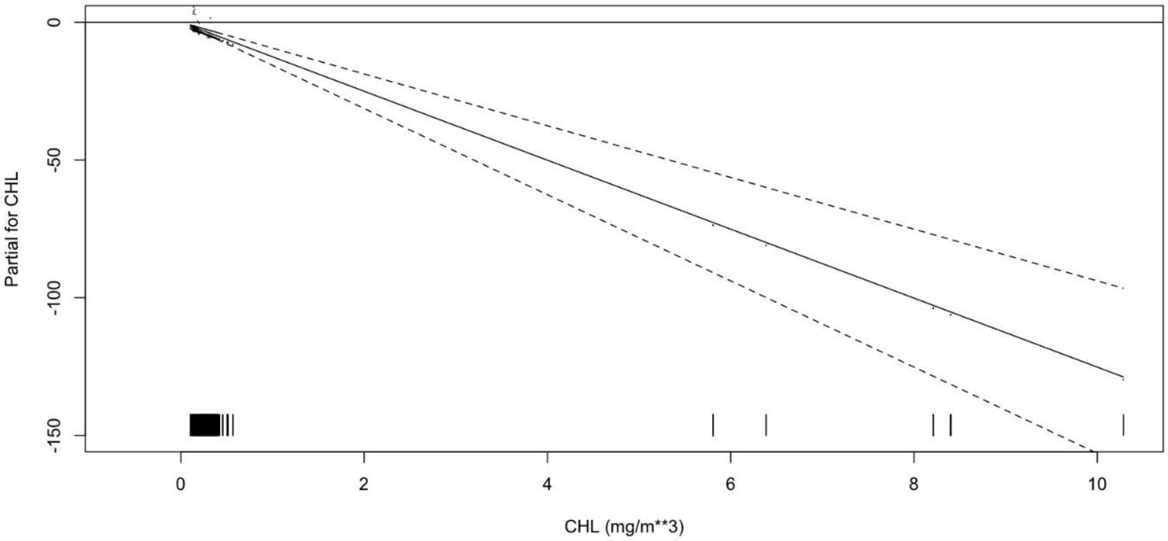


Figure 10. GAM-predicted linear spline of the response variable used / available habitat for beaked whales as a function of the CHL predictor. Dashed lines delimit the 95% confidence interval of the functions and dots on the graph area represent the residuals.

5.5 Discussion

To date, cetacean surveys in the PEEZ have been dedicated to coastal areas and little attention has been given to high seas. It is true that sampling offshore waters is logistically challenging (Abdulla *et al.*, 2009; Evans *et al.*, 2012; Kiszka *et al.*, 2007; Viddi *et al.*, 2010) and that anthropogenic threats to marine ecosystems are more intensive near the shore, and therefore, knowledge on coastal distribution has been considered more demanding (McIntyre, 1999). However, it is also true that these threats are increasingly affecting offshore areas upon which cetaceans highly depend (Abdulla *et al.*, 2009; Evans *et al.*, 2012; Hooker & Gerber, 2014; Hoyt, 2011). Therefore, there is an urgent need to obtain baseline data for high seas, where knowledge gaps still prevail. This is the case for the PEEZ, where in addition, offshore waters are particularly important, due to the topographic and oceanographic dynamics, which creates environmental conditions (seamounts, thermal fronts, upwellings and eddies) that are favourable to cetacean occurrence (Hoyt, 2011). The habitat diversity leads to high species richness and in the present study, a total 10,668 nmi sampled in the Mainland and Madeira Island PEEZ (mostly offshore waters) resulted in 218 sightings (ER = 2.04 sightings / 100 nmi) accounting for at least nine cetacean species from the three large groups (dolphin, baleen and toothed whales).

The importance of high seas to cetacean distribution was evident in the habitat preferences analysis for all the species / groups, with the exception of common dolphins. This was the most-sighted species with a very high prevalence in northern coastal areas on the Mainland, which is consistent with previous records for the Continental area (Brito & Sousa, 2011). Despite this, Moura *et al.* (2012) considered that on the Portuguese coast, the northern region had a low probability of common dolphin occurrence. Therefore, and taking into account the present study, a deeper analysis should be undertaken and dedicated surveys are recommended.

Bottlenose dolphins were the second most-sighted species, overlapping in their habitat range with common dolphins in several variables. Where both species have been reported in coastal areas, in the Continent and the Island (Brito & Sousa, 2011; www.museudabaleia.org), the conducted sampling design also demonstrated their occurrence in the high seas, suggesting the existence of coastal and oceanic populations. This is probably an effect of the seamounts along transects, as model results appear to indicate. The NE Atlantic seamounts have been characterised as highly productive (WWF, 2003) and these dolphins are known to be opportunistic feeders that take advantage of the local features that aggregate their preferred prey, the schooling fish (Cañadas *et al.*, 2002; Moura *et al.*, 2012). Both species were observed feeding in the same areas at the same time, thus, as they share the same prey, they probably also share feeding grounds in the area. However, these results indicate that although these species share common environmental habitats, they also differ in spatial preferences: common dolphins prefer northern latitudes (and probably seamounts in northern latitudes) compared to bottlenose dolphins.

The occurrence of sperm whales and suggested site fidelity in Madeira Island is known and has been previously described (www.museudabaleia.org), but the present study indicates that their habitat range extends towards the open ocean and reaches middle latitudes. In fact, even though most sightings were in waters near Madeira Island, this species demonstrated a wide habitat range concerning the analysed variables. Moreover, the occurrence of this species in the high seas apparently does not depend on seamounts, as these features did not positively influence their distribution.

The other two groups of beaked and baleen whales occupied deeper waters with narrower habitat ranges, which is consistent with knowledge on the groups, and in concrete, on the animals in the area (www.museudabaleia.org). Moreover, results show seamounts determine the distribution of these groups in high seas. These features were suggested to play an important role for the distribution of all cetaceans, but especially for beaked whales (Azzellino *et al.*, 2012; Cañadas *et al.*, 2002; Moulins *et al.*, 2007). This group mostly inhabits offshore areas and, as they are deep-divers with elusive behaviour, are particularly difficult to observe. However, from the conservation point of view, data on their occurrence and hotspots are essential, as they are very sensitive to anthropogenic threats (Kashner *et al.*, 2012; MacLeod *et al.*, 2008). In this area, beaked whales had a considerably high total ER, which might indicate that they are common in the area, contrary to what is suggested by the few records of their occurrence, which may result from their occurrence mainly in offshore waters where little effort has been undertaken (MacLeod *et al.*, 2008). Beaked whales distribution appears to be mostly shaped by seamounts in mid- to southern latitudes. Furthermore, some evidence exists for habitat segregation among beaked and sperm whales: the former occurs in deeper areas further from the coast and closer to seamounts and the latter prefers areas surrounding Madeira Island and is not influenced by the presence of seamounts. Even though these two species share the same type of prey (cephalopods), they are both size-limited predators, and only take a narrow range of prey relative to their body size. Since they are very different sizes, they do not necessarily consume the same sizes of prey (MacLeod *et al.*, 2006). Therefore, if different sizes of cephalopods occur at different depths, the two toothed whales will have different habitat preferences leading to a substantial reduction of their niches overlap. This habitat segregation has been suggested previously in the Mediterranean Sea (Azzellino *et al.*, 2008; Tepsich *et al.*, 2014).

Whereas it is frequent to use the knowledge on distribution based solely on static variables in designing MPAs, due to the static nature of such administrative boundaries, the dynamism of habitat suitability should be considered, especially in areas where dynamic processes highly influence the productivity and therefore the distribution of cetaceans (Hooker *et al.*, 2011). Moreover, when dealing with high seas, dynamic variables derived from remote-sensing data are more accurate and have less missing data than in coastal regions (Robinson, 2010). In the present study, most of the effort was in the high seas and given the dynamism of the area, such variables had to be considered in the modelling process and, as expected, definitely impacted cetacean distribution.

The CHL had a positive effect on baleen whale distribution, which was probably influenced by phytoplankton blooms in the area that favoured large zooplankton swarms (such as krill), the prey of mysticetes species (Viddi *et al.*, 2010). On the contrary, beaked whales distribution was negatively related with CHL, which can be a result of the spatial-temporal lag occurring between phytoplankton blooms and its influence on preys of this teutophagus group. Though it has been proved that CHL is an important predictor in cetacean habitat modelling, determining the appropriate spatial and temporal scale to use is challenging when dealing with high trophic level species, due to the existing lag of the bottom-up control (Frederiksen *et al.*, 2006; Grémillet *et al.*, 2008; O'Hern & Biggs, 2009). This may also justify the low deviance explained obtained in the models for this variable. The GAM models also predicted positive peaks for low SST values for the three groups of bottlenose dolphins, sperm and beaked whales. For bottlenose dolphins and beaked whales, this appears to be related to small and frequent offshore upwelling systems near the seamounts. Even though sperm whales show a clear preference for the Madeira Island region, where SST is generally high in the area, they possibly occur mainly in niches with small upwelling phenomena, and therefore, in more productive waters (Robinson, 2010). The impact of MSLA was particularly interesting in the distribution of beaked whales. There were two major peaks of positive influence of MSLA in this group, which might indicate a preference for areas that are particularly active in terms of mesoscale eddies (causing downwelling and upwelling phenomena) (Robinson, 2010). The peaks correspond to a high positive MSLA, which is possibly associated with anti-cyclonic eddies and small upwelling phenomena; and a high negative MSLA, which is possibly associated with cores of cyclonic eddies. All these areas are associated with high productivity (Robinson, 2010). This relationship was reported previously and is very important, as it can be a very good indicator of species movements, especially in such a sensitive group as beaked whales (Baird *et al.*, 2011).

Both the techniques that were used for habitat analysis were complementary in providing a general picture of habitat preferences and potential suitable habitats where dedicated surveys are recommended in the region. Whereas quantile analysis provided an idea of the habitat range and distribution of occurrences along the sampled habitat, the GAM models and subsequent *GAMvelope* approach allowed the understanding of the influence of the variables on cetacean distribution and potential trends of cetacean occurrence in response to these variables (Austin, 2007; Elith & Leathwick, 2009; Pearce and Boyce, 2006; MacLeod *et al.*, 2008). Moreover, they both highlighted and corroborated the importance of sampling different habitat profiles and including static and dynamic environmental and geographical variables that are representative of the habitat diversity in the area (Hooker *et al.*, 2011; Redfern *et al.*, 2006). While the results from this study must be carefully interpreted, they suggest that with more analysis, more data and dedicated surveys in potential suitable habitats, cetacean hotspots could be identified.

5.6 Conclusions

To our knowledge, these are the first published results on cetacean occurrence and habitat preferences for the offshore waters in the Mainland and Madeira Island PEEZ. Although the use of OPOs partially limited the temporal and spatial resolution of data collection, considering the scarce knowledge in the PEEZ, this approach proved to be a very efficient and cost-effective method to collect data. In fact, several studies have successfully used OPOs to collect scientific data, namely for cetacean distribution (Kiszka *et al.*, 2007; MacLeod *et al.*, 2008;

Moulins *et al.*, 2007; 2008; Moura *et al.*, 2012; Viddi *et al.*, 2010) and oceanographic data (Lüger *et al.*, 2006).

The presented preliminary results show the potential of the area as habitat for cetaceans, given the environmental space. Furthermore, this initial assessment on habitat preferences is an important contribution as it determines possible suitable habitats for cetacean occurrence in the high seas, where little or no published data exist and where dedicated surveys are required to inform decision-makers.

Though it has been proved that offshore MPAs are efficient (Alemany *et al.*, 2012), their design and implementation is challenging when dealing with highly mobile species such as cetaceans, and dynamic open ocean areas with several logistic constrains (Evans *et al.*, 2012; Kiszka *et al.*, 2007; Viddi *et al.,* 2010). To succeed, this conservation measures need to be supported by a solid knowledge on the entire year-round cetacean distribution and their areas of special importance such as feeding and breeding grounds (UNEP-MAP RAC/SPA, 2011). Also, given the wide range of cetacean habitat, a network of protected areas instead of single MPAs is suggested (UNEP-MAP RAC/SPA, 2011). Moreover, considering the oceanographic dynamics in the area, which highly influences cetacean distribution, efforts have to be made in order to predict hotspots given the habitat conditions (Hooker *et al.*, 2011). With increasing scientific advances in the forecasting and modelling of oceanographic conditions, the implementation of dynamic MPAs is achievable and recommended (Hooker *et al.*, 2011).

Ultimately, considering the potential of the area for habitat modelling, improving the dataset would provide robust and predictive models, which are essential to efficiently determine the distribution and to plan an effective MPA network in the future PEEZ. Therefore, a long-term monitoring network, with an increase in the sampling season and sampled routes in the PEEZ, including dedicated surveys in potential suitable habitats, is recommended. Such a monitoring program would provide relevant knowledge to support decision-makers in developing an effective Marine Spatial Management for Cetacean Conservation in the area and address several resolutions identified by ACCOBAMS, including: Resolution 3.22 – achieve a representative network of MPAs and respond to targets defined by Convention on Biological Diversity (CBD) and Resolution 4.12 – obtain comprehensive cetacean population estimates and distribution in the ACCOBAMS area (ACCOBAMS, 2013).

5.7 Acknowledgements

This study was developed during a Master program from the Faculty of Sciences of the University of Porto, Portugal and was supported by the Interdisciplinary Centre of Marine and

Environmental Research (CIIMAR – Porto, Portugal) and by the CIMA Research Foundation (Savona, Italy). A special acknowledgement goes to Filipe Alves and Xavier Couvelard from CIIMAR Madeira for the relevant and important support and sharing of knowledge throughout this study. We also thank the contribution and dedication of all the volunteers during the monitoring campaigns: Olga Azevedo, Zara Valquíria, Andreia Pereira, Cláudia Ferreira, Pedro Fernandes, Sofia Silva, Joana Romero and Ayça Eleman. We are also extremely grateful to TRANSINSULAR, the cargo ship company that provided all the logistic support, and to the ship crews for their hospitality. Finally, we want to thank the two anonymous reviewers, whose comments and suggestions on the early drafts of the manuscript, helped to improve it.

5.8 References

Abdulla A., Gomei M., Hyrenbach D., Notarbartolo-di-Sciara G. and Agardy T. (2009(. Challenges facing a network of representative marine protected areas in the Mediterranean: prioritizing the protection of underrepresented habitats. *ICES Journal of Marine Science*, 66: 22-28.

Alemany D., Iribarne O.O. and Acha E.M. (2013). Effects of a large-scale and offshore marine protected area on the demersal fish assemblage in the Southwest Atlantic. *ICES Journal of Marine Science*, 70: 123-134.

ACCOBAMS (2013). Resolutions. Available at: http://www.accobams.org/ (accessed in October, 2013).

Alves F., Quérouild S., Dinis A., Nicolau C., Ribeiro C., Freitas L., *et al*. (2013). Population structure of short-finned pilot whales in the oceanic archipelago of Madeira based on photo-identification and genetic analyses: implications for conservation. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 23(5): 758-776.

Alves F. and Dinis A. (2014). Survival and abundance of short-finned pilot whales in the archipelago of Madeira, NE Atlantic. *Marine Mammal Science*, 31(1): 106-121.

Armstrong C.W., Foley N.C., Tinch R. and van den Hove S. (2012). Services from the deep: Steps towards valuation of deep sea goods and services. *Ecosystem Services*, 2: 2–13.

Augusto J. (2007). Análise da estrutura social, da composição dos grupos e associações nos Golfinhos-Roazes (*Tursiops truncatus*) residents na região do Sado. Master thesis (MSc) in Marine Ecology. Faculdade de Ciências – Universidade de Lisboa, Lisboa. 60 pp.

Austin M. (2007). Species distribution models and ecological theory: A critical assessment and some possible new approaches. *Ecological Modelling*, 200: 1-19.

AVISO (2013). Sea-surface height products. Available at: http://www.aviso.oceanobs.com/en/data/products/sea-surface-height-products/global/madt.html#c5132 (accessed in November 2013).

Azzellino A., Gaspari S., Airoldi S. and Nani B. (2008). Habitat use and preferences of cetaceans along the continental slope and the adjacent pelagic waters in the western Ligurian Sea. *Deep-Sea Research I*, 55: 296-323.

Azzellino A., Panigada S., Lanfredi C., Zanardelli M., Airoldi S. and Notarbartolo di Sciara G. (2012). Predictive habitat models for managing marine areas: Spatial and temporal distribution of marine mammals within the Pelagos Sanctuary (Northwestern Mediterranean sea). *Ocean and Coastal Management*, 67: 63–74.

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

Baird R.W., Schorr G.S., Webster D.L., Mahaffy S.D., McSweeney D.J., Hanson M.B., and Andrews R.D. 2011. Open-Ocean Movements of a Satellite-Tagged Blainville's Beaked Whale (*Mesoplodon densirostris*): Evidence for an Offshore Population in Hawaii? *Aquatic Mammals*, 37(4): 506-511.

Barbet-Massin M., Jiguet F., Albert C.H. and Thuiller W. (2012). Selecting pseudo-absences for species distribution models: how, where and how many? *Methods in Ecology and Evolution*, 3(2): 327-338.

Brito C. and Sousa A. (2011). The Environmental History of cetaceans in Portugal: Ten Centuries of Whale and Dolphin Records. *PLoS ONE*, 6(9): e23951.

Brotons L., Thuiller W., Araújo M.B. and Hirze A.H. (2004). Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. *Ecography*, 27: 437-448.

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

Certain G., Ridoux V., van Canneyt O. and Bretagnolle V. (2008). Delphinid spatial distribution and abundance estimates over the shelf of the Bay of Biscay. *ICES Journal of Marine Science*, 65: 656-666.

Correia A.M. (2013). Cetacean monitoring in Northeastern Atlantic Ocean: Occurrence and distribution of cetacean species in the Canary Basin. Master thesis (MSc) in Ecology, Environment and Territory. Sciences Faculty of Porto University, Porto, Portugal. 93 pp.

Cox T.M., Ragen T.J., Read A.J., Vos E., Baird R.W., Balcomb K, *et al.* (2006). Understanding the impacts of anthropogenic sound on beaked whales. *Journal of Cetacean Research and Management*, 7(3):177-187.

Dalla Rosa L., Ford J.K.B. and Trites A.W. (2012). Distribution and relative abundance of humpback whales in relation to environmental variables in coastal British Columbia and adjacent waters. *Continental Shelf Research*, 36: 89-104.

Davis R.W., Ortega-Ortiz J.G., Ribic C.A., Evans W.E., Biggs D.C., Ressler P.H., *et al.* (2002). Cetacean habitat in the northern oceanic Gulf of Mexico. *Deep Sea Research Part I: Oceanographic Research Papers*, 49: 121-142.

EarthRef (2013). Seamount Catalog. Available at: http://earthref.org/SC/. (accessed in April 2013)

Elith J. and Leathwick J.R. (2009). Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *The Annual Review of Ecology, Evolution, and Systematics*, 40: 677-697.

ESRI. 2011. ArcGIS Desktop: Release 10. Redlands, CA: Environmental Systems Research Institute.

Esteban R., Verborgh P., Gauffier P., Giménez J., Afán I., Cañadas A., *et al.* (2013). Identifying key habitat and seasonal patterns of a critically endangered population of killer whales. *Journal of the Marine Biological Association of the United Kingdom*, 1-9.

Estrutura de Missão para a Extensão da Plataforma Continental – EMEPC (2013). Extensão Plataforma Continental. Available at: http://www.emepc.pt/. (accessed in October 2013)

European Commission (2013). *Commission Staff working document - impact assessment. Accompanying the document - Proposal for a directive of the European Parliament and of the Council establishing a framework for maritime spatial planning and integrated coastal management*. Brussels, Belgium.

Evans P.G.H., Pierce G.J. and Wright A.J. (2012). Marine mammal studies to address future challenges in conservation management. *Journal of the Marine Biological Association of the United Kingdom*, 92: 1639-1644.

Frederiksen M., Edwards M., Richardson A.J., Halliday N.C. and 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(6): 1259-1268.

GEBCO (2012). Gridded bathymetry data. Available at: http://www.gebco.net/data_and_products/gridded_bathymetry_data/. (assessed February 2012).

Google (2013). Google Play. Available at: https://play.google.com/store/apps/details?id=com.google.android.maps.mytracks. (assessed April 2013).

Grémillet D., Lewis S., Drapeau L., Lingen C.D.V.D, Huggett J.A., Coetzee J.C, *et al.* (2008). Spatial match–mismatch in the Benguela upwelling zone: should we expect chlorophyll and sea-surface temperature to predict marine predator distributions? Journal of Applied Ecology, 45(2): 610-621.

Hooker S.K., Cañadas A., Hyrenbach D., Corrigan C., Polovina J.J. and Reeves R.R. (2011). Making protected area networks effective for marine top predators. *Endangered Species Research*, 13: 203-218

Hooker S.K. and Gerber L.R. (2014). Marine Reserves as a Tool for Ecosystem-Based Management: The Potential Importance of Megafauna. *Bioscience*, 54(1): 27-39.

Hoyt E. (2011). *Marine Protected Areas for Whales, Dolphins and Porpoises*. Second Edition, Earthscan. New York, USA.

Kaschner K., Quick N.J., Jewell R., Williams R. and Harris C.M. (2012). Global Coverage of Cetacean Line-Transect Surveys: Status Quo, Data Gaps and Future Challenges. *PLoS ONE*, 7(9): e44075.

Kiszka J., Macleod K., Van Canneyt O., Walker D. and Ridoux V. (2007). Distribution, encounter rates, and habitat characteristics of toothed cetaceans in the Bay of Biscay and adjacent waters from platform-of-opportunity data. *ICES Journal of Marine Science*, 64: 1033-1043.

Lüger H., Wanninkhof R., Wallace D.W.R. and Körtzinger A. (2006). CO2 fluxes in the subtropical and subarctic North Atlantic based on measurements from a volunteer observing ship. Journal of Geophysical Research, 111: C06024.

MacLeod C.D., Santos MB., Lopez A. and Pierce G.J. (2006). Relative prey size consumption in toothed whales: implications for prey selection and level of specialisation. *Marine Ecology Progress Series*, 326: 295-307.

MacLeod C.D., Mandleberg L., Schweder C., Bannon S.M., and Pierce G.J. (2008). A comparison of approaches for modelling the occurrence of marine animals. *Hydrobiologia*, 612: 21-32.

Marubini, F., Gimona, A., Evans, P. G. H., Wright, P. J., and Pierce, G. J. 2009. Habitat preferences and interannual variability in occurrence of the harbour porpoise *Phocoena phocoena* off northwest Scotland. Marine Ecology Progress Series, 381: 297–310.

Mason E. (2009). High-resolution modelling of the Canary Basin oceanic circulation. Doctorate thesis (PhD) in Oceanography. Universidad de Las Palmas de Gran Canaria, Las Palmas de Gran Canaria. 245 pp.

MATLAB R2012a version 7.14.0 (2012). Computer software, The MathWorks Inc., Natick, Massachusetts.

McIntyre A.D. (1999). Conservation in the sea–looking ahead. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 9: 633-637.

Morato T., Varkey D.A., Damaso C., Machete M., Santos M., Priet, R., *et al.* (2008). Evidence of a seamount effect on aggregating visitors. *Marine Ecology Progress Series*, 357: 23-32.

Morato T., Hoyle S.D., Allain V. and Nicola S.J. (2010). Seamounts are hotspots of pelagic biodiversity in the open ocean. *Proceedings of the National Academy of Sciences*, 107(21): 9707-9711.

Moulins A., Rosso M., Nani B. and Würtz M. (2007). Aspects of the distribution of Cuvier's beaked whale (*Ziphius cavirostris*) in relation to topographic features in the Pelagos Sanctuary (north-western Mediterranean Sea). *Journal of the Marine Biological Association of the United Kingdom*, 87: 177-186.

Moulins A., Rosso M., Ballardini M. and Würtz M. (2008). Partitioning of the Pelagos Sanctuary (north-western Mediterranean Sea) into hotspots and coldspots of cetacean distributions. *Journal of the Marine Biological Association of the United Kingdom*, 88(6): 1273-1281.

Moura A.E., Sillero N. and Rodrigues A. (2012). Common dolphin (*Delphinus delphis*) habitat preferences using data from two platforms of opportunity. *Acta Oecology*, 38: 24-32.

NASA (2013a). Giovanni – Ocean Color Radiometry Online Visualization and Analysis. Available at: http://gdata1.sci.gsfc.nasa.gov/daac-bin/G3/gui.cgi?instance_id=ocean_month. (assessed in January 2014).

NASA (2013b). Ocean Color Data – MODISA. Available at: http://oceandata.sci.gsfc.nasa.gov/. (assessed in November 2013).

O'Hern J.E. and Biggs D.C. (2009). Sperm Whale (*Physeter macrocephalus*) Habitat in the Gulf of Mexico: Satellite Observed Ocean Color and Altimetry Applied to Small-Scale Variability in Distribution. *Aquatic Mammals*, 35: 358-366.

Pearce J.L. and Boyce M.S. (2006). Modelling distribution and abundance with presence- only data. *Journal of Applied Ecology*, 43: 405-412.

Pirotta E., Matthiopoulos J., MacKenzie M., Scott-Hayward L. and Rendell L. (2011). Modelling sperm whale habitat preference: a novel approach combining transect and follow data. *Marine Ecology Progress Series*, 436: 257-272.

R Development Core Team (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna Austria. ISNB: 3-900051-07-0, URL: http://www.R-project.org/.

Redfern J.V., Ferguson M.C., Becker E.A., Hyrenbach K.D., Good C., Barlow J., *et al.* (2006). Techniques for cetacean-habitat modeling. *Marine Ecology Progress Series*, 310: 271-295.

Robinson I.S. (2010). *Discovering the Ocean from Space – The unique applications of satellite oceanography*. Springer. Bath, UK.

Schlacher T.A., Rowden A.A., Dower, J.D. and Consalvey M. (2010). Seamount science scales undersea mountains: new research and outlook. *Marine Ecology*, 31: 1-13.

Sergio F., Newton I., Marchesi L. and Pedrini P. (2006). Ecologically justified charisma: preservation of top predators delivers biodiversity conservation. *Journal of Applied Ecology*, 43: 1049-1055.

Sergio F., Caro T., Brown D., Clucas B., Hunter J., Ketchum J., *et al.* (2008). Top Predators as Conservation Tools: Ecological Rationale, Assumptions, and Efficacy. *Annual Review of Ecology, Evolution, and Systematics*, 39: 1-19.

Silva M.A., Prieto R., Magalhães S., Cabecinhas R., Cruz A., Gonçalves J.M. and Santos R.S. (2003). Occurrence and distribution of cetaceans in the waters around the Azores (Portugal), Summer and Autumn 1999–2000. *Aquatic Mammals*, 29.1: 77-83.

Spyrakos E., Santos-Diniz T.C., Martinez-Iglesias G., Torres-Palenzuela J.M. and Pierce G.J. (2011). Spatiotemporal patterns of marine mammal distribution in coastal waters of Galicia, NW Spain. *Hydrobiologia,* 670: 87-109.

Stockin K.A., Pierce G.J., Binedell V., Wiseman N. and Orams M.B. (2008). Factors affecting the occurrence and demographics of common dolphins (*Delphinus* sp.) in the Hauraki Gulf, New Zealand. *Aquatic Mammals*, 34: 200-211.

Tepsich P., Rosso M., Halpin P.N. and Moulins A. (2014). Habitat preferences of two deep-diving cetacean species in the northern Ligurian Sea. *Marine Ecology Progress Series*, 508: 247-260.

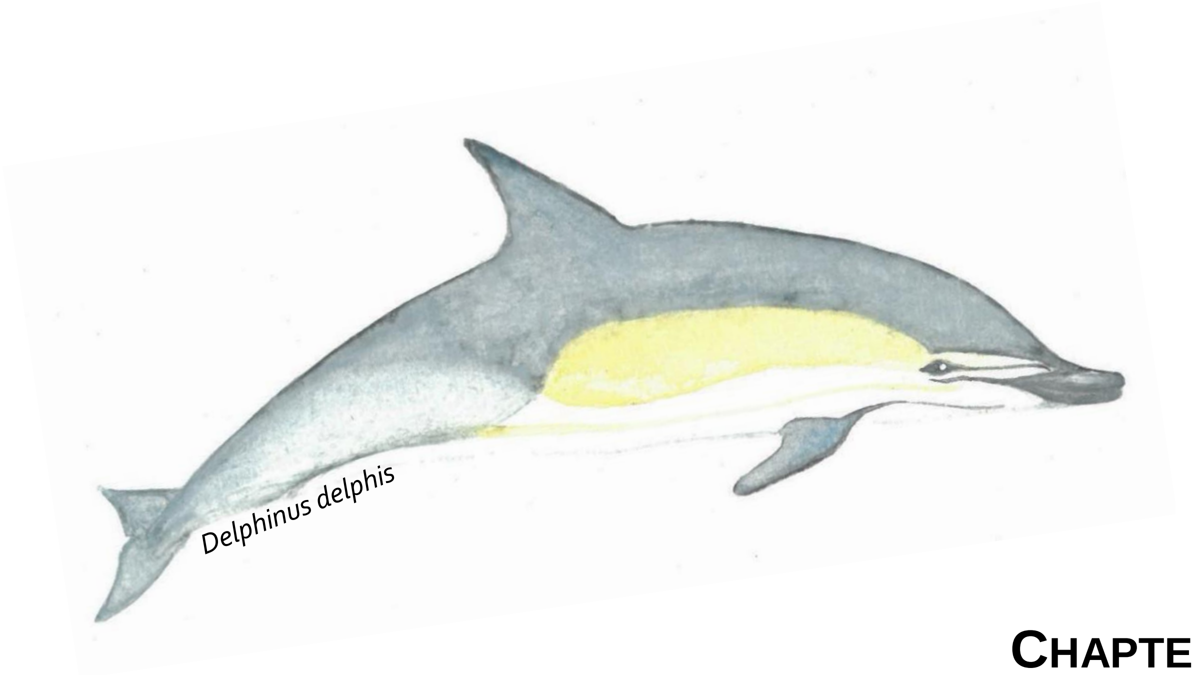
Torres L.G., Read A.J. and Halpin P. (2008). Fine-scale habitat modelling of a top marine predator: do prey data improve predictive capacity? *Ecological Applications*, 18: 1702-1717.

UNEP-MAP RAC/SPA (2011). *Guidelines for the Establishment and Management of Marine Protected Areas for Cetaceans*. By Giuseppe Notarbartolo di Sciara. Ed. ACCOBAMS-RAC/SPA, Tunis, 36pp.

Viddi F.A., Hucke-Gaete R., Torres-Florez J.P. and Ribeiro S. (2010). Spatial and seasonal variability in cetacean distribution in the fjords of northern Patagonia, Chile. *ICES Journal of Marine Science*, 67: 959-970.

Visser F., Hartman K.L., Pierce G.J., Valavanis V.D. and Huisman J. (2011). Timing of migratory baleen whales at the Azores in relation to the North Atlantic spring bloom. *Marine Ecology Progress Series*, 440: 267-279.

WWF (2003). *Seamounts of the North-East Atlantic. Oceanic Seamounts – An Integrated Study*. Frankfurt am Main, Germany.



CHAPTER VI.

Paper published in the Journal of the Marine Biological Association of the United Kingdom, 2019. doi: 10.1017/S0025315419000249

6 Distribution and habitat modelling of common dolphins (*Delphinus delphis*) in the eastern North Atlantic

Ana M. Correia[1,2], Ágatha Gil[1,2], Raul Fonseca Valente[1,2], Massimiliano Rosso[1,3], Graham J. Pierce[4,5,6], Isabel Sousa-Pinto[1,2]

[1]Interdisciplinary Centre of Marine and Environmental Research (CIIMAR). 4450-208 Matosinhos, Portugal.
[2]Department of Biology, Faculty of Sciences, University of Porto (FCUP). 4169-007 Porto, Portugal.
[3]CIMA Research Foundation. 17100 Savona, Italy.
[4]Instituto de Investigacións Mariñas (CSIC). 36208 Vigo, Pontevedra, Spain.
[5]Oceanlab, University of Aberdeen. AB41 6AA, UK.
[6]CESAM and Department of Biology, University of Aveiro. 3810-193 Aveiro, Portugal.

6.1 Abstract

The eastern North Atlantic (ENA) has many highly productive areas where several species of cetaceans have been recorded, with the common dolphin (*Delphinus delphis*) being one of the most frequently sighted species. However, its spatial and temporal distribution in high seas is poorly known. The study presents the results from 5 years of cetacean monitoring in the ENA (2012–2016) aboard cargo ships that follow the routes from Continental Portugal to the Macaronesian archipelagos and north-west Africa. Common dolphin was the most frequently sighted cetacean with 192 occurrences registered on effort and an overall encounter rate of 0.36 sightings / 100 nmi. The species was distributed in coastal and offshore waters, but absent from the Canaries and Cape Verde islands. Statistical ‘habitat’ models were developed to describe and explain the occurrence of sightings of the species: variables affecting detection of dolphins had a small impact and there were clear spatiotemporal distribution patterns, influenced to some degree by environmental variables. Predicted probability of occurrence was highest in coastal waters of continental Portugal and around the Azores. The models, combined with maps of distribution, were useful to identify important areas for the species, which could be the focus of future conservation efforts. Common dolphin presence was related to depth, distance to coast and seamounts, seabed slope, chlorophyll concentration, sea-surface temperature and sea level anomalies; the possible ecological significance of these relationships is explored.

Keywords: Cetaceans; ecological modelling; high seas; Macaronesia; spatial distribution; temporal distribution

6.2 Introduction

The eastern North Atlantic Ocean (ENA) includes the four archipelagos of the biogeographic region of Macaronesia: Azores, Madeira, Canaries and Cape Verde. The region has a complex topography including seamounts, hills, banks, abyssal platforms, canyons, and a rugged coastline along European and African continents. Moreover, it is characterized by dynamic oceanographic processes: strong coastal upwelling phenomena, formation of numerous eddies and fronts, and the presence of several Atlantic oceanic currents (Caldeira *et al.*, 2002; Mason, 2009; Sala *et al.*, 2013). This complexity and diversity of habitat conditions plays a major role in the distribution of primary production, and therefore, in the distribution of biomass across the trophic levels of the marine food chain. Cetacean distribution in space and time is generally considered to be shaped by environmental factors that condition prey availability at different spatial and temporal scales (for a review, see Redfern *et al.*, 2006). Nonetheless, when looking at distribution based on observational data, it is necessary to account for factors affecting detectability in order to obtain reliable information (e.g. Pierce *et al.*, 2010). These factors include the conditions of the platform of observation, survey design, state of the weather during the survey, distance to the sighted animal(s), species detected, size of the group, and, ultimately, the ability of the observer to detect and identify the species. In the ENA, at least 36 cetacean species have been recorded, both resident and migrating, in coastal and oceanic areas (e.g. Alves *et al.*, 2013, 2018, 2019; Berrow *et al.*, 2015; Carrillo *et al.*, 2010; Correia *et al.*, 2015; Dinis *et al.*, 2016, 2017; Djiba *et al.*, 2015; Goetz *et al.*, 2015; Hammond *et al.*, 2013; Hazevoet & Wenzel, 2000; Hazevoet *et al.*, 2010; Jungblut *et al.*, 2017; Silva *et al.*, 2014; Tobeña *et al.*, 2016; Weir, 2010; Weir & Pierce, 2013). All cetaceans in European Union (EU) waters receive protection under the Habitats Directive (Council Directive 92/43/EEC) and the Marine Strategy Framework Directive (MSFD, Directive 2008/ 56/EC). These directives demand both monitoring of cetacean population status (e.g. distribution, abundance) and enactment of conservation measures if population status is found to be unfavourable (see Santos & Pierce, 2015, for a discussion of the application of the MSFD to cetaceans). Marine conservation in the ENA is also covered by several international organizations and agreements, including the International Council for the Exploration of the Sea (ICES, http://www.ices.dk/), the Convention for the Protection of the Marine Environment of the North-East Atlantic (OSPAR, http://www.ospar.org/), the Agreement on the Conservation of Small Cetaceans of the Baltic and North Seas (ASCOBANS) and the Agreement on the Conservation of Cetaceans of the Black Sea, Mediterranean Sea and contiguous Atlantic area (ACCOBAMS).

In the ENA, common dolphins (*Delphinus delphis* Linnaeus, 1758), are among the most frequently sighted cetacean species (Alves *et al.*, 2018; Correia *et al.*, 2015; Goetz *et al.*, 2015; Hammond *et al.*, 2013; Jungblut *et al.*, 2017; Silva *et al.*, 2014; Tobeña *et al.*, 2016). Their distribution and habitat characteristics have been modelled in relation to geographic, physiographic, oceanographic and fishing-related variables, and several studies have identified well-defined habitat preferences related to the abundance of prey, for example productive areas (i.e. upwelling regions), with low to medium sea-surface temperatures, mostly coastal and shallow but often deeper waters, and/or areas that concentrate their preferred prey (e.g. Cañadas & Hammond, 2008; Correia *et al.*, 2015; Goetz *et al.*, 2015; Halicka, 2016; Moura *et al.*, 2012; Pierce *et al.*, 2010; Tobeña *et al.*, 2016). Their apparently patchy distribution suggests that common dolphins, although widely distributed, have a well-defined habitat and they may be dietary specialists in the sense of feeding on schooling fish (Marçalo *et al.*, 2018; Moura *et al.*, 2012). Common dolphins usually target high energy prey and/or locally abundant pelagic schooling fish and some of their prey have high commercial value, such as sardines, bluewhiting, anchovy, sprat and horse mackerel, which often results in interactions of feeding dolphins with fisheries (e.g. Marçalo *et al.*, 2018; Meynier *et al.*, 2008; Santos *et al.*, 2013; 2014). In fact, negative impacts of fishery by-catch mortality and/or prey depletion due to overfishing of common dolphin prey have been widely reported. For example, in the Bay of Biscay, by-catch has been suggested to have reached unsustainable levels, inconsistent with the maintenance of common dolphin populations at a favourable status (Peltier *et al.*, 2016). In the Mediterranean, overfishing is probably one of the causes for the estimated 50% decline in abundance of this species in the last 45 to 35 years, leading the Mediterranean subpopulation of common dolphins to be listed as endangered in the IUCN Red List of Threatened Species (Cañadas & Vázquez, 2017; Piroddi *et al.*, 2011).

Common dolphin occurrence in coastal areas of the ENA (Djiba *et al.*, 2015; Goetz *et al.*, 2015; Hammond *et al.*, 2013; Moura *et al.*, 2012; Weir, 2010; Weir & Pierce, 2013) and around the islands of Macaronesia (Alves *et al.*, 2018; Carrillo *et al.*, 2010; Halicka, 2016; Hazevoet & Wenzel, 2000; Silva *et al.*, 2014; Tobeña *et al.*, 2016) is reasonably well reported, but in the high seas, where logistic constraints impede systematic surveys for cetacean monitoring, data are still lacking and spatial and temporal distribution of this species is poorly known (Correia *et al.*, 2015; Jungblut *et al.*, 2017). This baseline knowledge is fundamental to further assess the conservation status of the species and the impacts of human activities on its distribution, and to efficiently manage the status of common dolphins in the North Atlantic. In 2012, a monitoring project started collecting cetacean occurrence data in the ENA using cargo vessels as observation platforms of opportunity (OPOs) along routes from Continental Portugal to the Macaronesian archipelagos and north-west Africa (Correia *et al.*, 2015). In the present study, the occurrences recorded in the surveys from 2012 to 2016 were used to analyse the spatial and temporal distribution of common dolphins. Four different models were developed to describe (i) the influence of detectability factors (observation effects model), (ii) dolphin distribution across space and time (spatiotemporal model), (iii) the influence of topographic and oceanographic features (environmental model) and (iv) a combination of all the above (final habitat model). We evaluate the usefulness of data collected from surveys on OPOs to develop habitat models and to identify important areas for conservation across a wide area of ocean. The results are expected to contribute to status evaluations by international organizations that have responsibility or interest in the conservation of cetaceans, and to support legal instruments for the management of the area.

6.3 Material and methods

6.3.1 Study area

As part of the CETUS Project (http://www.cetusproject.com/), data on cetacean occurrence were collected within the ENA. The study area included the coastal waters of mainland Portugal and of north-west Africa, the waters in between (oceanic) and within the Macaronesian archipelagos: the Azores and Madeira (Portugal) and the Canary (Spain) and Cape Verde islands (Figure 1).

From 2012 to 2016, surveys for cetacean occurrence took place during 99 round-trips aboard cargo ships belonging to TRANSINSULAR, a Portuguese maritime transport company. The cargo ships were used as OPOs and each followed one of three different routes, all starting and ending in mainland Portugal, to the Azores, Madeira and Cape Verde respectively, with a total of 15 ports visited, 10 of them routinely (Figure 1).

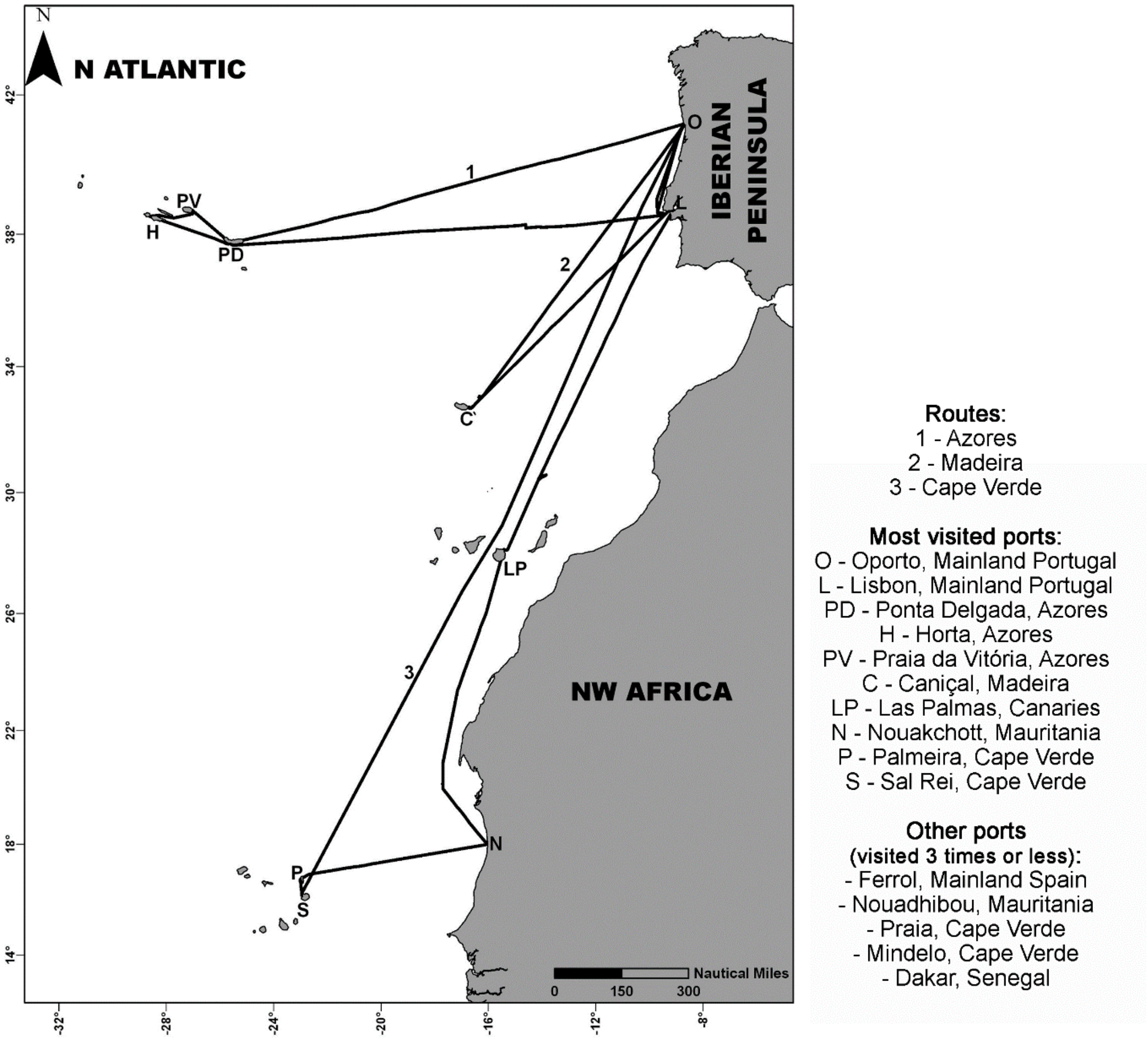


Figure 1. The study area within the eastern North Atlantic, with surveyed transects and visited ports.

Most surveys were conducted during summer months (from July to October) with favourable weather conditions for cetacean sampling, especially considering North Atlantic offshore areas where sea conditions are generally rough during the rest of the year (Table 1).

Table 1. Survey effort, sightings of common dolphin (*Delphinus delphis*), group size and total encounter rates, for each sampled route and season of survey.

Route	Year	Season	No. of trips / No. of surveys	Survey effort	Total sightings / Sightings on effort	Group size min-max (mean ± SD)	ER
Madeira	2012	July-October	9 / 19	5025	17 / 14	1-40 (12.21 ± 10.19)	0.28
	2013	July-October	13 / 29	5616	30 / 22	1-120 (15.91 ± 29.16)	0.39
	2014	August-October	11 / 23	3938	22 / 16	2-100 (18.31 ± 24.54)	0.41
	2015	July-October	18 / 44	6009	30 / 21	1-80 (20.95 ± 19.29)	0.35
	2016	July-October	16 / 46	4887	28 / 19	2-100 (12.95 ± 22.62)	0.39
Azores	2014	July-September	6 / 32	5556	30 / 19	2-40 (8.16 ± 9.50)	0.34
	2015	July-October	7 / 33	3444	37 / 14	3-80 (21.86 ± 20.53)	0.41
	2016	July-October	7 / 31	3920	26 / 21	1-50 (16.48 ± 15.77)	0.54
Cape Verde	2015	May-October	7 / 46	8723	29 / 24	3-2500 (168.29 ± 510.22)	0.28
	2016	February/ August-December	5 / 42	6203	34 / 22	2-40 (13.32 ± 10.58)	0.35
TOTAL			99/ 345	53321	283 / 192	1-2500 (34.58 ± 185.04)	0.36

A trip is considered a round-trip starting and ending in mainland Portugal while a survey is a leg between two ports. Survey effort is presented in nautical miles (nmi) rounded to the unit. For the group size, the minimum (min), maximum (max), mean and standard deviation (SD) values presented are based on the best estimate of the number of animals per sighting on effort, accessed by the observer.

6.3.2 Data collection

In situ

For each route, two observers were trained in use of survey protocols by the project team and then boarded TRANSINSULAR cargo ships to visually monitor cetaceans throughout the trips. Travel speed generally varied from 11 to 16 knots. Surveys were performed from sunrise to sunset, whenever weather conditions were favourable (with sea state and wind speed up to 4, on the Douglas and Beaufort scales respectively, and visibility over 1 km) and the ship was sailing outside the ports. Surveys stopped occasionally during periods when observers were not allowed at the observation stands, i.e. during safety drills, cleaning of the deck or manoeuvres. Observers stood in the navigation bridge and wings of the bridge, at an approximate height of 20 m above sea level (depending on the loading of the ship) and searched for cetacean presence through 180°, centred on the ship's heading, with and without binoculars (magnification of 7 × 50 mm, with scale and compass). When cetaceans were sighted, the species was identified and number of individuals recorded. When it was not possible to determine the exact number of individuals, a minimum and maximum number of animals was recorded, as well as the most probable number of individuals according to the observer's perception (best estimate). Besides cetacean occurrence, data on the presence of other top predators (e.g. turtles, sharks, tuna), as well as information on weather conditions and marine traffic, were collected. For more details on sampling protocol, see Correia *et al.*

(2015). Since the present paper is focused on common dolphins (*Delphinus delphis* Linnaeus, 1758), results for other species will be presented elsewhere.

Remote sensing

For the statistical habitat modelling, in addition to weather conditions and spatiotemporal variables needed for both observational and spatiotemporal models, habitat variables were derived from satellite data at several temporal and spatial scales (see Table 2). Slope was derived from bathymetry data. For distance to seamounts, topographic features classified as seamounts, banks, hills, ridges and rises in GEBCO (GEBCO, 2017) were delimited, using contour lines created every 50 m, and defining a polygon from the outermost closed contour line around the geographic location of the top of the features. Then, the distance from the base of the seamounts and from the coastline (distance to coast) to the sightings was calculated. Both slope and distances were computed using ArcGIS 10.5 (ESRI, 2016).

For dynamic variables, satellite data were used. Chlorophyll-a and sea-surface temperature are ocean products derived from the satellite MODIS – Aqua Mapped data from NASA (NASA, 2017). The algorithms return the near-surface concentration of chlorophyll-a (from *in situ* remote sensing reflectance) and temperature (from measured radiances). Both variables were extracted at two different temporal and spatial scales. Chlorophyll-a was extracted for the calendar month and week in which the sightings occurred but also with four different time lags (one and two weeks and months of lag). For altimetry, the mean sea level anomalies were obtained from Ssalto/Duacs multimission altimeter products provided by AVISO (AVISO, 2017). The sea level anomalies are sea-surface heights computed with respect to a 20-year mean profile (1993–2012). When assembling data for sea level anomalies, delayed products were available only until 5 May 2016 and, as a consequence, near-real time products were used for July–October 2016. Near-real time final products become available six days after the date of measurement, but are less precise than delayed products, which become available around two months after collection, having been re-analysed and re-processed (AVISO, 2017). For this variable, weekly and monthly resolutions were computed by averaging daily products.

Table 2. Variables tested as predictors for statistical modelling and its characteristics.

Model	Variables	Source	Reference	Product name	Name used in the analysis	Spatial resolution	Temporal resolution	Unit
Observation effects	Sea-state	Sea-surveys	-	-	sea_state	-	-	Douglas scale
	Wind-state	Sea-surveys	-	-	wind_state	-	-	Beaufort scale
	Visibility	Sea-surveys	-	-	visibility	-	-	1-10 scale[a]
Spatiotemporal	Latitude	Sea-surveys (GPS)	-	-	lat	-	~10 seconds	Decimal degrees
	Longitude	Sea-surveys (GPS)	-	-	lon	-	~10 seconds	Decimal degrees
	Day of the year	Date of survey	-	-	day	-	Daily	Day
	Year	Year of survey	-	-	year	-	Yearly	Year
Environmental	Depth	GEBCO	GEBCO, 2017	bathy_30arc_second	depth	30 sec	-	Meters (m)
	Slope	GEBCO	GEBCO, 2017	-	slope	30 sec	-	Degrees (º)
	Distance to coast	-	-	-	dist_coast	-	-	Kilometres (km)
	Distance to seamounts	GEBCO	GEBCO, 2017	-	dist_sm	-	-	Kilometres (km)
	Chlorophyll	MODIS Aqua	NASA, 2017	CHL_chlor_a	CHL	4 km / 9 km	8 day / monthly	Density (mg m^{-3})
	Chlorophyll lag 1 week	MODIS Aqua	NASA, 2017	CHL_chlor_a	CHL_lag1w	4 km / 9 km	8 day / monthly	Density (mg m^{-3})
	Chlorophyll lag 2 weeks	MODIS Aqua	NASA, 2017	CHL_chlor_a	CHL_lag2w	4 km / 9 km	8 day / monthly	Density (mg m^{-3})
	Chlorophyll lag 1 month	MODIS Aqua	NASA, 2017	CHL_chlor_a	CHL_lag1m	4 km / 9 km	8 day / monthly	Density (mg m^{-3})
	Chlorophyll lag 2 months	MODIS Aqua	NASA, 2017	CHL_chlor_a	CHL_lag2m	4 km / 9 km	8 day / monthly	Density (mg m^{-3})
	Sea-surface temperature	MODIS Aqua	NASA, 2017	sst4_4_sst4	SST	4 km / 9 km	8 day / monthly	Celsius (ºC)
	Mean sea level anomalies	AVISO	AVISO, 2017	MSLA_h_DT_all_sat_merged_0.25 / MSLA_h_NRT_all_sat_merged_0.25	MSLA	0.25 degree	8 day / monthly	Centimetres (cm)
Final	All variables above							

[a] Visibility scale: 5 - 1 to 2 km; 6 - 2 to 4 km; 7 - 4 to 10 km; 8 - 10 to 20 km; 9 - 20 to 50 km; 10 - > 50 km. Below 5 (1 km of visibility), the survey stopped (off effort).

6.3.3 Data analysis

Total and on effort sightings of common dolphins per season of survey and route were computed, as well as the survey effort. On effort sightings are those recorded during survey effort, while total number includes off effort sightings recorded opportunistically. The group size (minimum, maximum, mean and standard deviation values) was accessed from the recorded best estimate for the number of individuals in the group (Table 1). For the remaining analyses, an individual sighting was used as the sampling unit, regardless of the group size. Encounter rates were computed as the total number of sightings on effort per 100 nautical miles (nmi) surveyed, for each season and route. Then, the spatial and temporal distributions of common dolphin occurrences were analysed for the entire study area (considering data from the three routes), computing geographic positions and monthly variation of sightings, survey effort and encounter rate.

Statistical modelling was performed using Generalized Additive Models (GAMs), which have been widely used to describe cetacean distribution and habitat characteristics. An approach based on used / available habitat was chosen (Correia *et al.*, 2015; Elith & Leathwick, 2009; Pearce & Boyce, 2006), with used (common dolphin sightings on effort) and available (survey route) habitat points combined to generate a binary (1,0) response variable. The set of available points was created as in Correia *et al.* (2015), through the creation of equidistant points (every 2.5 nmi) along all effort tracks. Using this methodology guarantees that areas that had a higher survey effort are given more points of available habitat, hence, survey effort is being taken into account in the models. The values of the variables to use as predictors in the modelling process were extracted from the set of used and available points (Table 2). For oceanographic variables, the pack of tools for ArcGIS, Marine Geospatial Ecology Tools (MGET) (Roberts *et al.*, 2010) was used.

Prior to modelling, Pearson correlation between explanatory variables was computed to avoid using highly correlated variables in the same model (threshold of 0.75) (after Marubini *et al.*, 2009). Distance to coast and depth were the only pair of variables highly positively correlated. Since both were of interest, a GAM model was fitted, with depth as predictor and distance to coast as response variable, and both depth and the residuals of this model were used as predictors in the common dolphin models (see Smith *et al.*, 2011). Moreover, multiple correlation among explanatory variables was assessed through the Variance Inflation Factor (VIF, with a threshold of 3) (Zuur *et al.*, 2010). After replacing distance to coast by the residual distance as described above, all remaining variables had VIF values < 3 and no additional variables were removed.

A binomial distribution was assumed for the response variable and a maximum of four splines was used (k-fold set to 4) to limit the complexity of smoothers describing effects of explanatory variables. Model fitting mainly involved backward selection, starting from an oversaturated model (Correia *et al.*, 2015; Quian, 2009; Viddi *et al.*, 2010). However, forward selection was undertaken when choosing between the different scales of the oceanographic variables (and different time lags for chlorophyll). Interactions between spatial and temporal variables were also explored in the fitting process to account for main and interaction effects: interaction between latitude with longitude and between year with day of the year. This was done by including these pairs of variables in two dimensional smoothers and visualizing the results as surface plots (in this case, the k-fold was set to 16 as to account for the interaction effect, i.e. four times four).

Following Correia *et al.* (2015), and to account for varying dolphin group size, a weight parameter was included in the models, corresponding to the best estimate of animals sighted for each observation. Given the wide range of group size and high uncertainty of the estimations, weights were attributed in categories: a small group – from one to five animals (weight = 1); a medium group – from six to 20 animals (weight = 2); a large group – more than 20 animals (weight = 3). A weight of 1 was set for points of available habitat.

Best models were selected by using the Akaike Information Criterion (AIC) as a measure of goodness of fit, choosing the model with the lowest AIC value at each step of the model fitting process, i.e. comparing otherwise identical models with or without a specific explanatory variable. If the difference in AIC values between two models was less than 2, a chi-squared test was applied. Whenever differences between AIC values were not statistically significant (based on δAIC > 2 or the chi-squared test result), the simplest model was maintained (following the principle of parsimony, e.g. Burnham & Anderson, 2002). Finally, at the end of the modelling process, the models were evaluated by creating two random subsets of data: fitting and evaluating sets (75% and 25% of the data, respectively). Prediction power of the models was determined using the Area Under the Curve (AUC) metric of the Receiving Operator Characteristic (ROC) curve (Beck & Shultz, 1986).

Four different models were developed, three of these to specifically evaluate, respectively (i) variables affecting cetacean detection (observation effects model), (ii) spatiotemporal variation (spatiotemporal model) and (iii) habitat preferences (environmental model). Model iv, the final habitat model, used a combination of all the variables tested (Table 2) and was then used to predict probabilities of common dolphin occurrence at the set of used/available points along the routes. Prediction was done using all the original data values for explanatory variables. Finally, predicted probabilities of dolphin occurrence at the points were represented in a map.

Maps were created in ArcGIS 10.5 (ESRI, 2016) using a Mercator projection (EPSG: 4326), graphs in Microsoft Excel 2016 and statistical modelling was carried out using R (R Development Core Team, 2012) with R Studio.

6.4 Results

6.4.1 Survey effort

Most of the survey effort was during summer months, from July to October. A total of 2073 sightings was collected and 26 species identified (at least to genus), with 17 species occurring along the Madeira route, 11 along the Azores route and 25 along the Cape Verde route. Sighted species included baleen whales, toothed whales, dolphins and porpoises, with most sightings being of dolphin species. With a total of 25,475 nmi surveyed, the route to Madeira was the most sampled, being surveyed since 2012 (Table 1). Survey effort was heterogeneous across the sampled transect with some gaps due to periods of bad weather conditions as well as areas crossed during night time (Figure 2).

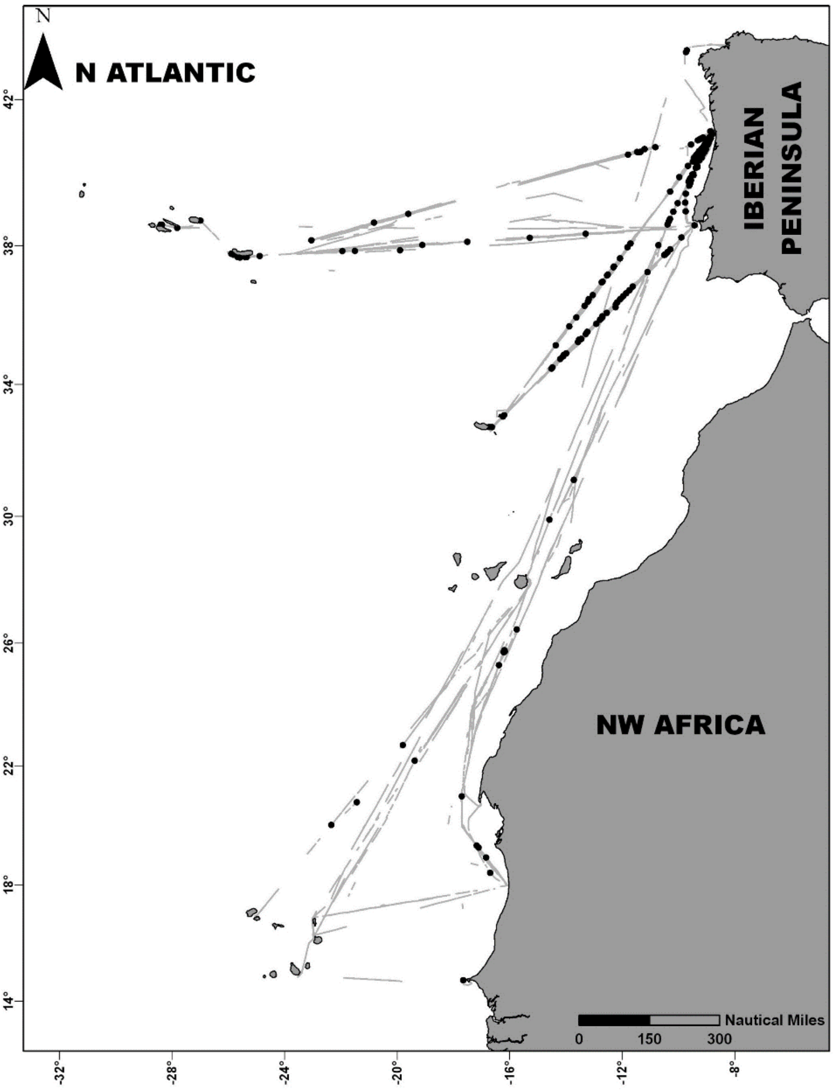


Figure 2. Spatial distribution of common dolphin (*Delphinus delphis*) occurrences with survey effort transects represented in grey lines. Only sightings on effort are represented.

6.4.2 Spatiotemporal distribution of common dolphins

Common dolphin (*Delphinus delphis* Linnaeus, 1758) was the most frequently sighted species (283 sightings, ~ 14% of the all species total), present over a wide latitudinal range, but mostly sighted in northern latitudes within the sampled area, with fewer occurrences south of Madeira Island (Table 1 and Figure 2).

There were 192 on effort sightings of common dolphins, giving an overall encounter rate of 0.36 sightings / 100 nmi (Table 1). Common dolphin groups varied in size between one and 2500 animals and encounter rates (by route and by year) ranged from 0.28 sightings / 100 nmi (2012 along the Madeira route and 2015 on the Cape Verde route) to 0.54 sightings / 100 nmi

(2016 on the Azores route) (Table 1). The largest group, of 2500 animals, was recorded off Dakar, in 2015 (Figure 2).

The highest monthly number of common dolphin sightings on effort (20) was in August 2016, while the highest monthly encounter rate (0.73 sightings/100 nmi) was recorded in October 2013, with 10 on effort sightings over 1370 nmi surveyed. No common dolphin sightings were registered in the months with the lowest survey effort (February, March and December 2016) (Figure 3).

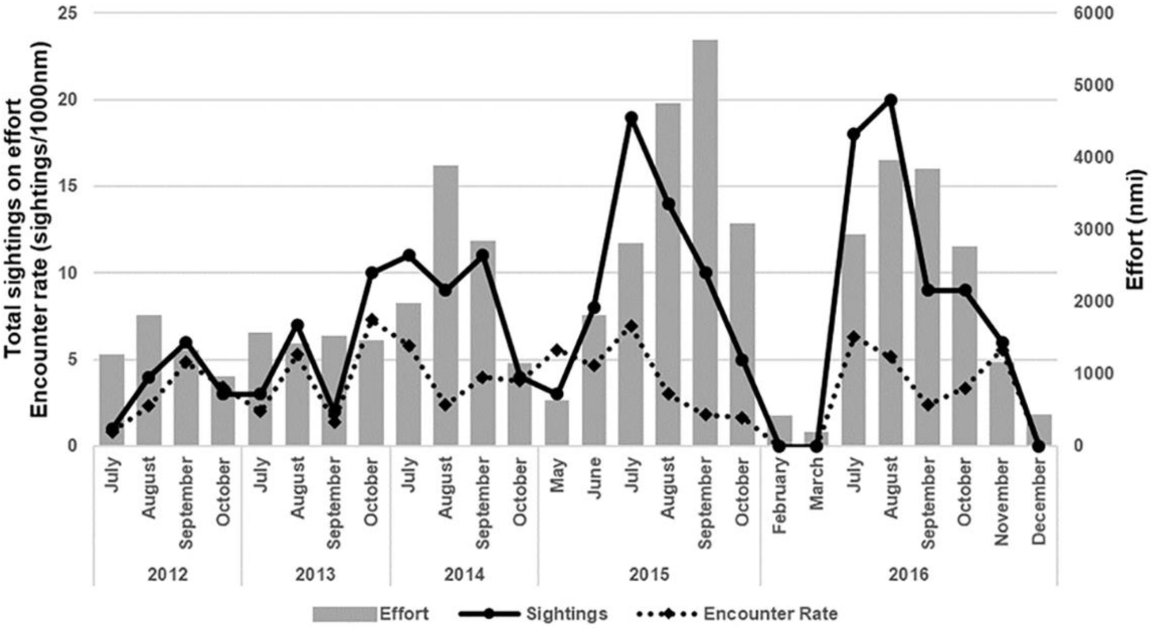


Figure 3. Temporal variation of common dolphin (Delphinus delphis) occurrence, encounter rate and monthly survey effort in nautical miles (nmi). Data from the entire study area, within the eastern North Atlantic, are summarized. Only sightings on effort are considered.

6.4.3 Modelling

Of the three initial models, the model fitted for observation effects had the lowest deviance explained (4.11%) and AUC (0.689), while the spatiotemporal model had a slightly higher deviance explained (16.5%) than the environmental model (15.5%). The final habitat model had the highest deviance explained (22.3%) and included variables from all the three models above (Table 3).

Table 3. Best GAM model results for common dolphin (*Delphinus delphis*).

Model Parameters	Estimate	edf	SE	z-value	Chi-square	p-value	Deviance explained (%)	r²	UBRE	AUC (CI 95%)
Observation effects										
Intercept	-4.18		0.06	-68.32		<0.001				
Smoother terms										
sea_state		1.61			26.34	<0.001				
wind_state		2.92			135.44	<0.001				
visibility		2.88			11.62	0.008				
Best model (n=20388; 192 presences): CD~s(sea_state)+s(wind_state)+s(visibility)							4.11	9.41^{E-3}	-0.82	0.689 (0.619-0.758)
Spatiotemporal										
Intercept	-4.83		0.10	-49.82		<0.001				
Smoother terms										
lat, lon		14.65			464.95	<0.001				
day,year		12.33			81.95	<0.001				
Best model (n=20388; 192 presences): CD~s(lat,lon)+s(day,year)							16.5	0.06	-0.84	0.809 (0.727-0.891)
Environmental										
Intercept	-4.81		0.21	-22.52		<0.001				
Smoother terms										
depth		1.84			160.32	<0.001				
resid_dist_coast		2.75			21.34	<0.001				
slope		2.87			9.69	0.017				
dist_sm		2.80			18.27	<0.001				
CHL_9km_monthly		2.58			19.12	0.005				
SST_4km_8day		2.83			22.42	<0.001				
MSLA_8day		2.83			9.83	0.015				
Best model (n= 16706; 165 presences): CD ~ resid_dist_coast+s(depth)+s(slope)+s(dist_sm)+ s(SST_4km_8day)+s(CHL_9km_monthly)+s(MSLA_8day)							15.5	0.05	-0.84	0.744 (0.651-0.838)
Final										
Intercept	-4.73		0.10	-49.89		<0.001				
CHL_9km_monthly	-0.45		0.07	-6.09		<0.001				
Smoother terms										
wind_state		2.87			45.16	<0.001				
visibility		2.93			17.98	<0.001				
lat,lon		14.39			104.02	<0.001				
day,year		11.03			66.88	<0.001				
depth		1.04			26.54	<0.001				
dist_sm		2.84			16.17	0.001				
MSLA_8day		2.97			28.66	<0.001				
Best model (n= 19658; 189 presences): CD ~ CHL_9km_monthly+s(wind_state)+s(visibility)+s(lat,lon)+s(day,year)+ s(depth)+s(dist_sm)+s(MSLA_8day)							22.3	0.09	-0.85	0.727 (0.639-0.814)

edf, effective degrees of freedom; SE, standard error; AUC, Area Under the Curve; CI 95%, 95% confidence interval for the AUC; N, total number of points (used/available) considered in the model fitting; CD, common dolphins; resid_dist_coast, residuals from the model for distance to coast with depth as predictor. For other parameters abbreviations, see Table 1.

All the three variables tested, namely sea state, wind state and visibility, contributed to the observation effects model. Sea state had a positive effect over the range Douglas 2–4, visibility had an overall positive influence, albeit with a negative effect apparent at intermediate visibilities (range 7–8), and wind-state had a negative influence over the range Beaufort 1–3 (Figure 4).

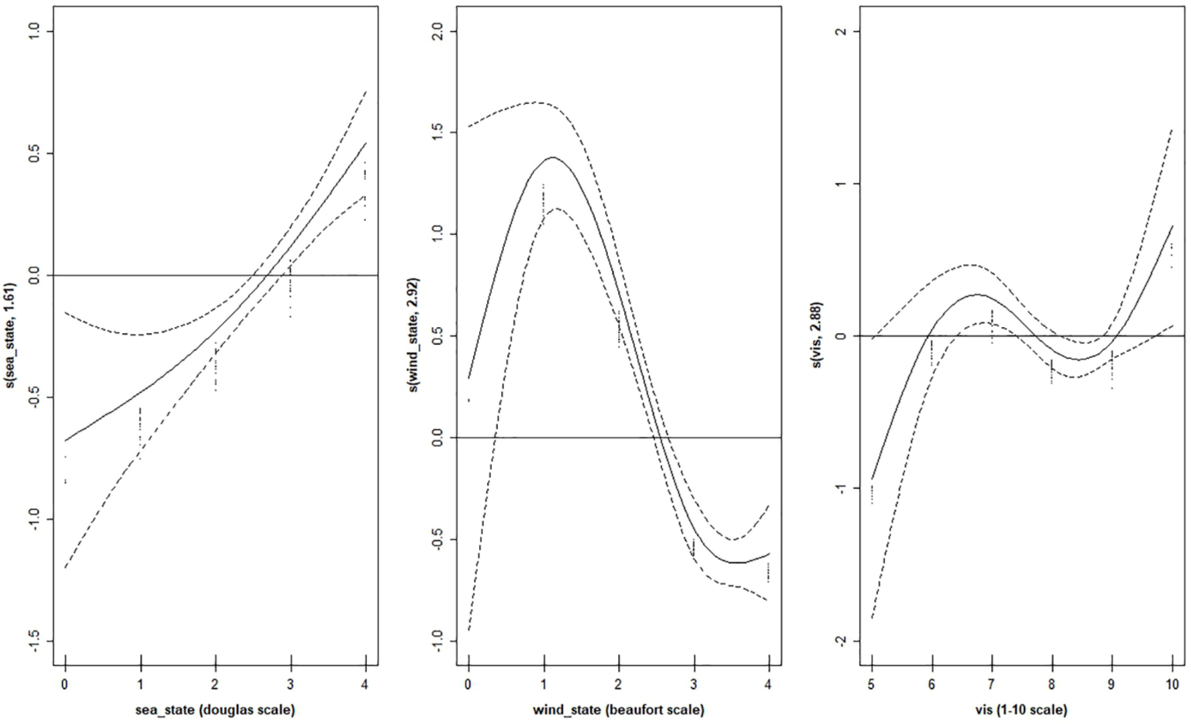


Figure 4. GAM predicted splines of the response variable dolphin presence as a function of the explanatory variables for the observation effects model produced for common dolphin (*Delphinus delphis*). The degrees of freedom are in parentheses on they-axis. Tick marks above thex-axis indicate the distribution of observations. Dashed lines delimit the 95% confidence intervals of the spline functions and dots on the graph area represent the residuals. For parameters abbreviations, see Table 2.

The spatiotemporal model included latitude×longitude and year × day effects (i.e. main effects and interactions). There were positive effects at several different geographic locations within the surveyed area: northern latitudes with eastern longitudes, corresponding to the proximities of continental Portugal; northern latitudes with western longitudes, corresponding to the Azores region; and a smaller peak at southern latitudes with eastern longitudes, along the African coast. As for the temporal variables, the surface of the year × day of year plot varies along the day of year axis with the same pattern seen across all years. A peak is observed in the beginning of the survey season (July), sightings rate decreasing thereafter and with a smaller peak at the end (October) (Figure 5).

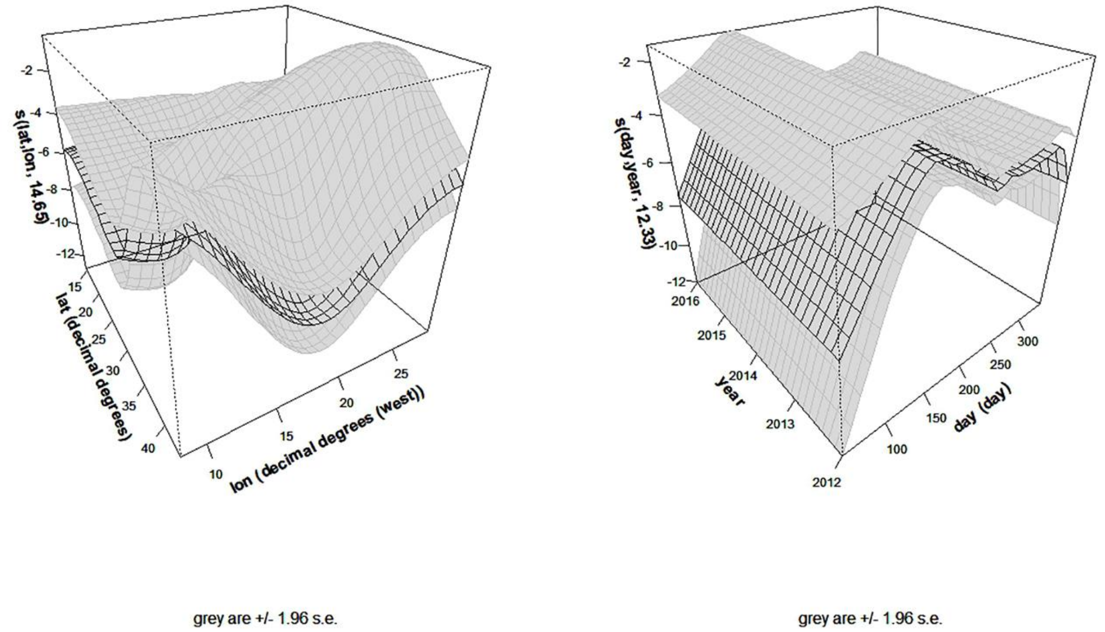


Figure 5. GAM predicted perspective graphs of the response variable dolphin presence as a function of the explanatory variables for the spatiotemporal model produced for common dolphin (*Delphinus delphis*). These correspond to variables introduced as interactions in the model, spatially (latitude × longitude) and temporally (day of the year × year), and represent in a surface the variation along the two variables. The degrees of freedom are in parentheses on the z-axis. Grey surfaces define the upper and lower limits of the 95% confidence interval. For parameters abbreviations, see Table 2.

The environmental model included seven environmental variables: depth, residuals from the model of distance to coast vs depth, slope, distance to seamounts, chlorophyll concentration, sea surface temperature and mean sea level anomaly. Depth had an almost linear negative correlation with common dolphin occurrence, i.e. there was a lower probability of sightings over deeper waters. As for the residuals from the model of distance to coast vs depth, GAM results indicate that, for a given depth, there is a positive influence of proximity to coastal areas. In relation to seabed slope, there was a peak in sightings probability at ~ 5° of slope, with predicted dolphin presence then decreasing over steeper slopes. Distance to seamounts had a negative effect up to 300 km and then a positive effect towards areas most distant from seamounts. Both chlorophyll and sea-surface temperature had a broadly negative effect, while for mean sea level anomaly there was a negative correlation between 0.07 cm and 0.15 cm but also a probable positive correlation at higher anomaly values (where, however, the confidence interval is wide). While sea surface temperature and mean sea level anomalies had the highest explanatory power at the finest spatial and temporal resolutions (8-day for both and 4 km for sea-surface temperature), chlorophyll presented a strong relationship with sightings at the lowest resolution, both spatially and temporally, and with no lags (Figure 6).

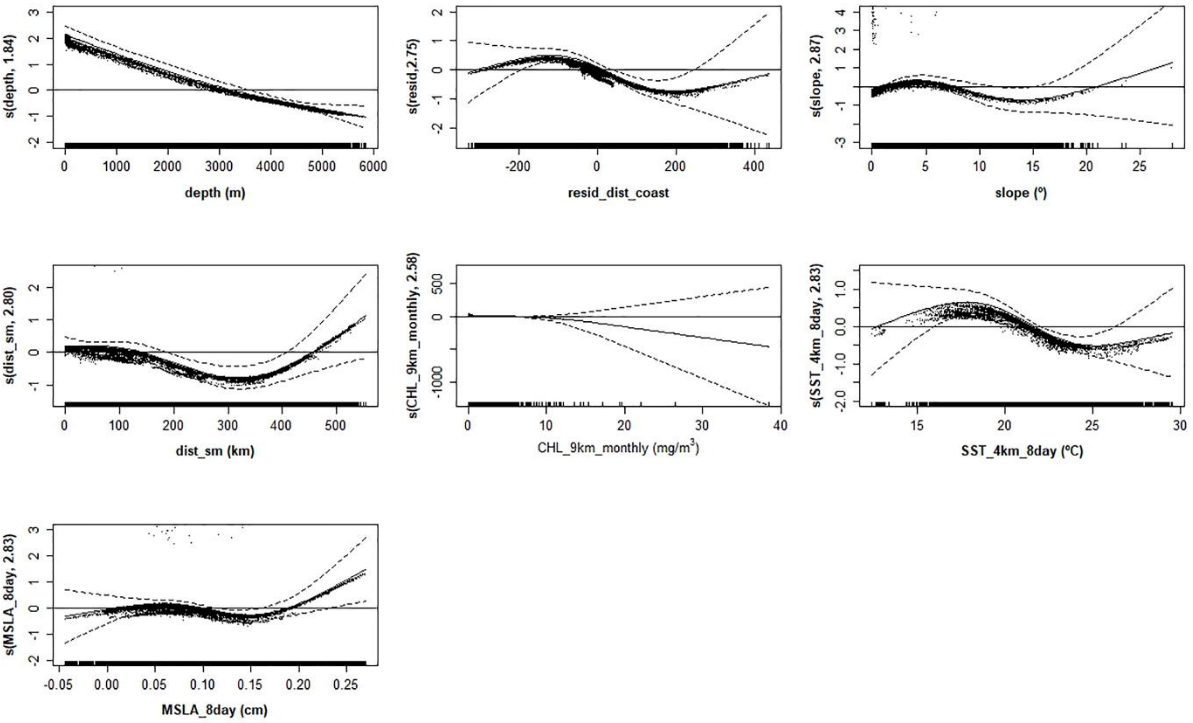


Figure 6. GAM predicted splines of the response variable dolphin presence as a function of the explanatory variables for the environmental model produced for common dolphin (*Delphinus delphis*). The degrees of freedom are in parentheses on the y-axis. Tick marks above the x-axis indicate the distribution of observations. Dashed lines delimit the 95% confidence intervals of the spline functions and dots on the graph area represent the residuals. resid_dist_coast – residuals from the model for distance to coast with depth as predictor. For other parameters abbreviations, see Table 2.

The final habitat model, where all the variables were tested during the fitting process, included 10 variables with two interactions among variables, namely the spatial (latitude with longitude) and temporal (day of the year with year of survey) variables. By introducing the dynamic variables, chlorophyll and sea-surface temperature, the total number of observations decreases (from 192 to 165), and consequently the number of available habitat points also decreases, as these variables were collected from satellite data and presented several data gaps (Table 3). While combining all predictors, the effects illustrated by the smooth curves for the variables included remain similar to their forms in the previous models. Dolphin presence was negatively and linearly related to chlorophyll concentration. The relationship between sightings and depth was approximately linear and also negative. The other variables had non-linear fits, with more complex relationships with the response variable. In general, probability of common dolphin detection was highest with low wind speed (low values on the Beaufort scale) and very good visibility. Common dolphin occurrence was more likely in areas further than 300 km from seamounts and at locations of intermediate and high positive sea level anomalies. Occurrence varied spatially (with peaks in Portuguese and African coastal areas and Azorean islands) with a relatively consistent seasonal pattern over the years of the survey (increase in the beginning of the season and small peak at the end) (Figure 7).

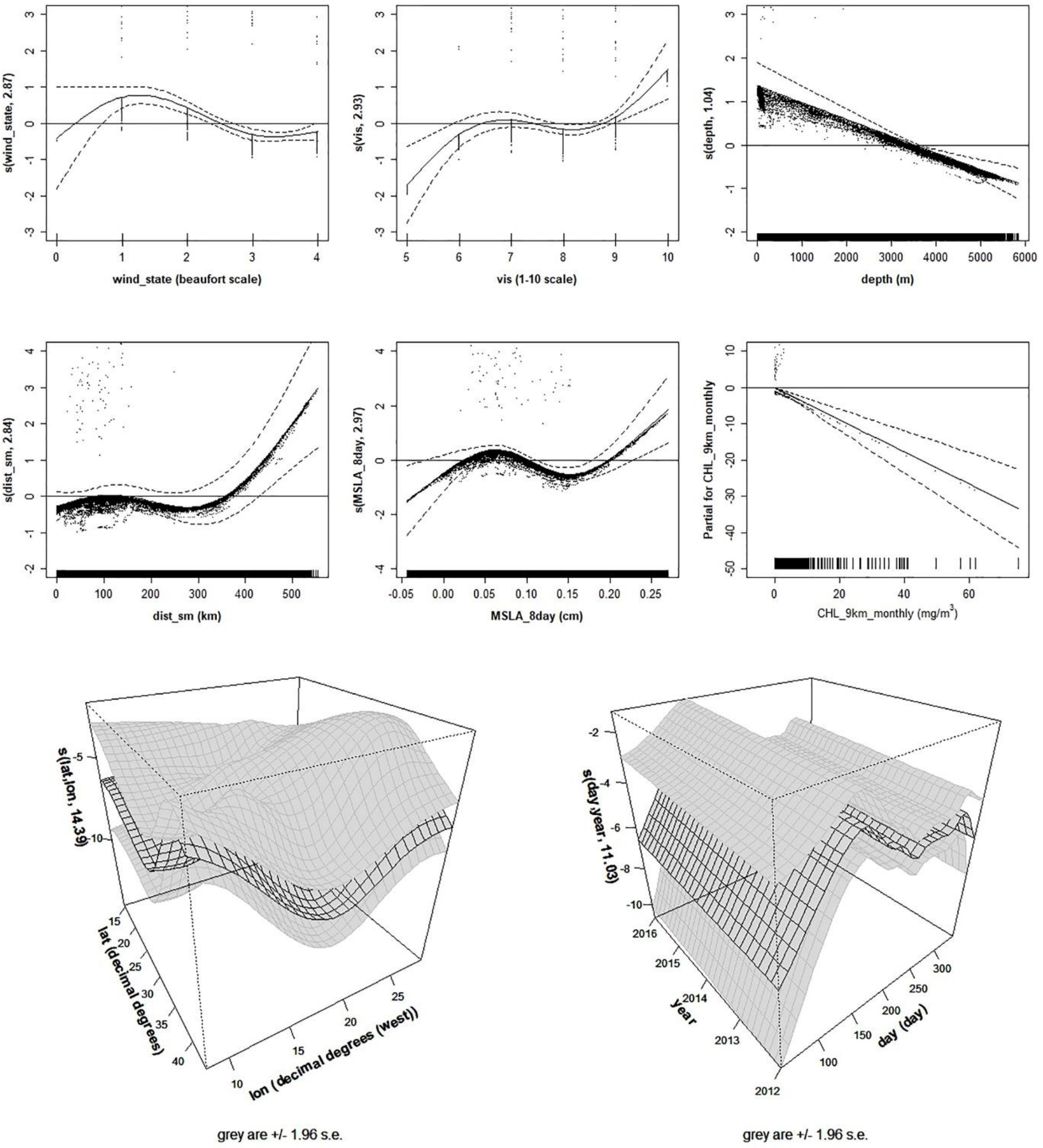


Figure 7. GAM predicted splines of the response variable dolphin presence as a function of the explanatory variables for the final model produced for common dolphin (*Delphinus delphis*). The degrees of freedom for non-linear fits are in parentheses on the y-axis. Tick marks above the x-axis indicate the distribution of observations. Dashed lines delimit the 95% confidence intervals of the spline functions and dots on the graph area represent the residuals. Perspective graphs correspond to variables introduced as interactions in the model, spatially (latitude × longitude) and temporally (day of the year × year), and represent in a surface the variation along the two variables. In these graphs, the degrees of freedom are in parentheses on the z-axis and grey surfaces define the upper and lower limits of the 95% confidence interval. For parameters abbreviations, see Table 2.

When mapping probability of occurrence predicted by the final GAM habitat model, at the set of available and used points along the route, two main areas stood out as having the highest values for predicted probability of common dolphin occurrence (28–47%): coastal continental Portugal and the Azores archipelago. The areas of Madeira Island and in the open ocean close to continental Portugal and in front of the Nouadhibou port in Mauritania had intermediate probabilities of dolphin occurrence (10–28%) (Figure 8).

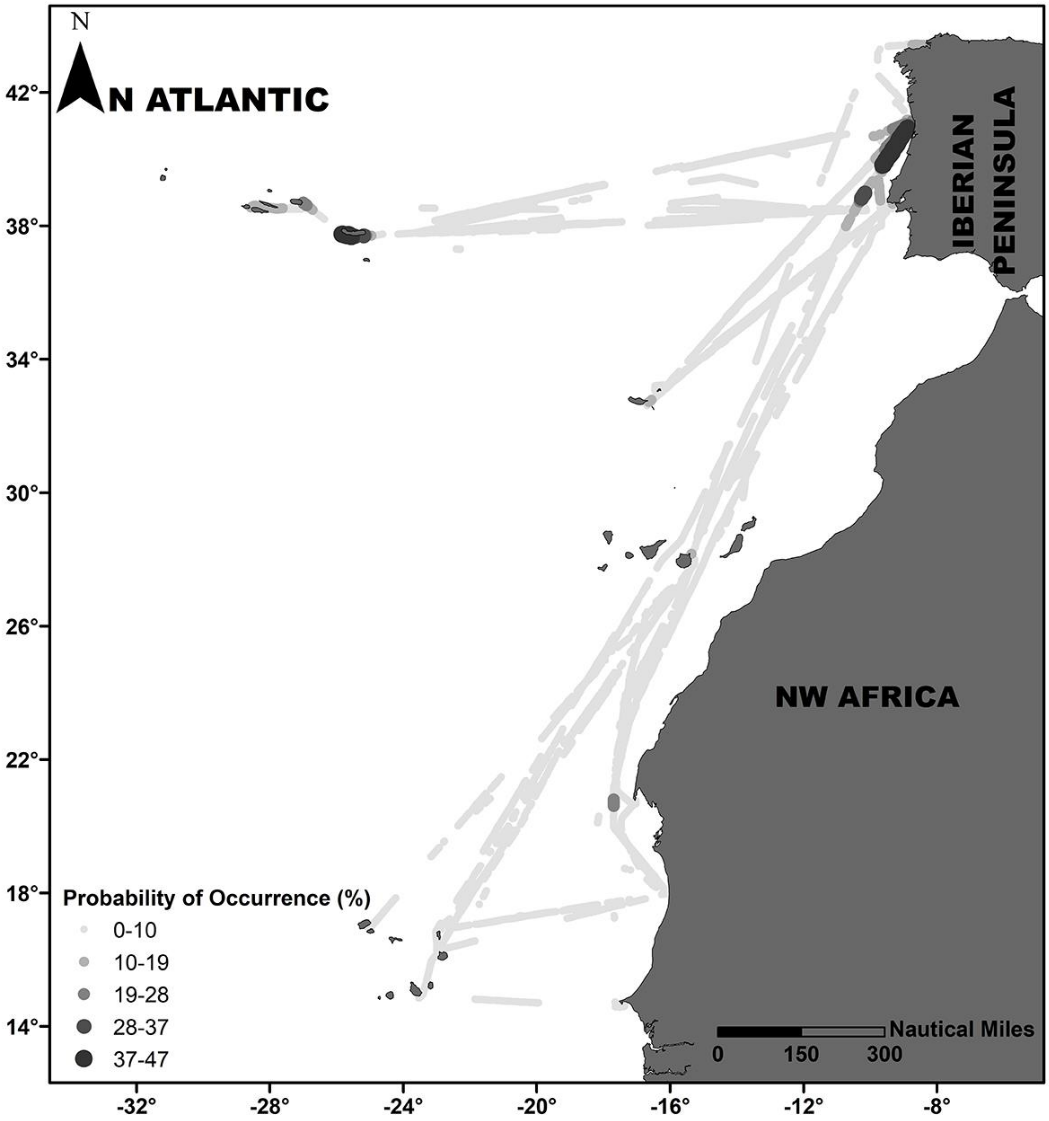


Figure 8. GAM predicted probabilities of occurrence of common dolphin (*Delphinus delphis*) for the set of the response variable points.

6.5 Discussion

This study presents the results from a 5-year data set on common dolphin (*Delphinus delphis* Linnaeus, 1758) occurrence from systematic surveys for cetacean monitoring in the ENA, with a great amount of effort carried out along a wide latitudinal range of about 30° latitude, mostly in poorly surveyed areas such as the high seas. Survey effort was concentrated in summer months, which is very common in marine surveys dependent on weather conditions (Kaschner *et al.*, 2010; 2012; Redfern *et al.*, 2006). Hence, results presented here reflect common dolphin distribution mainly for this period and few conclusions can be drawn for the remaining months of the year.

Common dolphin was the most frequently encountered species, accounting for 14% of the sightings across 26 species. This species has been reported as being among the most abundant in the area, however most studies present data mainly for coastal areas and islands (Alves *et al.*, 2018; Goetz *et al.*, 2015; Hammond *et al.*, 2013; Silva *et al.*, 2014; Tobeña *et al.*, 2016). On the contrary, the present study sampled mostly areas in the high seas. The biggest group of common dolphin, comprising ~ 2500 individuals, was recorded off Dakar in 2015. Large pods of dolphins have been registered previously in the coastal areas of north-west Africa (Bowman Bishaw Gorham, 2003; Camphuysen *et al.*, 2012; Djiba *et al.*, 2015; Weir *et al.*, 2014). The group size was highly variable, which is consistent with published results for coastal areas (e.g. Djiba *et al.*, 2015), islands (e.g. Alves *et al.*, 2018) and high seas (e.g. Correia *et al.*, 2015). Group size has been correlated with the water depth and, in the case of common dolphins, larger pods, frequently with calves, often occur closer to the coast (Cañadas & Hammond, 2008).

Spatially, common dolphin occurrences were most frequently registered over the shelf of continental Portugal and around the Azores and Madeira islands. There were also sightings along the entire Madeira route, which may be a consequence of higher survey effort but also an effect of the complex topography (Correia *et al.*, 2015; Schlacher *et al.*, 2010). Along the routes to the Azores and Cape Verde, there were areas with a total absence of sightings. No sightings of common dolphin were recorded in the Canaries and Cape Verde archipelagos. Our results for the Canaries are consistent with those from Carrillo *et al.* (2010) who reported the seasonal presence of common dolphins in the Canary Islands from December to May, the species being absent from June to November.

The year to year variation in common dolphin encounter rates did not present any clear pattern, which may relate to the spatial heterogeneity of survey effort. In fact, encounter rates peaked in different seasons in different years. In 2016, no encounters were registered in the months of February, March and December, but during these months only the route to Cape Verde was monitored and effort was very low.

Putative explanatory variables were chosen for the modelling process according to the effects they may have on the presence of common dolphins (based on the literature) but also reflecting availability. Observation effects were modelled to test whether the weather conditions that were likely to affect detection of dolphins strongly influenced the models. While detectability factors are not always included or tested in habitat modelling, their inclusion should provide more reliable results (Pierce *et al.*, 2010). While the variables tested did significantly affect the probability of seeing common dolphins, the observational effects model (as might be expected) had the lowest deviance explained of all the models (4.11%). Contrary to what was expected, sea state was positively correlated with common dolphin occurrence with probability of sighting increasing with higher wave height, at least in the range Douglas 2–3. This is probably due to the fact that common dolphins tend to surf down the leading edge of waves (possibly to save energy) and thus may be visible at the surface for longer if the waves are higher and wider. Nonetheless, this variable was then excluded from the final habitat model as it did not significantly affect common dolphin presence when considering the effects of the remaining predictors. Although weather conditions affect the detection of cetaceans which in turn influences model results (Pierce *et al.*, 2010), in this case, observation effects had a very low explanatory power; hence deviance explained in the final model is mainly related with the other predictors.

The spatiotemporal and environmental models had similar values of deviance explained, 16.5% and 15.5% respectively, likely to a large extent capturing the same variation since the best final model explained only 22.3% of deviance. Some habitat variables were excluded from the best final model while geographic location and temporal variables (days and years) were retained, presumably thus accounting for the effects of other habitat variables not being considered (Correia *et al.*, 2015; Elith & Leathwick, 2009; Pirotta *et al.*, 2011; Spyrakos *et al.*, 2011). Over three quarters of the variation in presence remains unexplained. In part this may be because relevant habitat variables were not included but it is also likely that many of the observed animals were travelling through less-preferred habitat.

In general, common dolphin probability of occurrence was higher in continental regions (continental Portugal and African coast) and in the area of Azores. As for seasonality, there seems to be a higher probability of occurrence at the beginning and the end of the survey season (July and October). However, this temporal trend should be interpreted with caution as there was substantial temporal heterogeneity in survey effort, which may be a source of noise in the analysis. If occurrence really is lower in the middle of the survey season, the question is whether this indicates animals moving out of the survey area (or at least away from the survey trackline) or a change in behaviour (e.g. aggregation, surfacing or response to boats). Nevertheless, the surface in the temporal perspective plot shows that common dolphin presence varies through the days of the year, with a pattern that remains relatively constant between years, pointing to a seasonal pattern. Seasonality of common dolphin occurrence in the different archipelagos of Macaronesia has been reported, in general, with higher abundances in cold months and a negative tendency during the summer months: in Madeira (Alves *et al.*, 2018; Halicka, 2016), Azores (Silva *et al.*, 2014; Tobeña *et al.*, 2016) and in the

Canary Islands (Carrillo *et al.*, 2010). The decrease of abundance in summer months is consistent with results presented here.

For the environmental variables, different spatial and temporal scales were tested. It has been shown that spatial and temporal scales affect model results and it is important to understand at which scale the impacts of the variable are significant for the presence of the species (González *et al.*, 2018; Fernandez *et al.*, 2017; 2018). Some of the variables included in the environmental model were dropped from the final combined habitat model during the fitting process. This probably reflects the fact that their effects are already explained by spatial and temporal variables and thus does not mean they are unimportant. However, depth, distance to seamounts, chlorophyll and sea level anomalies remained statistically significant in the final habitat model, increasing the overall deviance explained and having a clear influence in the spatiotemporal patterns.

Depth had an almost linear negative correlation with common dolphin presence. In the environmental model, the residual effect of the distance to coast (after taking depth into account) is negative, i.e. there is a preference for coastal waters. However, in the final habitat model this effect is probably being captured by longitude. A preference for shallower and coastal waters has been reported for common dolphins in several different studies, a result most likely due to the distribution of their preferred prey (Alves *et al.*, 2018; Cañadas & Correia *et al.*, 2015; Hammond, 2008; Meynier *et al.*, 2008; Moura *et al.*, 2012; Santos *et al.*, 2013; 2014; Stockin *et al.*, 2008), although strictly speaking we cannot prove whether diet choice follows from habitat choice or vice versa. Another suggestion for the coastal distribution is the presence of calves within the group (Alves *et al.*, 2018; Cañadas & Hammond, 2008; Stockin *et al.*, 2008). However, since this information was not collected in the present study, such a relationship could not be investigated. Most survey effort in previous studies was coastal, so the preferences of common dolphins could be reflecting sampled rather than preferred areas; in the present study, this is not the case as most effort was in deeper, offshore waters.

Although seamounts have a positive effect in cetacean presence, especially in the high seas where these structures act as oases of productivity in rather oligotrophic waters (Schlacher *et al.*, 2010), they did not seem to strongly influence common dolphin distribution. In fact, the model results indicate the highest probability of occurrences furthest from the seamounts (more than 300 km distance), which probably relates to the preference for coastal areas that are located furthest from the seamounts.

Sea surface temperature acts as a good indicator of upwelling phenomena that are characterized by the cold productive waters at the surface (Caldeira *et al.*, 2002; Mason, 2009). In the environmental model, an increase in sea surface temperature negatively affects common dolphin presence, pointing to a preference for colder waters. The ENA is characterized by strong coastal upwellings (Caldeira *et al.*, 2002; Mason, 2009), that are characterized by colder surface waters. This may explain the apparent preference for colder waters. The preference of common dolphins for more productive areas associated with strong upwellings has been reported before, as well as a tendency to prefer colder waters rather than warmer (sub-) tropical waters (Cañadas & Hammond,2008; Halicka, 2016; Jefferson *et al.*, 2009; Moura *et al.*, 2012; Stockin *et al.*, 2008). However, when including all the other variables, the sea-surface temperature does not significantly affect common dolphin distribution. This is probably because the sea-surface temperature pattern in the area is related to latitude, with a decrease of temperature from north to south, and distance to coast, with an abrupt decrease of temperature during coastal upwellings.

The surveyed area is highly dynamic and habitat is influenced by several current systems (Caldeira *et al.*, 2002; Mason, 2009). The sea level anomalies reflect this dynamism, probably not fully captured by spatial and temporal variables, and are related to productivity, being affected by upwelling and downwelling phenomena and currents that aggregate or disperse prey (Baird *et al.*, 2011; Davis *et al.*, 2002; Robinson, 2010). Two different temporal scales were tested for the altimetry data, with the 8-day resolution leading to the model with highest deviance explained. The fit indicates that common dolphin presence is more strongly affected at a weekly than a monthly scale, probably due to the high dynamism in the area. This also means that models would probably benefit from a better spatial resolution for altimetry, as the one available is rather low (0.25°, ~ 28 km). In the study area, there is a complex relationship between sea level anomalies and common dolphin presence, with a decrease in probability of occurrence at intermediate positive anomalies and an increase at more highly positive anomalies. This complex relationship may however indicate overfitting in the model.

In the case of the chlorophyll concentration, different temporal lags were also tested, besides the different spatial and temporal scales. The rationale is that chlorophyll is a proxy for productivity and there is a temporal lag (and possibly also spatial displacement) between chlorophyll blooms and high abundance of common dolphin prey (Frederiksen *et al.*, 2006; Grémillet *et al.*, 2008). Nonetheless, and contrary to the result for sea-surface temperature and altimetry, the chlorophyll had the highest explanatory power at the lowest resolution, both spatially and temporally, and with zero lag. Chlorophyll negatively affected common dolphin presence, contrary to what was expected (Cañadas & Hammond, 2008; Halicka, 2016; Moura *et al.*, 2012; Tobeña *et al.*, 2016). However, the influence of chlorophyll reflected in these results has to be interpreted with caution, partly due to the wide confidence limits around the fitted line but mostly because, as with all the explanatory variables, we are describing partial effects, once effects of all other variables in the model have been taken into account. Also, most of the survey is in the high seas, comprising mostly oligotrophic areas, with a low representation of effort in coastal areas which leads to a highly heterogeneous distribution of records within the range of chlorophyll values. In previous studies that reported positive relationships between chlorophyll and common dolphin presence, survey effort was mostly concentrated in coastal areas, thus providing a wider range of chlorophyll values, making this a good proxy for productivity (Cañadas & Hammond, 2008; Halicka, 2016; Moura *et al.*, 2012; Tobeña *et al.*, 2016). However, in this study, a wide range of depth values was sampled while the surveys passed through mainly oligotrophic waters which resulted in a small range of chlorophyll values sampled, hence depth being a better proxy for areas of upwelling (i.e. more productive areas). Moreover, timings of the chlorophyll blooms vary across the area and common dolphin distribution may not be affected by production at certain times of the year, or in certain areas where other factors are more important. Hence, although the inclusion of chlorophyll concentration improves the overall model result, it is not very useful for the ecological interpretation of the distribution when working over such a wide area. To test the effect of chlorophyll, models would probably perform better when working in narrower areas and with a more homogeneous effort across the range of available chlorophyll values.

Maps of the predicted probabilities along the routes illustrate the model results, highlighting the areas where sighting probabilities reach the highest values: coastal continental Portugal and the Azores archipelago, with slightly lower probabilities in Madeira and in the open-ocean areas close to continental Portugal and in front of Nouadhibou port in Mauritania.

This study shows that common dolphins have core areas of occurrence, thus supporting the idea that the species is more of an ecological specialist than a generalist (Marçalo *et al.*, 2018; Moura *et al.*, 2012). The explanatory power of the models developed was relatively low (under 25%) and, in fact, we have to be realistic about how much we can expect a model to explain about the distribution of a highly mobile species in such a wide area. Moreover, we are grouping animals that are potentially using the area for different purposes (e.g. foraging or travelling). Also, we have to be aware that cetaceans spend a great amount of time underwater so that, with visual observational data, we are only getting a sample of their occurrence. Finally, we do not have a complete knowledge about all the environmental variables that may influence distribution and we cannot assume that cetaceans occurring in the area have perfect knowledge about prey distribution and its variation across seasons and years, so models based on resource selection functions will only tell us where animals are more likely to be, based on an incomplete knowledge of all the predictors involved. Nevertheless, all models performed considerably better than a random model (AUC > 0.5) and provide new information on common dolphin preferences in the area between the months of July and October, especially in the high seas region. Spatial and temporal predictors had a slightly stronger influence than environmental variables on common dolphin distribution. In this wide study area, with surveys occurring over five years and with heterogeneous effort, it is likely that the spatial pattern and the seasonality of common occurrence are linked to different habitat characteristics, also reflecting the effects of several environmental variables. However further work would be needed to determine which environmental variables are involved. Hence, in this context, the models, combined with the spatial and temporal distribution of occurrences, are more successful in identifying important areas of conservation than explaining the ecological rationale for the common dolphin distribution.

This study has several limitations, mostly related with effort heterogeneity, both temporally and spatially: surveys evidently depend on the company's schedule and the surveys along the three routes began in different years, with the Madeira route starting first (2012), therefore having a higher survey effort than the other two transects. Such differences in effort along line-transects are an almost unavoidable disadvantage of using OPOs (Correia *et al.*, 2015; Kiszka *et al.*, 2007; MacLeod *et al.*, 2008; Moura *et al.*, 2012).

Nonetheless, this work shows that the use of OPOs to systematically monitor cetaceans provides important data to fill data gaps in space and time, especially in areas that are logistically challenging for dedicated surveys and where baseline knowledge is needed, i.e. the high seas. It constitutes an important contribution to the knowledge of common dolphin distribution in the ENA, with records in poorly surveyed areas and insights in habitat preferences based on a 5-year dataset of systematic surveys and a great amount of effort. However, more surveys are still needed to fill knowledge gaps, mainly in relation to seasonal variation, as the results presented here mainly reflect temporal variation from July and October, failing to provide a year-round distribution of common dolphins in the area.

6.6 Acknowledgements

We thank the contribution and dedication of all the volunteers during the monitoring campaigns. We are extremely grateful to TRANSINSULAR, the cargo ship company that provided all the logistic support, and to the ships' crews for their hospitality. We thank the reviewers for their comments and suggestions.

6.7 References

Alves F., Quérouil S., Dinis A., Nicolau C., Ribeiro C., Freitas L., *et al.* (2013). Population structure of short-finned pilot whales in the oceanic archipelago of Madeira based on photo-identification and genetic analyses: implications for conservation. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 23: 758-776.

Alves F., Ferreira R., Fernandes M., Halicka Z., Dias L. and Dinis A. (2018). Analysis of occurrence patterns and biological factors of cetaceans based on long-term and fine-scale data from platforms of opportunity: Madeira Island as a case study. *Marine Ecology*, 39: e12499.

Alves F., Alessandrini A., Servidio A., Mendonça A.S., Hartman K.L., Prieto R., *et al.* (2019). Complex biogeographical patterns support an ecological connectivity network of a large marine predator in the north-east Atlantic. *Diversity and Distributions,* 25: 269-284.

AVISO (2017). Sea-surface Height Products. AVISO+ Satellite Altimetry Data, Toulouse, France. AVISO. Available at https://www. aviso.altimetry.fr/en/ data/products/sea-surface-height-products/global/msla-h.html (accessed September 2017).

Baird R.W., Schorr G.S., Webster D.L., Mahaffy S.D., McSweeney D.J., Hanson M.B. and Andrews R.D. (2011). Open-ocean movements of a satellitetagged Blainville's beaked whale (*Mesoplodon densirostris*): evidence for an offshore population in Hawaii? *Aquatic Mammals*, 37: 506-511.

Beck J.R. and Shultz E.K. (1986). The use of relative operating characteristic (ROC) curves in test performance evaluation. *Archives of Pathology & Laboratory Medicine*, 110: 13-20.

Berrow S., Suárez P.L., Jann B., O'Brien J., Ryan C., Varela J. and Hazevoet C.J. (2015). Recent and noteworthy records of Cetacea from the Cape Verde Islands. *Zoologia Caboverdiana*, 5: 111-115.

Bowman Bishaw Gorham. *Chinguetti Development Project (2003) Seabird and Cetacean Surveys in the Vicinity of the Chinguetti Oil Field, Offshore Mauritania*. Woodside Australian Energy, 66 pp.

Burnham K.P. and Anderson D.R. (2002). *Model Selection and Multimodel Inference: a Practical Information Theoretic Approach*. Springer Verlag: New York, USA.

Caldeira R.M.A., Groom S., Miller P., Pilgrim D. and Nezlin N.P. (2002). Sea-surface signatures of the Island mass effect phenomena around Madeira Island, Northeast Atlantic. *Remote Sensing of Environment,* 80, 336-360.

Camphuysen C.J., van Spanje T.M. and Verdaat H. (2012). *Ship Based Seabird and Marine Mammal Surveys off Mauritania, Nov–Dec 2012 – Cruise Report*. Mauritanian Institute for Oceanographic Research and Fisheries – IMROP, 73 pp.

Cañadas A. and Hammond P.S. (2008). Abundance and habitat preferences of the short-beaked common dolphin *Delphinus delphis* in the southwestern Mediterranean: implications for conservation*. Endangered Species Research,* 4: 309–331.

Cañadas A. and Vázquez J.A. (2017). Common dolphins in the Alboran Sea: facing a reduction in their suitable habitat due to an increase in sea surface temperature. *Deep-Sea Research. Part II, Topical Studies in Oceanography*, 141: 306-318.

Carrillo M., Pérez-Vallazza C.P. and Alvarez-Vázquez R. (2010). Cetacean diversity and distribution off Tenerife (Canary Islands). *Marine Biodiversity Records,* 3: 1-9.

Correia A.M., Tepsich P., Rosso M., Caldeira R. and Sousa Pinto I. (2015). Cetacean occurrence and spatial distribution: habitat modeling for offshore waters in the Portuguese EEZ (NE Atlantic). *Journal of Marine Systems,* 143: 73-85.

Davis R.W., Ortega-Ortiz J.G., Ribic C.A., Evans W.E., Biggs D.C., Ressler P.H., *et al.* (2002) Cetacean habitat in the northern oceanic Gulf of Mexico. *Deep-Sea Research Part I: Oceanographic Research Papers*, 49: 121-142.

Dinis A., Alves F., Nicolau C., Ribeiro C., Kaufmann M., Cañadas A. and Freitas L. (2016). Bottlenose dolphin *Tursiops truncatus* group dynamics, site fidelity, residency and movement patterns in the Madeira Archipelago (North-East Atlantic). *African Journal of Marine Science,* 38: 151-160.

Dinis A., Marques R., Dias L., Sousa D., Gomes C., Abreu N. and Alves F. (2017). Site fidelity of Blainville's beaked whale (*Mesoplodon densirostris*) off Madeira Island (Northeast Atlantic). *Aquatic Mammals,* 43: 387-390.

Djiba A., Bamy I.L., Bilal A.S.O. and Van Waerebeek K. (2015). Biodiversity of cetaceans in coastal waters of northwest Africa: new insights through platform-of-opportunity visual surveying in 2011–2013. *IOC Technical Series*, 115: 283-297.

Elith J. and 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.

ESRI (2016) ArcGIS Desktop: Release 10.5. Redlands, CA: Environmental Systems Research Institute.

Fernandez M., Yesson C., Gannier A., Miller P.I. and Azevedo J.M.N. (2017). The importance of temporal resolution for niche modelling in dynamic marine environments. *Journal of Biogeography*, 44: 2816-2827.

Fernandez M., Yesson C., Gannier A., Miller P.I. and Azevedo J.M.N. (2018). A matter of timing: how temporal scale selection influences cetacean ecological niche modelling. *Marine Ecology Progress Series,* 595: 217:231.

Frederiksen M., Edwards M., Richardson A.J., Halliday N.C. and 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.

GEBCO (2017) Gridded Bathymetry Data. General Bathymetric Chart of the Oceans, Monaco. Available at http://www.gebco.net/data_and_products/gridded_bathymetry_data/ (accessed September2017).

Goetz S, Read FL, Ferreira M, Portela JM, Santos MB, Vingada J, Siebert U, Marçalo A, Santos J, Araújo H, Monteiro S, Caldas M, Riera M, Pierce GJ (2015) Cetacean occurrence, habitat preferences and potential for cetacean–fishery interactions in Iberian Atlantic waters: results from cooperative research involving local stakeholders. Aquatic Conservation: Marine and Freshwater Ecosystems 25, 138–154.

González L., Pierce G.J., Autret E. and Torres-Palenzuela J.M. (2018). Multi-scale habitat preference analyses for Azorean blue whales. *PLoS ONE*, 13: e0201786.

Grémillet D., Lewis S., Drapeau L., Lingen C.D.V.D., Huggett J.A., Coetzee J.C., *et al.* (2008). Spatial match–mismatch in the Benguela upwelling zone: should we expect chlorophyll and sea-surface temperature to predict marine predator distributions? *Journal of Applied Ecology,* 45: 610-621.

Halicka Z. (2016). Temporal Distribution of the Short-Beaked Common Dolphin (*Delphinus delphis*) in the South of Madeira Island (Portugal) and Relationship with Oceanographic Variables. Master thesis (MSc) in Marine Biology. Universidade do Algarve, Portugal. 52 pp.

Hammond P.S., Macleod K., Berggren P., Borchers D.L., Burt L., Cañadas A., *et al.* (2013) Cetacean abundance and distribution in European Atlantic shelf waters to inform conservation and management. *Biological Conservation,* 64, 107-122.

Hazevoet C.J. and Wenzel F.W. (2000). Whales and dolphins (Mammalia, Cetacea) of the Cape Verde Islands with special reference to the humpback whale *Megaptera novaeangliae* (Borowski,1781). *Contributionsto Zoology,* 69: 197-211.

Hazevoet C.J., Monteiro V., López P., Varo N., Torda G., Berrow S. and Gravanita B. (2010). Recent data on whales and dolphins (Mammalia: Cetacea) from the Cape Verde Islands, including records of four taxa new to the archipelago. *Zoologia Caboverdiana*, 1: 75-99.

Jefferson T.A., Fertl D., Bolaños-Jiménez J.and Zerbini A.N. (2009). Distribution of common dolphins (*Delphinus* spp.) in the western Atlantic Ocean: a critical re-examination. *Marine Biology,* 156: 1109-1124.

Jungblut S., Nachtsheim D.A., Boos K. and Joiris C.R. (2017). Biogeography of top predators – seabirds and marine mammals – along four latitudinal transects in the Atlantic Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography*, 141: 59-73.

Kaschner K., Tittensor D.P., Ready J., Gerrodette T. and Worm B. (2010). Current and future patterns of global marine mammal biodiversity. *PLoS ONE,* 6: e19653.

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

Kiszka J., Macleod K., Van Canneyt O., Walker D. and Ridoux V. (2007). Distribution, encounter rates, and habitat characteristics of toothed cetaceans in the Bay of Biscay and adjacent waters from platform-of-opportunity data. *ICES Journal of Marine Science*, 64: 1033-1043.

MacLeod C.D., Mandleberg L., Schweder C., Bannon S.M. and Pierce G.J. (2008) A comparison of approaches for modelling the occurrence of marine animals. *Hydrobiologia,* 612: 21-32.

Marçalo A., Nicolau L., Giménez J., Ferreira M., Santos J., Araújo H., *et al.* (2018). Feeding ecology of the common dolphin (*Delphinus delphis*) in Western Iberian waters: has the decline in sardine (Sardina pilchardus) affected dolphin diet? *Marine Biology,* 165: 44.

Marubini F., Gimona A., Evans P.G.H., Wright P.J. and Pierce G.J. (2009). Habitat preferences and interannual variability in occurrence of the harbour porpoise *Phocoena phocoena* off northwest Scotland. *Marine Ecology Progress Series,* 381: 297-310.

Mason E. (2009). High-resolution Modelling of the Canary Basin Oceanic Circulation. Doctorate thesis (PhD) in Oceanography, Universidad de Las Palmas de Gran Canaria, Las Palmas de Gran Canaria. 245 pp.

Meynier L., Pusineri C., Spitz J., Santos M.B., Pierce G.J. and Ridoux V. (2008). Intraspecific dietary variation in the short-beaked common dolphin Delphinus delphis in the Bay of Biscay: importance of fat fish. *Marine Ecolology Progress Series*, 354: 277-287.

Moura A.E., Sillero N. and Rodrigues A. (2012). Common dolphin (*Delphinus delphis*) habitat preferences using data from two platforms of opportunity. *Acta Oecologica,* 38: 24-32.

NASA (2017) Ocean Color Data — MODISA. NASA. Available at https://oceandata.sci.gsfc.nasa.gov/ (accessed September 2017).

Pearce J.L. and Boyce M.S. (2006). Modelling distribution and abundance with presence only data. *Journal of Applied Ecology*, 43: 405-412.

Peltier H., Authier M., Deaville R., Dabin W., Jepson P.D., van Canneyt O., *et al.* (2016). Small cetacean bycatch as estimated from stranding schemes: the common dolphin case in the northeast Atlantic. *Environmental Science & Policy*, 63: 7-18.

Pierce G.J., Caldas M., Cedeira J., Santos M.B., Llavona A., Covelo P., *et al.* (2010) Trends in cetacean sightings along the Galician coast, north-western Spain, 2003–2007, and inferences about cetacean habitat preferences. *Journal of the Marine Biological Association of the United Kingdom*, 90: 1547-1560.

Piroddi C, Bearzi G, Gonzalvo J and Christensen V (2011) From common to rare: the case of the Mediterranean common dolphin. Biological Conservation 144, 2490–2498.

Pirotta E., Matthiopoulos J., MacKenzie M., Scott-Hayward L. and Rendell L. (2011). Modelling sperm whale habitat preference: a novel approach combining transect and follow data. *Marine Ecology Progress Series*, 436: 257-272.

Quian S.S. (2009). *Environmental and Ecological Statistics with R*. Toronto: Chapman & Hall.

R Development Core Team (2012). R: A Language and Environment for Statistical Computing. Vienna: R Foundation. www.R-project.org/.

Redfern J.V., Ferguson M.C., Becker E.A., Hyrenbach K.D., Good C., Barlow J., *et al.* (2006) Techniques for cetacean-habitat modeling. *Marine Ecology Progress Series*, 310: 271-295.

Roberts J.J., Best B.D., Dunn D.C., Treml E.A. and Halpin P.N. (2010). Marine geospatial ecology tools: an integrated framework for ecological geoprocessing with ArcGIS, Python, R, MATLAB, and C++. *Environmental Modelling & Software*, 25: 1197-1207.

Robinson I.S. (2010). *Discovering the Ocean from Space – The Unique Applications of Satellite Oceanography*. Bath: Springer.

Sala I., Caldeira R.M.A., Estrada-Allis S.N., Froufe E. and Couvelard X. (2013). Lagrangian transport pathways in the northeast Atlantic and their environmental impact. *Limnology & Oceanography: Fluids & Environments*, 3: 40:60.

Santos M.B. and Pierce G.J. (2015). Marine mammals and good environmental status: science, policy and society; challenges and opportunities. *Hydrobiologia*, 750: 13-41.

Santos M.B., German I., Correia D., Read F.L., Cedeira J.M., Caldas M., *et al.* (2013) Long-term variation in common dolphin diet in relation to prey abundance. *Marine Ecology Progress Series*, 481: 249-268.

Santos M.B., Saavedra C. and Pierce G.J. (2014). Quantifying the predation on sardine and hake by cetaceans in the Atlantic waters of the Iberian Peninsula. *Deep-Sea Research. Part II, Topical Studies in Oceanography*, 106: 232-244.

Schlacher T.A., Rowden A.A., Dower J.D. and Consalvey M. (2010) Seamount science scales undersea mountains: new research and outlook. *Marine Ecology,* 31: 1-13.

Silva M.A., Prieto R., Cascão I., Seabra M.I., Machete M., Baumgartner M.F. and Santos R.S. (2014). Spatial and temporal distribution of cetaceans in the mid-Atlantic waters around the Azores. *Marine Biology Research*, 10: 123-137.

Smith J.M., Pierce G.J., Zuur A.F., Martins H., Martins M.C., Porteiro F. and Rocha F. (2011). Patterns of investment in reproductive and somatic tissues in the loliginid squid *Loligo forbesii and Loligo vulgaris* in Iberian and Azorean waters. *Hydrobiologia*, 670: 201-221.

Spyrakos E., Santos-Diniz T.C., Martinez-Iglesias G., Torres-Palenzuela J.M. and Pierce G.J. (2011). Spatio temporal patterns of marine mammal distribution in coastal waters of Galicia, NW Spain. *Hydrobiologia*, 670: 87-109.

Stockin K.A., Pierce G.J., Binedell V., Wiseman N. and Orams M.B. (2008). Factors affecting the occurrence and demographics of common dolphins (*Delphinus* sp.) in the Hauraki Gulf, New Zealand. *Aquatic Mammals*, 34: 200-211.

Tobeña M., Prieto R., Machete M. and Silva M.A. (2016). Modeling the potential distribution and richness of cetaceans in the Azores from Fisheries Observer Program data. *Frontiers in Marine Science*, 3: 202.

Viddi F.A., Hucke-Gaete R., Torres-Florez J.P. and Ribeiro S. (2010). Spatial and seasonal variability in cetacean distribution in the fjords of northern Patagonia, Chile. *ICES Journal of Marine Science*, 67: 959-970.

Weir C.R. (2010). A review of cetacean occurrence in West African waters from the Gulf of Guinea to Angola. *Mammal Review*, 40: 2-39.

Weir C.R. and Pierce G.J. (2013). A review of the human activities impacting cetaceans in the eastern tropical Atlantic. *Mammal Review*, 43: 258-274.

Weir C.R., Coles P., Ferguson A., May D., Baines M., Figueirdo I. *et al.* (2014). Clymene dolphins (*Stenella clymene*) in the eastern tropical Atlantic: distribution, group size, and pigmentation pattern. *Journal of Mammalogy* 95: 1289-1298.

Zuur A.F., Ieno E. and Elphick C. (2010). A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution,* 1: 3-14.

CHAPTER VII.

Paper prepared for submission

7 Predicting cetacean distributions in the eastern North Atlantic to support marine management

Ana M. Correia[1,2], Diana S. Guedes[3], Ágatha Gil[1,2], Raul Valente[1,2], Massimiliano Rosso[1,4], Isabel Sousa-Pinto[1,2], Neftalí Sillero[3], Graham J. Pierce[5,6,7]

[1]Interdisciplinary Centre of Marine and Environmental Research (CIIMAR). 4450-208 Matosinhos, Portugal.
[2]Department of Biology, Faculty of Sciences, University of Porto (FCUP). 4169-007 Porto, Portugal.
[3]CICGE, Centro de Investigação em Ciências Geo-Espaciais, Faculty of Sciences, University of Porto (FCUP). 4430-146 Vila Nova de Gaia, Portugal.
[4]CIMA Research Foundation. 17100 Savona, Italy.
[5]Instituto de Investigacións Mariñas (CSIC). 36208 Vigo, Pontevedra, Spain.
[6]Oceanlab, University of Aberdeen. AB41 6AA, UK.
[7]CESAM and Department of Biology, University of Aveiro. 3810-193 Aveiro, Portugal.

7.1 Abstract

Using six years of cetacean occurrence data collected along cargo ship routes between Iberian Peninsula, northwestern (NW) African coasts and the Macaronesian islands, we developed ecological niche models (ENMs) to predict the distribution of the eight most sighted taxa in the area (seven species and one genus), representative of all cetacean major groups: small delphinoids, toothed and baleen whales. We used two different modelling approaches: presence/pseudo-absence with Generalized Additive Models (GAMs), including sightings collected during on effort sampling; and presence/background with Maximum entropy models (Maxent), including all sightings. Variables used for model fitting included topographic (depth, slope, distance to coast, distance to seamounts), oceanographic (chlorophyll, temperature and sea level anomalies), detectability (sea-state, wind-state and visibility), geographic (latitude), and seasonal (day of the year) features.

The best model for the most frequently sighted species, *Delphinus delphis*, explained around 20% of deviance by GAM. Predictions pointed for a more northern and coastal distribution with occurrence peaks in continental (Portugal and NW Africa) and island shores (excluding Cape Verde). Distributions predicted by Maxent varied among the modelled species, showing narrow or broad ranges (e.g., *Ziphius cavirostris* VS *Stenella frontalis*), coastal or oceanic (e.g., *Tursiops truncatus* VS *Globicephala* sp.), and northern or southern species (e.g., *Delphinus delphis* VS *Physeter macrocephalus*). Both modelling techniques were consistent and complementary, allowing an in-depth analysis on the ecological significance of each predictor and on the spatial distribution of species across the study area.

Our results are an important contribution to the knowledge of cetacean distribution at basin-scale as baseline data for the European conservation agenda, especially in less-surveyed areas in open-ocean. Our models allow the definition of priority areas, monitoring plans, and conservation measures within the EEZs, essential to comply with the Marine Strategy Framework Directive requirements.

Keywords: Ecological niche modelling, GAMs, Maxent, basin-scale modelling.

7.2 Introduction

One of the main research questions for cetacean conservation is related to managing data deficiency. Over 35% of cetacean species are data deficient (www.iucnredlist.org). Cetacean conservation is consequently hindered as “data deficient” is often treated as “no concern” or leads to delayed management actions (Parsons *et al.*, 2016). How can research address data gaps and provide useful data for decision makers? How to get comprehensive data on wide-ranging species, such as cetaceans, that cross long-distances without any physical barriers? How to sample remote areas like open-ocean waters where long-term monitoring programs are financially and logistically challenging? The solution may lie on the use of observation platforms of opportunity (OPOs), remote sensing data and ecological niche modelling.

Recently, the use of OPOs to collect long-term data on cetacean occurrence has increased considerably (Alves *et al.*, 2018a; Arcangeli *et al.*, 2016; Tobeña *et al.*, 2016). The sampling protocols and the techniques used in the data processing and analysis have been refined: data collected from OPOs is now frequently used to assess cetacean occurrence and relative abundances, as well as to conduct ecological niche modelling (Correia *et al.*, 2015; Kiszka *et al.*, 2007; Moura *et al.*, 2012). Modelling techniques have been widely used to forecast cetacean distribution and understand its relationship with habitat characteristics (Prieto *et al.*, 2017; Storrie *et al.*, 2018; Tepsich *et al.*, 2014). Ecological niche models (ENMs) are an optimal solution for data-poor areas, as model results can be projected beyond the surveyed area (Redfern *et al.*, 2017). Results have been successfully applied in the definition of monitoring plans, management strategies and creation of Marine Protected Areas (MPAs) (Cañadas *et al.*, 2005; García-Barón *et al.*, 2019; Passadore *et al.*, 2018).

Presence-only models without survey effort data may provide reliable information on cetacean occurrence ranges (Do Amaral *et al.*, 2015; Friedlaender *et al.*, 2011; Thorne *et al.*,2012). However, effort-based data can provide better insights into species habitat characteristics as they account for surveyed habitat and allow the definition of pseudo-absences close to true absences (Brotons *et al*., 2004; Praca *et al.*, 2009; Tepsich *et al*., 2014). Besides selecting the best model technique, it is also important to decide on the appropriate explanatory variables and spatiotemporal scales to include in the models. So far, the most frequently used predictors in ecological niche modelling are static variables (such as topography and derived variables), as they are easier to use for management purposes (e.g., definition of MPAs). Nonetheless, oceanographic processes play a fundamental role in determining the distribution of cetaceans, for example because they can indicate local abundance of prey species. Hence, a combination of static and dynamic variables should be considered when modelling cetacean distribution, as well as for management efforts (Breen *et al.*, 2017; Pennino *et al.*, 2017; Tobeña *et al.*, 2016). In fact, the definition of dynamic MPAs has been previously suggested (Hooker *et al*., 2011). Another fundamental factor is the spatial and/or temporal scales of each variable. The scales can influence highly model results and application. Therefore, multi-scale models and/or the testing of several scales is recommended (Fernandéz *et al*., 2018). Overall, the best model approach and methodology must be selected given the data available, sampled area and the aims of the models (Redfern *et al.*, 2006).

In the eastern North Atlantic, the area encompassing the Iberian and northwestern (NW) African coasts and the Macaronesia, 36 species of cetaceans have been described, with the eight most frequently sighted species being representative of the main guilds of cetaceans: small dolphins (*Tursiops truncatus*, *Delphinus delphis*, *Stenella coeruleoalba* and *Stenella frontalis*), large dolphins (*Globicephala* sp.), beaked whales (*Ziphius cavirostris*), sperm whales (*Physeter macrocephalus*) and baleen whales (*Balaenoptera acutorostrata*) (Correia *et al.*, unpublished data). This is an area with a wide latitudinal and longitudinal range encompassing a great habitat variability (Caldeira *et al*., 2002; Mason, 2009; Sala *et al*., 2013). The species profiles vary among sub-regions (Correia *et al.*, unpublished data), but cetaceans move and migrate across the entire area (Alves *et al.*, 2018b; Valente *et al.*, 2019). Therefore, to fully understand and protect cetacean species in this area, distribution patterns need to be analysed at the basin-scale. However, similarly to other areas in the globe, the eastern North Atlantic is a data-poor area without oceanic data on cetacean occurrence (Correia *et al.*, 2015; Hammond *et al.*, 2013; Jungblut *et al.*, 2017). Besides *P. macrocephalus* and *B. acutorostrata* (“vulnerable” and “least concern”, respectively), the remaining most frequently sighted species are “data deficient” in European waters, and the status of both *S. frontalis* and *Globicephala* sp. are not assessed at the European level (www.iucnredlist.org). This lack of cetacean occurrence data in the area has already hindered cetacean conservation and management (Santos & Pierce, 2015; Directive, 2008/56/EC of the European Parliament and of the Council of 17 June 2008). Therefore, ecological niche modelling may be the most appropriate tool to address this issue of data deficiency.

Within this study, we aimed to relate habitat characteristics to the distribution of the eight most frequently sighted species within the eastern North Atlantic, by using ecological niche models at basin-scale, with effort-based data collected since 2012 by OPOs along long-distance line-transects (CETUS Project, www.cetusproject.com). As a consequence, the outputs of the analysis will also reflect the value of data collected from OPOs to this kind of studies. We used two different modelling techniques: a presence/pseudo-absence approach using Generalized additive models (GAMs) to analyse cetacean-habitat relationships; and a presence/background approach using Maximum entropy models (Maxent), to map cetacean species distributions over the entire study area. Besides analysing the overall distribution patterns, we discussed occurrence and distribution within each Exclusive Economic Zone (EEZs) for added-value to marine conservation and management.

7.3 Material and Methods

7.3.1 Data collection

Cetacean occurrence data were collected within the CETUS Project (www.cetusproject.com), a cetacean monitoring program in the eastern North Atlantic, running since 2012. Through a collaboration with TRANSINSULAR, a Portuguese company for maritime transport, cargo ships are used as OPOs to collect data along commercial routes between continental Portugal, Macaronesian archipelagos and NW Africa (Figure 1). Every year, each ship receives a team of two marine mammal observers (MMOs) for cetacean surveys following a standard sampling protocol, from sunrise to sunset (Correia *et al.*, 2015; Hammond *et al.*, 2013; Tepsich *et al.*, 2014). MMOs stand in the wings of the navigation bridge (at a height of between 13.5 and 16 metres above sea level, considering maximum draught and speed, and depending on the ship), occasionally monitoring from inside of the ship when weather is uncomfortable (i.e., strong winds or moderate rain) but still suitable to survey. Each observer stands on one side of the vessel and they switch every 60 minutes (approximately) to avoid fatigue and data biases. Moreover, in turns, both take one-hour breaks for meals and two optional rests of up to 40 minutes (one in the morning and another during the afternoon).

Monitoring is performed from the front of the vessel with a field of view of approximately 180º, each MMO covering 90º; at mealtimes and resting periods, the lone MMO covers the entire 180º range from one of the sides. Observers look for cetacean presence with the naked eye, performing occasional scans with binoculars (with a compass and a distance scale with seven or eight reticles, 7 × 50 mm). The route of the ship during sampling and the positions marked by the MMOs are recorded using a tablet with an inbuilt GPS and running the application MyTracks (https://my-tracks.pt.aptoide.com), which registers date and time, speed and direction of the route. Weather conditions are assessed at the beginning and end of each survey leg (a continuous period of sampling) and every time there are significant changes in the conditions. The following variables are recorded: sea state (using the Douglas scale), wind speed (using the Beaufort scale), visibility (on a categorical scale of values from 1 – 10 covering visibility ranges from 0 m to more than 50000 m, estimated based on the definition of the horizon line and reference points at a known range, e.g., ships with an AIS system) and the occurrence of rain. For the marine traffic, small and big vessels (less than and over 20 m in length), detected with or without binoculars, at a 360º range, are registered at the beginning and end of each survey, hourly and at every sighting. Sampling effort stops whenever MMOs cannot be in the survey stand (e.g. during safety drills, manoeuvres) or when weather conditions are unfavourable for cetacean monitoring: Beaufort or Douglas values >4, visibility <1 km or heavy rain. Any data collected until effort resumes are considered opportunistic (off-effort).

Whenever a cetacean species is sighted, both observers gather on the side of the boat where the animals where spotted and mark the end of an on-effort transect in order to correctly collect the data on the occurrence. After registering the sighting, each MMO returns to its side of the vessel, and a new on-effort transect starts. Species identification is attempted to the species level, however, the taxonomy registered is always at the level to which the MMOs are confident of their identification. For group size measures, the observers provide an estimate of the minimum, maximum and assumed (best estimate) number of individuals in a sighting. Moreover, whenever possible, information on the heading of the group and its behaviour towards the ship (i.e. approaching, indifferent or avoiding) is also collected. The occurrence of pelagic megafauna other than cetaceans is collected opportunistically, with observers recording only taxonomic information and number of individuals.

For ecological niche modelling, in addition to weather conditions and spatiotemporal variables, we derived habitat variables (static and dynamic) from satellite data at several temporal and spatial scales (see Supplemental file 1). Slope was derived from bathymetry data. For distance to seamounts, we delimited topographic features classified as seamounts, banks, hills, ridges and rises in GEBCO (GEBCO 2017), using contour lines created every 50 m and defining a polygon from the outermost closed contour line around the geographic location of the top of the features. Then, we calculated the distance from the base of the seamounts and from the coastline (distance to coast) to the sightings. Both slope and distances were computed using ArcGIS 10.5 (ESRI 2016). Chlorophyll-a and sea-surface temperature are ocean products derived from the satellite MODIS – Aqua data from NASA (NASA, 2017). The algorithms return the near-surface concentration of chlorophyll-a (from in situ remote sensing reflectance) and temperature (from measured radiances). We extracted both variables at two different spatial (four and nine kilometres) and temporal (eight-day and monthly) scales. For altimetry, the mean sea level anomalies were obtained from Ssalto/Duacs multimission altimeter products provided by AVISO (AVISO, 2017). The sea level anomalies are sea-surface heights computed with respect to a twenty-year mean profile (1993-2012). We used delayed products, available around two months after collection, after re-analysis and re-processing (AVISO, 2017). For this variable, eight-day and monthly resolutions were computed by averaging daily products.

7.3.2 Ecological niche models

We applied modelling techniques to the eight most frequently sighted species: common dolphin (*D. delphis*), spotted dolphin (*S. frontalis*), striped dolphin (*S. coeruleoalba*), bottlenose dolphin (*T. truncatus*), Cuvier's beaked whale (*Z. cavirostris*), pilot whale (*Globicephala* sp.), sperm whale (*P. macrocephalus*), and minke whale (*B. acutorostrata*). We used GAMs to describe cetacean habitat characteristics using only cetacean occurrence data collected on-effort and all variables. We used Maxent models to predict and map suitable and unsuitable habitat across the entire study area using all cetacean occurrences (on and off-effort) and a selection of variables. Variables were chosen to reflect spatiotemporal occurrence, detectability factors and environmental factors (Supplemental file 1).

Generalized additive models (GAMs)

For GAMs, we chose a presence/pseudo-absence approach based on used/available habitat (Pearce & Boyce, 2006; Elith & Leathwick, 2009; Correia *et al.*, 2015), with used (cetacean occurrence) and available habitat (survey route) points combined to generate a binary (1,0) response variable. The set of available points was generated as in Correia *et al.* (2015), through the creation of equidistant points (every 5 km) along all effort tracks. In this way, survey effort is included in the models as areas with higher survey effort provide more points of available habitat. The values of the explanatory variables values were sampled for the set of used and available points. To sample oceanographic variables, we used the pack of tools for ArcGIS, Marine Geospatial Ecology Tools (MGET) (Roberts *et al.*, 2010).

Prior to modelling, we computed Pearson correlations between explanatory variables to allow us to identify and exclude highly collinear explanatory variables from being included in the same model (threshold of R=0.75) (after Marubini *et al.*, 2009). Distance to coast and depth were the only pair of variables that were highly correlated. Both were of interest, hence, we fitted a GAM model with depth as predictor and distance to coast as response variable. Subsequently, both depth and the residuals of this model were used in the habitat models (see Smith *et al.*, 2011). Moreover, we assessed multiple correlation among explanatory variables through the Variance Inflation Factor (VIF, applying a threshold of 3) (Zuur *et al.*, 2010). All VIF values were lower than the threshold, so no additional variables were removed.

We set a binomial distribution for the response variable and a maximum of four splines (k-fold set to 4) to limit the complexity of smoothers describing the effects of explanatory variables. Model fitting started with assembly of an oversaturated model followed by a backward selection (Quian, 2009; Viddi *et al.*, 2010; Correia *et al.*, 2015). However, we undertook forward selection when choosing between the different scales of the oceanographic variables.

Following Correia *et al.* (2015), and to account for varying dolphin group size, we included the best estimate of the number of animals sighted in a group as a weighting parameter in the models. Given the wide range of group size and high uncertainty of the estimations for the species sighted in large groups (common dolphins, spotted dolphins, striped dolphins, bottlenose dolphins and pilot whales), we attributed weights in categories, considering the best estimate count of animals sighted for each observation: a small group – from one to five animals (weight = 1); a medium group – from six to 20 animals (weight = 2); a large group – more than 20 animals (weight = 3). A weight of 1 was set for points of available habitat.

We selected the best models by using the Akaike Information Criterion (AIC) as a measure of goodness of fit, choosing the model with the lowest AIC value at each step of the model fitting process, i.e. comparing otherwise identical models with or without a specific explanatory variable. If the difference in AIC values between two models was less than 2, a Chi-squared test was applied. Whenever differences between AIC values were not statistically significant (based on δAIC >2 or the chi-square test result), we kept the simplest model (following the principle of parsimony, e.g. Burnham & Anderson, 2002). In the case of the forward selection, for the oceanographic variables, whenever models did not differ with statistical significance, we chose the highest resolution (four kilometres over nine kilometres for spatial resolution, and eight-day over monthly for temporal resolution). Finally, we evaluated the models by creating two random subsets of data: fitting and evaluating sets (75% and 25% of the data, respectively). Prediction power of the models was determined using the Area Under the Curve (AUC) of the Receiving Operator Characteristic (ROC) plot (Beck & Shultz, 1986; Liu *et al.*, 2005). Random models have an AUC equal to 0.5; the closer an AUC is to 1, the more discriminant the model is.

Models were developed using the 'mgcv' package in R 3.4.4. (R Core Team, 2018) with R Studio.

Maximum entropy models (Maxent)

We modelled the ecological realised niches (see Sillero, 2011) of the eight species using the Maximum Entropy method implemented in Maxent software (Philips *et al.*, 2006; 2017), a correlative niche algorithm for presence-only and background records. This method distinguishes between suitable and unsuitable habitats (Sillero, 2011). Maxent starts with a uniform probability distribution (gain=0) and alters one weight at a time to maximise the likelihood of the occurrence data set, converging to the optimum probable distribution (Philips *et al.*, 2006; 2017). The output ranges from 0.0 to 1.0, representing the habitat suitability (not the occurrence probability, as presence-absence algorithms do). Maxent generates background sample of points, randomly selected from the whole study area, without any relationship to the presence or absence of the species (Philips *et al.*, 2009; Elith *et al.*, 2011). This selection of background points may strongly affect the model (Fourcade *et al.*, 2014). However, model performance increases if background points are extracted from areas near to species presences (Philips, 2009). For this reason, we clipped the environmental variables with four different buffer sizes (5, 10, 20 and 50 km) around the cetacean presence points, allowing Maxent to select the random background points within the buffer area. We then projected the models to the whole study area. We defined the buffer sizes following visibility conditions (height of the observation deck, visibility range) during at-sea surveys and species ecology (dolphins vs whales) (Fourcade *et al.*, 2014): up to 5 km, most animals are spotted under favourable conditions and jumping dolphins near the ship are sighted even in off-effort conditions; at a 10 km range, whales' blows are seen and some jumping dolphins can still be spotted under favourable weather conditions; 20 km is the most common visibility range during CETUS surveys; and 50 km is the maximum visibility range ever recorded (to which ships were still spotted at the horizon line at a distance of ~ 50 km confirmed in the AIS system of the cargo vessel).

We selected five variables with a Pearson correlation lower than 0.75 (Supplemental file 1): slope, chlorophyll-a, distance to seamounts, sea surface temperature and depth. Distance to coast (correlated with depth) and latitude (correlated with sea surface temperature) were then excluded. We did not include mean sea level anomalies due to its very low spatial resolution. We averaged chlorophyll-a and sea surface temperature variables for the fieldwork period in Macaronesia region using 8-day resolution files with Raster Calculator in QGIS. The spatial resolution of climate variables chosen was 4 km.

We performed Maxent runs with auto features, using 70% of the points as training data and 30% as test data. All duplicated records were eliminated. We built 100 model replicates for each species and gathered the arithmetic mean and the standard deviation for each set of 100 replicate models. We ran Maxent in clog-log format (Philips *et al.*, 2017).

Models performance were analysed with the AUC of the ROC plot (Liu *et al.*, 2005). In addition, we calculated a set of 100 null models for each species, following the methodology by Raes & Ter Steege (2007). For this, we created 100 different datasets with the same number of random points as the species presences following a Poisson distribution. We obtained the AUC values of the ROC plots for each set of 100 null models. Then, we compared the training AUC values between species models and null models using a Kruskal-Wallis test. Null models were calculated in R 3.4.4. (R Core Team, 2018) using 'dismo' package (Hijmans *et al.*, 2017).

The importance of each environmental variable was determined by the average percentage of contribution and permutation importance of each variable to the models through factor analysis: (1) a jackknife analysis of the average AUC using training and test data; and (2) a calculation of the average percentage contribution of each variable to the models. For this purpose, the variables were excluded in turn and a model was created with the remaining variables; then a model was created using each individual variable.

For two species, we analysed predictions within the EEZs in detail. For this analysis, we selected *T. truncatus* as it is an indicator species in the Marine Strategy Framework Directive (Santos & Pierce, 2015) and *Z. cavirostris* since the species is categorized as data deficient in European waters, with few occurrence records previously published, especially in offshore areas (www.iucnredlist.org).

7.4 Results

7.4.1 Sightings and survey-effort

A total of 124428 km of survey effort was distributed along three main routes: continental Portugal to Azores, to Madeira and to Cape Verde (with stopovers in the Canary Islands and West Africa) (Figure 1). We collected 2807 sightings and 1266 were analysed within this study for the eight most frequently sighted species (919 collected on-effort, 347 recorded opportunistically and 11 sightings with two associated species): *D. delphis* (394 sightings, 262 on-effort), *S. frontalis* (226 sightings, 167 on-effort), *S. coeruleoalba* (154 sightings, 119 on-effort), *T. truncatus* (134 sightings, 92 on-effort), *Z. cavirostris* (64 sightings, 51 on-effort), *Globicephala* sp. (59 sightings, 44 on-effort), *P. macrocephalus* (152 sightings, 116 on-effort) and *B. acutorostrata* (92 sightings, 75 on-effort). Since some single sightings included two of the selected species (i.e., species were sighted in association), those are accounted twice above when presenting number of sightings by species (Figure 1).

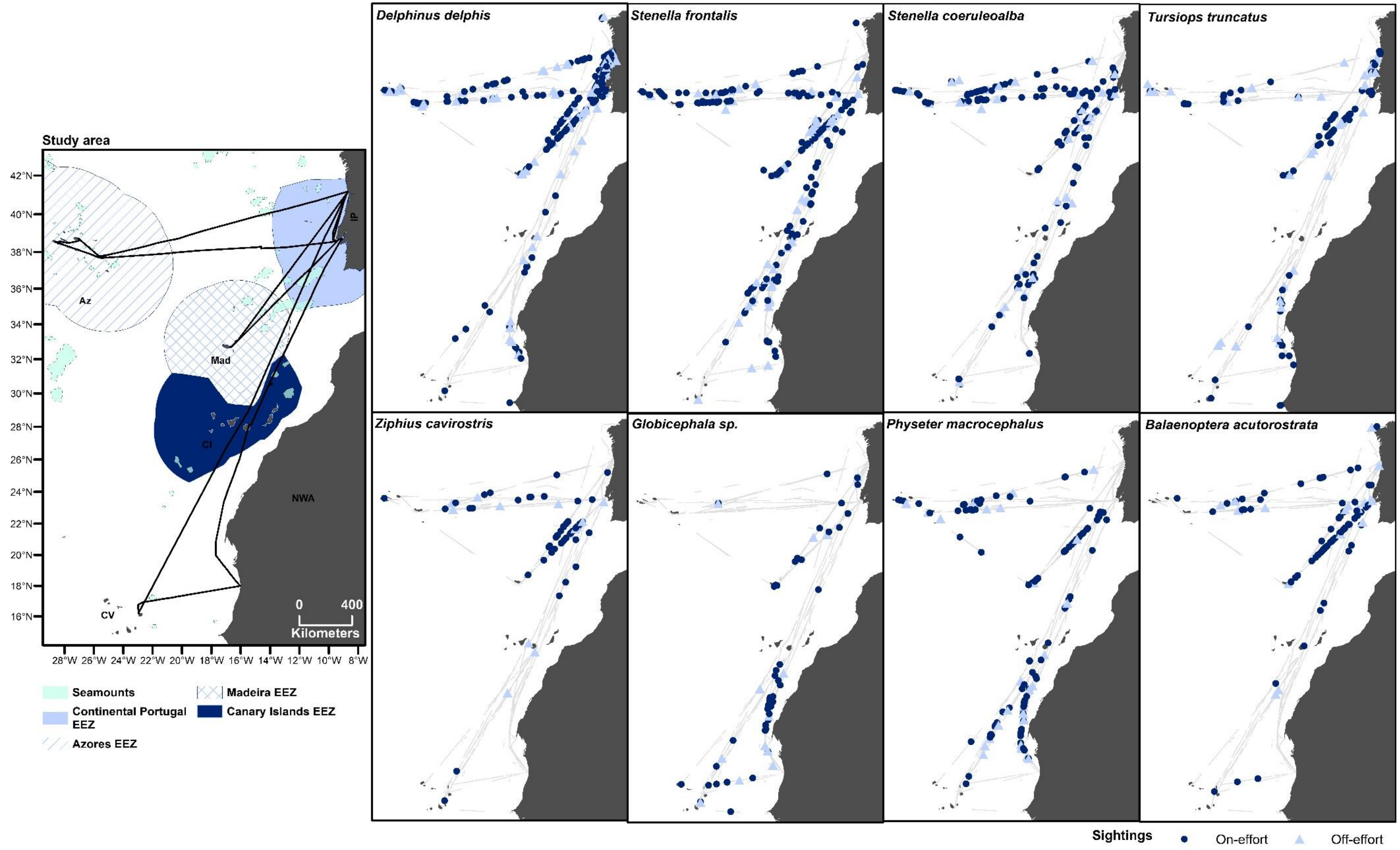


Figure 1. Study area with the surveyed transects and sightings of the eight most frequently sighted species, used in the environmental niche models. The approximate tracks of the routes monitored are presented in black lines in the study area map. On-effort legs are represented by grey lines in the species maps. EEZ – Exclusive Economic Zones; IP – Iberian Peninsula; NWA – Northwest Africa; Az – Azores; Mad – Madeira; CI – Canary Islands; CV – Cape Verde.

7.4.2 Ecological niche models

All eight GAM models included at least one variable from each category (detectability, spatiotemporal and environmental). All models performed better than a random model (AUC > 0.5) and explained deviances ranged from 7.45% (*S. frontalis*) to 26.8% (*Globicephala* sp.) (Table 1). Detailed model results are presented in the supplementary material (Supplemental file 2).

Table 1. Results from the best GAM models developed for the eight most frequently sighted species.

Model	Presences / Pseudo-absences	Deviance explained	AUC (CI 95%)
DD ~ s(day) + s(wind) + s(depth) + s(slope) + s(res) + s(dist_sm) + s(sst_4km_8d) + s(chl_9km_m) + s(msla_8d)	228 / 19570	20.2%	0.81 (0.74-0.87)
SF ~ s(day) + s(lat) + s(sea) + s(depth) + s(slope) + s(dist_sm) + s(sst_4km_8d) + chl_4km_m + s(msla_8d)	150 / 19439	7.45%	0.61 (0.52-0.71)
SC ~ s(day) + s(lat) + wind + s(vis) + slope + s(dist_sm) + s(sst_4km_m) + chl_4km_8d + s(msla_m)	93 / 17697	10.1%	0.70 (0.57-0.83)
TT ~ s(lat) + s(wind) + s(depth) + s(dist_sm) + chl_4km_m + s(msla_m)	79 / 22344	12.1%	0.75 (0.66-0.84)
ZC ~ s(day) + s(lat) + wind + s(depth) + sst_9km_8d + s(msla_m)	48 / 21095	13.9%	0.74 (0.62-0.86)
Gsp. ~ s(lat) + wind + s(depth) + s(sst_4km_m) + s(chl_4km_8d) + s(msla_8d)	30 / 17707	26.8%	0.93 (0.88-0.98)
PM ~ s(lat) + s(wind) + vis + s(res) + s(dist_sm) + s(sst_4km_m) + s(chl_9km_8d) + s(msla_8d)	87 / 18806	17.8%	0.73 (0.63-0.83)
BA ~ s(day) + s(wind) + s(vis) + s(depth) + dist_sm + s(msla_8d)	75 / 23419	7.48%	0.74 (0.63-0.84)

DD – *D. delphis*; SF – *S. frontalis*; SC – S. coeruleoalba; TT – *T. truncatus*; ZC – *Z. cavirostris*; Gsp. – *Globicephala* sp.; PM – *P. macrocephalus*; BA – *B. acutorostrata*; res – residuals from GAM model distance to coast ~ depth; AUC – Area under the curve; CI – Confidence interval. Number of presences decreases from the total number of sightings when chlorophyll-a and sea surface temperature is included in the model, due to non-available measures in some of the sightings points. The abbreviations used for the variables are defined in Supplemental file 1.

The best Maxent models were obtained with the buffer size of 50 km (results not showed for the other buffers). The eight Maxent models had mean training AUC values close to 0.8 and test AUC close to 0.7 (Table 2). AUC values of all species models were significantly higher than null models (Kruskal-Wallis with p-values < 0.001).

Table 2. Results from the Maxent models developed for the eight most frequently sighted species with the 50 km buffer.

Species	Training AUC	Training records	SD training AUC	Test AUC	Test records	SD test AUC	Percent contribution				
							chl	depth	dist.sm	slope	sst
DD	0.850	207	0.0084	0.826	88	0.0226	3.03	**47.81**	12.79	1.99	34.39
SF	0.746	144	0.0119	0.690	61	0.0293	26.52	9.39	10.54	17.72	**35.84**
SC	0.782	101	0.0124	0.719	43	0.0249	12.65	16.14	5.62	5.10	**60.49**
TT	0.819	75	0.0124	0.766	31	0.0423	14.29	**52.20**	20.30	4.47	8.74
ZC	0.832	45	0.0209	0.775	18	0.0567	34.73	16.19	**36.83**	0.54	11.71
Gsp	0.794	40	0.0212	0.707	17	0.0579	12.92	**69.25**	4.49	10.37	2.97
PM	0.800	93	0.0128	0.728	39	0.0298	21.37	27.99	6.91	12.89	30.84
BA	0.817	63	0.0164	0.757	26	0.0389	5.69	21.03	**37.24**	1.45	34.58

DD – *D. delphis*; SF – *S. frontalis*; SC – *S.* coeruleoalba; TT – *T. truncatus*; ZC – *Z. cavirostris*; Gsp. – *Globicephala* sp.; PM – *P. macrocephalus*; BA – *B. acutorostrata*; AUC – Area under the curve; SD – Standard Deviation. The abbreviations used for the variables are defined in Supplemental file 1. In bold are presented the most important variable for each model.

For two selected species, the GAM predicted splines are presented in Figure 2 and 3 (for the remaining species, see Supplemental file 3) with Maxent predicted distribution models represented in Figure 4.

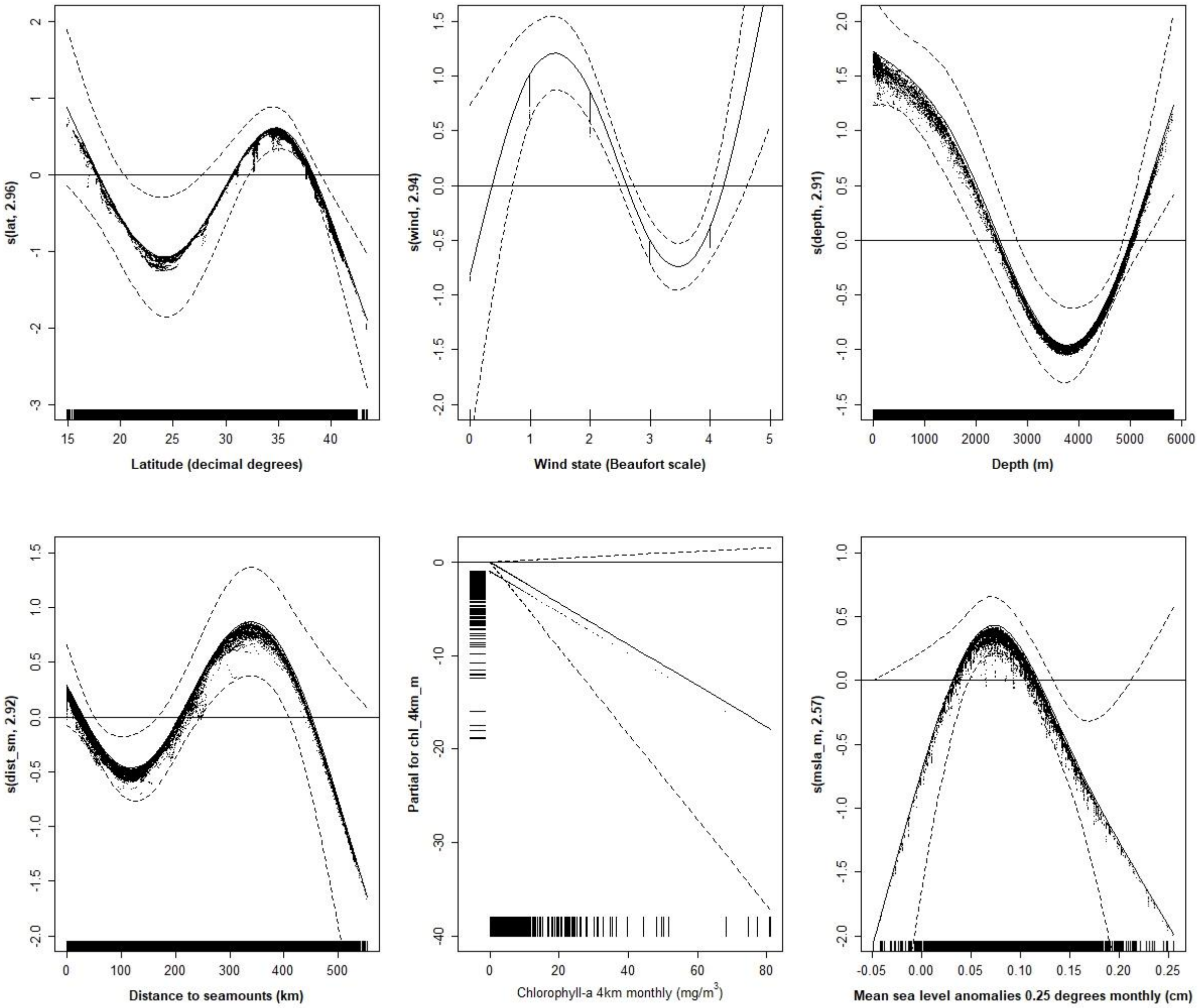


Figure 2. GAM predicted splines of the response variable species presence as a function of the explanatory variables for the environmental model produced for common bottlenose dolphin (*T. truncatus*). The degrees of freedom are in brackets on the y-axis. Tick marks above the x-axis indicate the distribution of observations. Dashed lines delimit the 95% confidence intervals of the spline functions and dots on the graph area represent the residuals. The abbreviations used for the variables are defined in supplemental file 1.

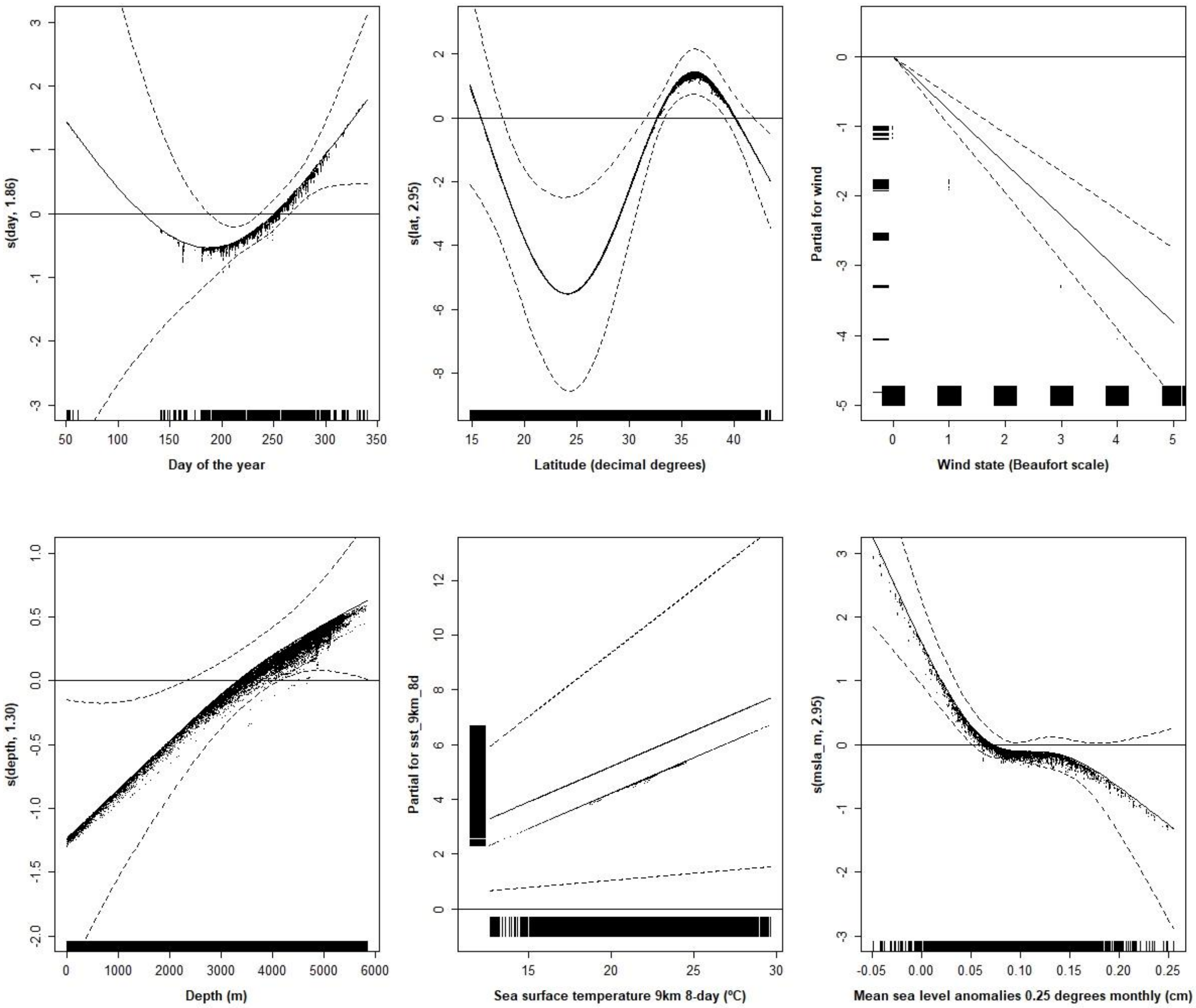


Figure 3. GAM predicted splines of the response variable species presence as a function of the explanatory variables for the environmental model produced for Cuvier's beaked whale (*Z. cavirostris*). The degrees of freedom are in brackets on the y-axis. Tick marks above the x-axis indicate the distribution of observations. Dashed lines delimit the 95% confidence intervals of the spline functions and dots on the graph area represent the residuals. The abbreviations used for the variables are defined in supplemental file 1.

Predicted distributions for the four EEZs (Azores, Canary Islands, Madeira and Portugal) are presented in detail in Figure 5.

The best GAM model was obtained for *Globicephala* sp. with an AUC of 0.93 and 26.8% of explained deviance (Table 1). This species was found to prefer southern latitudes, with occurrence peaking at ~2000 m depth. The best Maxent model corresponded to the most sighted species (*D. delphis*) with a training AUC value of 0.85 and test AUC of 0.826. This was the second species with the best GAM model (AUC of 0.81 and 20.2% deviance explained) (Tables 1 and 2). The species was found to be associated with lower depths and sea surface temperatures; its occurrence decreased from the beginning to the end of the summer months (Supplemental file 3). The species predicted distribution was mainly in coastal continental Portugal and NW Africa, and around the Macaronesian archipelagos (Figure 4).

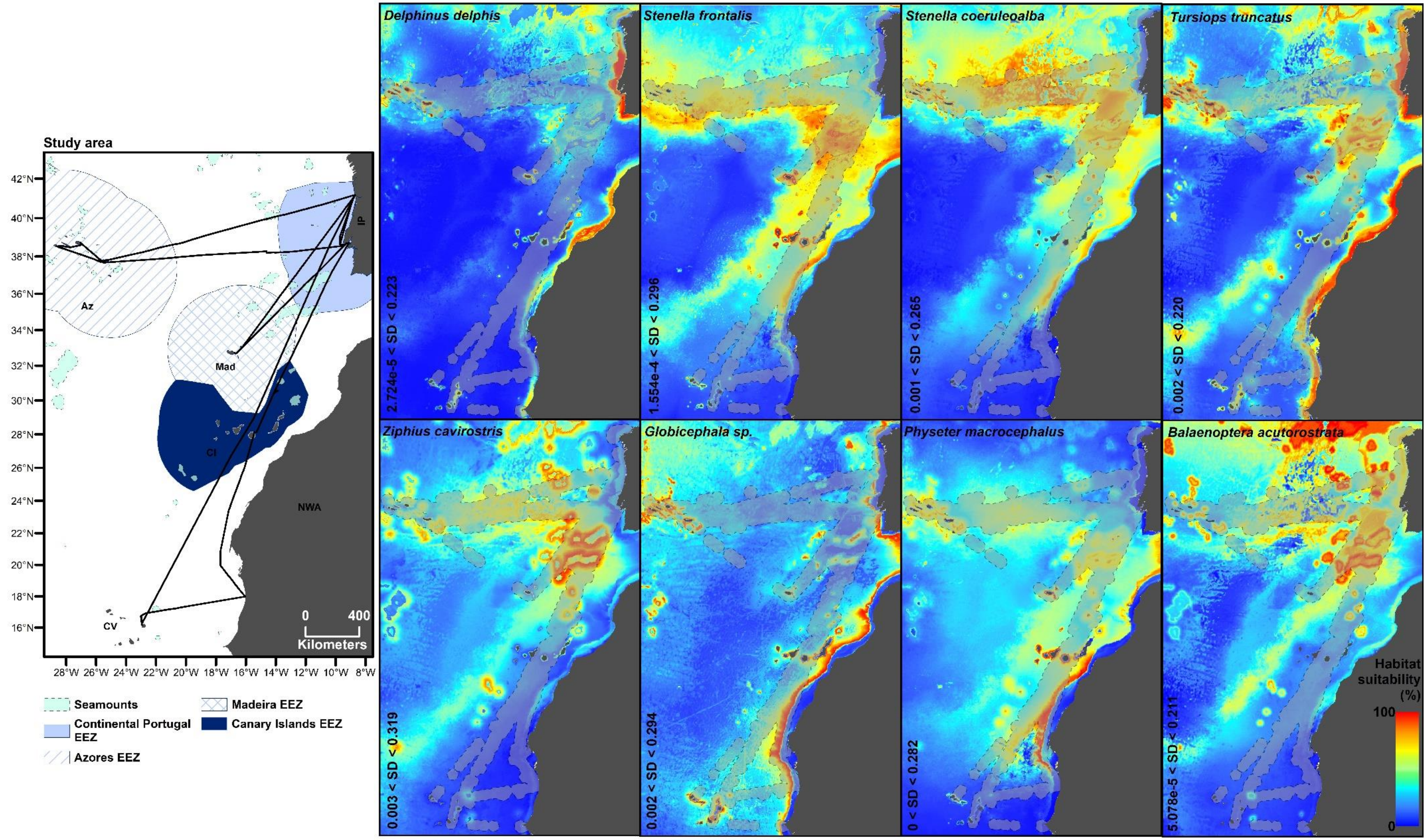


Figure 4. Averaged maps of the eight realised niche models obtained with Maxent. The range of standard deviation (SD) of each model is at the bottom left margin of the maps. The 50 km buffer around each presence point is in light grey delimited by dashed lines.

According to Maxent models, the variable that more contributed to *B. acutorostrata* and *Z. cavirostris* models was distance to seamounts. For *D. delphis*, *Globicephala* sp. and *T. truncatus* models, the most important variable was depth; for *P. macrocephalus*, *S. coeruleoalba* and *S. frontalis* was sea surface temperature (Table 2). Maxent predicted realised niches located in more oceanic waters (for *B. acutorostrata*, *Globicephala* sp., *P. macrocephalus*, S. *coeruleoalba*, *S. frontalis*, and *Z. cavirostris*) and southern areas (for *P. macrocephalus*). On the contrary, *D. delphis* and *T. truncatus* presented suitable habitats mostly in coastal waters; and *B. acutorostrata* and *D. delphis* predictions pointed to a more northern distribution. The models of *Z. cavirostris* and *P. macrocephalus* illustrated narrower distributions, while *S. coeruleoalba*, *S. frontalis* and *T. truncatus* more broad-range distributions (Figures 2 to 4 and Supplemental file 3).

T. truncatus presented peaks of habitat preferences in different areas: latitudes ~15º and ~35º, lower and higher depths, closer and further from seamounts and slightly positive anomalies in altimetry (Figure 2). The predicted realised niche pointed to a broad-range distribution with a strong presence in the four EEZs, mostly in coastal waters and particularly at Azores and continental Portugal (Figure 5). The presence of *Z. cavirostris* increased from the beginning to the end of the summer months, peaked at ~ 35º, increased with depth and sea surface temperature and decreased towards positive anomalies (Figure 3). The predicted distribution reflected a narrow ecological niche, with the species occurring mostly in oceanic areas of the EEZs. The maps highlighted a highly suitable habitat, completely shaped by seamounts, mid-way between southwest Portugal and Madeira island, that overlaps both these EEZs (Figure 5).

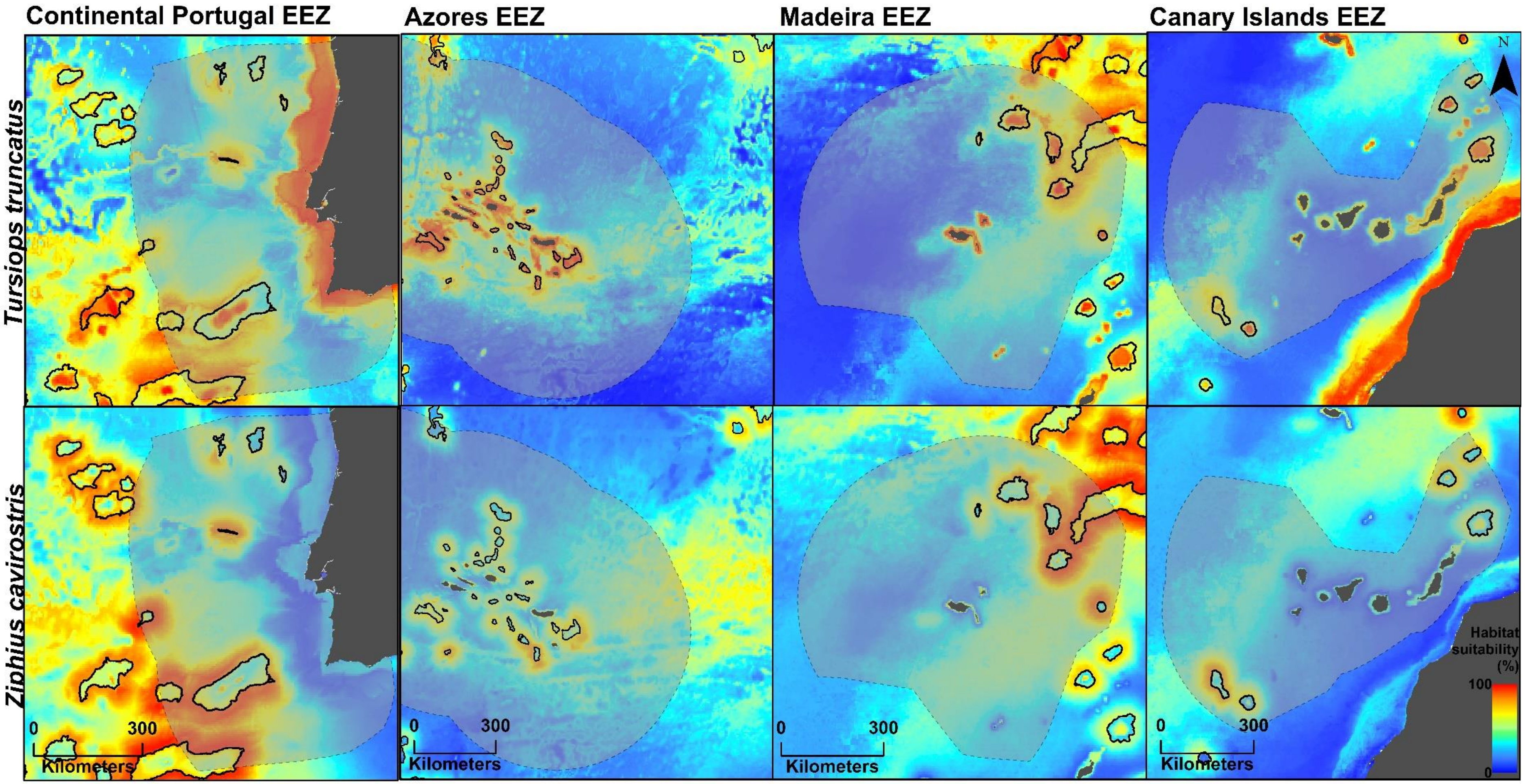


Figure 5. Realised niche models for *Z. cavirostris* and *T. truncatus* in the four European EEZs obtained with Maxent. The areas classified as EEZs are in light grey delimited by dashed lines. The seamounts are delimited at their base by black lines

7.5 Discussion

To our knowledge, this is the first study presenting modelled distribution of cetaceans at basin-scale in the eastern North Atlantic. Many studies focus on cetacean distribution patterns across smaller areas than their ranging capabilities, overlooking the complexity of their biogeographical patterns (Alves *et al.*, 2018b; García-Barón, 2019). Alves *et al.* (2018b) illustrated the connectivity of Macaronesia and Iberian Peninsula for one cetacean species (*Globicephala macrorhynchus*), presenting its wide-range movements across the entire area, and its spatial structuring. The methodology here applied to collect effort-related data across a wide latitudinal and longitudinal range, using OPOs, proved to be valuable to study patterns of distribution at the basin-scale and to develop relevant ecological niche models. This study design can be used in other areas of the globe where distribution of species with a wide range of occurrence needs to be addressed, and financial and logistics challenges are limiting data collection.

Alves *et al.* (2018b) also highlighted the advantages of ecological niche modelling and satellite-linked telemetry to assess the key drivers of the species biogeographical patterns. However, when working with highly mobile species at such a wide scale, we have to be aware of the limitations of ecological niche modelling: we may not be considering all predictors shaping the species distribution, observation data portraits only a sample of cetacean occurrence (as they spend a great amount of time underwater), and we are grouping different populations (e.g., transient and resident, regional sub-populations) (Correia *et al.*, 2019). In fact, although our GAM and Maxent models provided important knowledge and performed better than random models (AUCs higher than 0.5), explained deviances were relatively low. Hence we need to be cautious and avoid over (or erroneous) interpretation of the results. On the other hand, Redfern *et al.* (2017) proved that using datasets from multiple ecosystems, i.e. with a wide range of spatial and temporal variability as in the case of the present study, improves transferability and allows the identification of potential suitable habitats in data-poor areas (if the species ecology remains similar to the ecosystems used to fit the model). Therefore, we are confident with the obtained model predictions of suitable habitats as we used a dataset comprising habitat variability over a wide latitudinal and longitudinal range across the eastern North Atlantic, and we tested several predictors at multiple scales (Fernandez *et al.*, 2018; García *et al.*, 2018; Redfern, 2006).

For the most frequently sighted species, the common dolphin, the best GAM model explained about a fifth of the species distribution. Although the species is the most frequently encountered in the area, also according to literature (Goetz *et al.*, 2014; Silva *et al.*, 2014; Hammond *et al.*, 2013; Tobeña *et al.*, 2016; Alves *et al.*, 2018a), its ecological niche is quite restricted. Its Maxent model performed better than other species, possibly related with the fact that the species is more of an ecological specialist (Correia *et al.*, 2019; Marçalo *et al.*, 2018; Moura *et al.*, 2012), and had the highest number of presences. Suitable habitats are located in areas close to the coast, e.g. Portugal, NW Africa and the surroundings of the Iberian archipelagos. In fact, GAMs indicated that their occurrence decreases with the distance to coast and water depth. The species preferred colder waters, reflected by the northern distribution of suitable habitats (northern colder waters). On the contrary, Fernández *et al.* (2013) reported common dolphin preferences for warmer waters. Also, preliminary analysis of the CETUS dataset, including only two years of surveys and the route from continental Portugal to Madeira, pointed to a positive tendency towards warmer waters (from 16ºC to 19ºC, stabilizing hereafter) (Correia *et al.*, 2015). The three studies have very different ranges of surveyed temperatures (increasing temperature ranges: Fernández *et al.*, 2013; Correia *et al.*, 2015; and the present study). This may indicate that within northern parts (colder) of the study area, at a finer scale, warmer waters are selected by the species. Common dolphin occurrence decreases during summer months. The same trend is observed in the Iberian archipelagos (Alves *et al.*, 2018a; Carrillo *et al.*, 2010; Silva *et al.*, 2014). Higher survey effort in seasons other than summer is needed to assess seasonal shifts in suitable habitat (i.e., answering the question: where do common dolphins go from fall to spring?).

GAM models for both *Stenella* species had fairly low deviance explained. Their predicted niche was widespread across the area, although in more oceanic waters than the other dolphins (common and bottlenose). When species are more ecological generalists, modelling techniques tend to perform less efficiently and present rather low values of deviance explained (Brotons *et al.*, 2004). None of the species presented suitable habitats along the Iberian coastline. In fact, in Portugal, both species are generally encountered beyond the continental platform, mostly in oceanic waters and in the archipelagos (Moura *et al.*, 2017).

Bottlenose dolphins are included in Annex II of the Habitats Directive (Directive 92/43/CEE), hence Member States are required to designated Special Areas of Conservation (SACs) for the protection of this species. Moreover, the species is also selected as priority to assess indicators for the Marine Strategy Framework Directive. This increases the need to have complete knowledge of the species distribution. Overall, the suitable habitats for bottlenose dolphins extend across the entire area. The distribution patterns evidence the coastal areas, both in the mainland (Iberia Peninsula and NW Africa) and in the archipelagos, with suitable habitats extending further to high seas in the mainland. The continental platform and upwelling systems are larger in the mainland than in the archipelagos (Mason, 2009) which may be the reason for a wider range of suitable habitats in coastal areas near Iberia Peninsula and NW Africa. This is even more evident in the north-central coastal area of Portugal which highlights the need to extend conservation efforts into areas further from the coast. On the other hand, the archipelagos present a narrower continental platform which probably restricts suitable habitats for bottlenose dolphins to areas very close to the coast. In the Pelagos Sanctuary (northwest Mediterranean Sea), a social segregation was found between bottlenose dolphins preferring narrower continental platforms and those preferring wider continental platforms (Carnabuci *et al.*, 2016). Depth was, by far, the most important variable for suitable habitats: hence, bottlenose dolphin will be mostly restricted to the continental platforms and, when in areas further from the coast, to the seamounts (less deep areas than the surroundings). Suitable habitat maps show important areas overlapping seamount regions. An assessment of the connectivity between offshore and near-shore populations, as well as between populations near the archipelagos and near the mainland, is needed, hence we suggest dedicated campaigns for biopsy collection of both populations. The majority of the predicted suitable habitat is still within the EEZs, and is restricted to either the Portuguese or the Spanish EEZ. This may facilitate protection measures as individuals occurring within those areas can be included in management units and do not require cooperation between nations to designate a management plan (Santos & Pierce, 2015). GAM model results indicate that the species present different habitat preferences depending on the area they occur. In southern latitudes (from 15ºN to 25ºN), preference is for southern areas. If coupling with the results from Maxent, bottlenose dolphins prefer mainland in NW Africa over the Canary Islands. From 25ºN to 35ºN though, preference increases towards northern waters, likely indicating the habitat in Madeiran waters is more suitable than in the Canary Islands. In waters up to roughly 4000 m depth, the species seems to prefer less deep waters, however from 4000 m to 6000 m preference tends towards deeper waters. We struggle to find an ecological justification for the later. As for the seamounts, it is likely that when in proximity to those areas, the species takes advantage of such structures and moves towards them to benefit from the upwelling phenomena and less deep areas (Pitcher *et al.*, 2007). However, moving away from the seamounts, coastal areas get closer and are preferred by the species (stronger upwelling phenomena, wider areas of shallow waters).

Overall, most suitable habitat for Cuvier’s beaked whale is located at the north of the study area, with the species being less encountered in southern waters. Preferred areas are located in very deep waters and quite far from the coast, as expected, given previous knowledge from literature (Heyning, 1989; Tepsich *et al.*, 2014), Within these areas, seamounts are the most important feature for the species and warmer temperatures are preferred. Results highlight a very important region: the seamounts of the Madeira-Tore, and specifically the Ampère/Coral Patch Seamounts and the Gorringe Bank (Dionísio & Arriegas, 2019). These structures are located between south of mainland Portugal and Madeira island. This area was suitable as well for the other species and given the evidences, it can be indicated as a potential MPA of Importance for Cetacean Conservation, with a specific reference to the concerns for the beaked whales (as to comply with ACCOBAMS 2010 Resolution 4.15; ACCOBAMS - CMS, 2013). The species is "data deficient" in the Atlantic European waters and research effort in high seas will be needed to assess this gap. We recommend dedicated campaigns including photo-ID and biopsy collection in the Madeira-Tore (prioritizing Ampère / Coral Patch and Gorringe Bank). This would result in baseline data on population demography and structuring, as their suitable habitat is quite restricted and knowledge in this area is scarce (with effort focused on coastal waters, e.g., Alves *et al.*, 2018a; Moura *et al.*, 2017; Silva *et al.*, 2014). Finally, we would recommend surveys during autumn or winter, as the species preferences increase towards the end of the summer season.

The best GAM model was obtained for pilot whales, with about a quarter of its distribution explained by the predictors. This species presented the most southern predicted suitable habitats, standing out in Cape Verde waters when comparing with the remaining species. Suitable habitats were also identified along the African coastline, at some distance from the shore (probably at the continental slope given the peak of preference at ~ 2000 m depth). Although identification was not achieved to the species level (as both short-finned and long-finned species pilot whales occur in the area and are almost indistinguishable at sea; Freitas *et al.*, 2012; Hazevoet *et al.*, 2010; Moura *et al.*, 2017), the majority of occurrences is probably from short-finned pilot whale (*Globicephala macrorhynchus*) given its southern range in comparison with the long-finned pilot whale (*Globicephala melas*). In fact, the relationship with latitude was clear in the GAM model, with a preference for southern areas. Alves *et al.* (2018b) proved connectivity of *G. macrorhynchus* in the Iberian archipelagos but connectivity with Cape Verde or the mainland Africa was not assessed. This should be investigated to fully understand the species movements and population structuring within the study area.

The south of the study area is more suited for sperm whales, as occurrence decreased with the increase of latitude in both modelling techniques. The most important variable for this species was temperature, possibly related to the temperature latitudinal gradient across the surveyed area. However, even though occurring in southern areas with warmer waters, there is an overall negative tendency with the increase of temperature. As such, it is likely that, within southern areas, sperm whales prefer colder waters (associated with upwelling systems; Robinson, 2010). Suitable habitat for the species is quite narrow within the area, distributing mostly around Madeira island and Canaries and off NW Africa. In African waters, the species distribution seems to be shaped by the coastline although located at some distance. As described in the literature, sperm whales tend to distribute along the continental slope region where their preferred prey (cephalopods) are more prevalent (Roe, 1969; Tepsich *et al.*, 2014).

Both Madeira island and Canaries have smaller continental platforms than mainland Africa, and so, the continental slope is closer to the coast, facilitating the access to this area for research (considering logistics and costs). Along the African continental slope, suitable habitat is more evident in the waters of Western Sahara and Mauritania. Western Sahara is the state in NW Africa with less information on cetacean occurrence (Correia *et al.*, unpublished data). On the other hand, in Mauritania, more research effort points to a high prevalence of sperm whales (Baines & Reichelt, 2014; Camphuysen, 2012; Correia *et al.*, unpublished data). The NW African waters have several management issues (mostly regarding intensive fishing), some of them arising from poorly managed (or inefficient) fishing agreements with the European Union (Corten, 2014; FAO, 2013; Nagel & Gray, 2012). Given the importance of this area for the sperm whales stock (and likely for pilot whales as well) in the eastern North Atlantic, together with the fact that African countries probably lack the necessary requirements (e.g., financial) to ensure conservation of the species in its waters, it is desirable that an international cooperation takes place (and it should be defined as European priority and responsibility in marine management).

Minke whales were the only baleen species selected, with other baleen whales having a very low prevalence of detections. The species was found mainly in the north of the study area, with suitable habitat restricted to oceanic waters The GAM model presented rather low deviance explained (7.48%). Since in the area both resident and migratory minke whales co-exist (Valente *et al.*, 2019), the habitat use is different among the individuals (migrators VS resident), hence it is expected that models perform worse. Within this area, the research effort should focus on the understanding of the habitat requirements for migratory individuals VS for resident individuals. Moreover, the movements of the whales (latitudinal and longitudinal) should be further investigated (Valente *et al.*, 2019).

Although it was possible to assess the effects of environmental variables, we must highlight that the data was influenced by detectability factors. In all species best GAM models, at least one detectability variable (wind state, sea state or visibility) was included. As expected, and overall, the species occurrence decreases with worse conditions (higher wind speed and wave heights, and worse visibility). Since the predictions obtained with Maxent models do not account for the weather conditions (as we lack spatial data for these variables), the areas of suitable habitat are probably being underestimated: more detections might be made if the weather allowed it.

Three environmental variables stood out as the most important shaping suitable habitats. Depth was the most meaningful variable for species probably depending on the continental platforms in coastal areas and seamounts in offshore waters (common dolphins and bottlenose dolphins) or species with specific peaks of preference for depth (around 2000 m for pilot whales). Seamounts revealed as the most important features for species occurring (almost) completely in oceanic waters (Cuvier's beaked whales and minke whales). Finally, water temperature was the most important variable for species with a very clear relationship with latitude (*Stenella* sp. and sperm whales). Given the wide latitudinal range of the study area, the gradient of temperature is very well defined from northern cold to southern warm waters (Mason, 2009). In fact, both variables were found to be correlated when considering presence/background data in Maxent (instead of used/available habitat for GAMs). As for chlorophyll, the ecological meaning of the relationships modelled with GAMs is hard to interpret, probably due to the existing temporal (and spatial) lag between the measured chlorophyll-a and the available prey to cetaceans (Frederiksen *et al.*, 2006; García *et al.*, 2018). This lag varies amongst cetacean species according to their preys (i.e., position of the preys at the trophic chain). In previous analyses of habitat preferences for common dolphins, we tested chlorophyll delays (i.e., chlorophyll measured one or two weeks or months before the date of cetacean observation) and the best model included chlorophyll without lag (Correia *et al.*, 2019). However, García *et al.* (2018) found chlorophyll delays to be useful when modelling the distribution of blue whales in the Azores archipelago. For such reasons, we advise testing chlorophyll lags in further ecological niche modelling approaches. Altimetry was important in all best GAM models, but the variable was not included when fitting the Maxent models due to its very low resolution. However, it is evident that currents (altimetry acts mainly as a proxy for currents; Robinson, 2010) play an important role in cetacean habitat preferences. To improve resolution, other sources of data or proxies for currents should be tested (e.g., *in situ* measures, current speed or direction)

Both ecological niche modelling techniques have their advantages and drawbacks and in the present study we used each one with a specific objective. Even though GAM is a more robust technique when working with less prevalence of sightings (compared, for example, with GLMs), the model still requires absence (or pseudo-absence) data (Barragán-Barrera *et al.*, 2019; Derville *et al.*, 2018; Fiedler *et al.*, 2018). As such, only effort-related sightings could be used to fit the model. On the other hand, Maxent requires only presence data and thus all sightings could be used which increased significantly the number of presences. Maxent has been successfully used in predicting cetacean distribution (do Amaral *et al.*, 2015; Barragán-Barrera *et al.*, 2019; Derville *et al.*, 2018; Fernandez *et al.*, 2018). However, Maxent has few methods to estimate the error of the predictions (Phillips *et al.*, 2006). For example, there is no measure of deviance explained, so the explanatory capacity of the model is unknown. Moreover, since predictors data has to be available for the entire area, the effect of the detectability variables is overlooked, and dynamic variables are averaged so seasonality is also lost. As such, coupling GAMs proved to be a good strategy in order to not over-interpret the Maxent predictions and still attain a good explanatory power (i.e., understand the influence of detectability and seasonality in the species occurrence) and a map output.

Although our results are based on data collected mostly during summer months, and hence interpretations are only applicable for this season, we advanced the knowledge of cetacean ecological niches in the oceanic waters of the eastern North Atlantic. Nonetheless, year-round monitoring would be essential to understand the seasonality of suitable habitats. It has been previously proved that habitat preferences may change according to seasons (Fernández *et al.*, 2013) and that successful management strategies may depend on the designation of dynamic Marine Protected Areas (Hooker *et al.*, 2011).

Further endeavours in ecological niche modelling should rely on the inclusion of other relevant environmental variables (e.g., currents, thermal fronts), use of delays in the oceanographic variables (e.g., chlorophyll delay), and test more spatial and temporal resolutions, ideally with a bigger dataset which will be available as CETUS Project is predicted to operate in the upcoming years. Modelling should also be performed at finer scales, at least for the areas where highly suitable habitat was predicted. This would require dedicated campaigns to collect high-resolution occurrence data. Coupling broad-scale and narrow-scale models would improve the understanding of suitable habitats distribution across the area (and consequently the conservation of cetacean species). Moreover, at the verge of dramatic changes in the environment, the ecological niche modelling approach should be used to estimate cetacean niches under future climate change scenarios, to support meaningful conservation measures for cetacean community in the eastern North Atlantic.

With the deadline of the Strategic Plan for Biodiversity 2010-2020 fast approaching, only 7.4% of the global ocean is protected against the established 10% target to be achieved by 2020 (according to the last report, UNEP-WCMC *et al.*, 2018). When comparing protection in areas within the EEZs with areas beyond national jurisdiction, the difference gap is enormous: 16.8% against 1.2% (UNEP-WCMC *et al.*, 2018). Within the European waters, several areas were designated as Sites of Community Importance (SCI) in the Natura 2000 network, although few have a management plan defined/reported (http://natura2000.eea.europa.eu/; www.protectedplanet.net). Our results have highlighted offshore seamounts highly suitable habitats for all eight species, and specifically for bottlenose dolphins and Cuvier's beaked whales. These are evidently areas that need the definition of management plans, including the implementation of monitoring programmes.

One of the hardest challenges in managing protected areas in high seas (and, specifically, when including areas beyond national jurisdiction) is to guarantee continuity: transboundary remote areas usually require complex and expensive logistics to get access. Cost-effective monitoring programmes, potentially relying on OPOs, such as the CETUS Project, or based on new technologies that allow remote monitoring (e.g., use of automated vehicles) may be the solution to guarantee the data collection. Also, ecological modelling approaches or other relatively cheap analysis (e.g, environmental DNA, photo-ID techniques) are potentially the adequate methods to support long-term and efficient management and conservation of these remote areas (Bohorquez *et al.*, 2019).

7.6 Acknowledgements

We thank the contribution and dedication of all the volunteers during the monitoring campaigns. This is also a product of the work of every observer that participated in the CETUS Project. We are extremely grateful to TRANSINSULAR, the cargo ship company that provided all the logistic support, and to the ships' crews for their hospitality.

7.7 References

ACCOBAMS – CMS (2013). *Report of the Fifth Meeting of the Parties to ACCOBAMS (Appendix 3).* ACCOBAMS – MOP5, 329pp.

AVISO (2017). Sea-surface Height Products. https://www.aviso.altimetry.fr/en/data/products/sea-surface-height-products/global/msla-h.html (accessed September 2017).

Alves F., Ferreira R., Fernandes M., Halicka Z., Dias L. and Dinis A. (2018a). Analysis of occurrence patterns and biological factors of cetaceans based on long-term and fine-scale data from platforms of opportunity: Madeira Island as a case study. *Marine Ecology*, 39: e12499.

Alves F., Alessandrini A., Servidio A., Mendonça A.S., Hartman K.L., Prieto R, *et al.* (2018b). Complex biogeographical patterns support an ecological connectivity network of a large marine predator in the north-east Atlantic. *Diversity and Distributions*, 1-16.

do Amaral K.B., Alvares D.J., Heinzelmann L., Borges-Martins M., Siciliano S. and Moreno I.B. (2015). Ecological niche modeling of *Stenella* dolphins (Cetartiodactyla: Delphinidae) in the southwestern Atlantic Ocean. *Journal of Experimental Marine Biology and Ecology*, 472: 166-179.

Arcangeli A., Campana I., Marini L. and MacLeod C.D. (2016). Long-term presence and habitat use of Cuvier's beaked whale (*Ziphius cavirostris*) in the Central Tyrrhenian Sea. *Marine Ecology*, 37: 269-282.

Baines M.E. and Reichelt M. (2014). Upwellings, canyons and whales: An important winter habitat for balaenopterid whales off Mauritania, northwest Africa. *Journal of Cetacean Research and Management,* 14: 57-67.

Barragán-Barrera D.C., do Amaral K.B., Chávez-Carreño P.A., Farías-Curtidor N., Lancheros-Neva R., Botero-Acosta N., *et al.* (2019). Ecological Niche Modeling of Three Species of *Stenella* Dolphins in the Caribbean Basin, With Application to the Seaflower Biosphere Reserve. *Frontiers in Marine Science*, 6: 10.

Beck J.R. and Shultz E.K. (1986). The use of relative operating characteristic (ROC) curves in test performance evaluation. *Archives of Pathology & Laboratory Medicine*, 110: 13-20.

Bohorquez J.J., Dvarskas A. and Pikitch E.K. (2019). Filling the Data Gap – A Pressing Need for Advancing MPA Sustainable Finance. *Frontiers in Marine Science*, 6: 45.

Breen P., Brown S., Reid D. and Rogan E. (2017). Where is the risk? Integrating a spatial distribution model and a risk assessment to identify areas of cetacean interaction with fisheries in the northeast Atlantic. *Ocean and Coastal Management*, 136: 148-155.

Brotons L., Thuiller W., Araújo M.B. and Hirzel, A.H. (2004). Presence−absence versus presence-only modelling methods for predicting bird habitat suitability. *Ecography*, 27: 437-448.

Burnham K.P. and Anderson D.R. (2002). *Model Selection and Multimodel Inference: a Practical Information Theoretic Approach*. Springer Verlag: New York, USA.

Caldeira R.M.A., Groom S., Miller P., Pilgrim D. and Nezlin N.P. (2002). Sea-surface signatures of the Island mass effect phenomena around Madeira Island, Northeast Atlantic. *Remote Sensing of Environment*, 80: 336-360.

Camphuysen C.J., van Spanje T.M. and Verdaat H. (2012). *Ship based seabird and marine mammal surveys off Mauritania, Nov-Dez 2012 – cruise report.* Mauritanian Institute for oceanographic research and fisheries - IMROP, 73pp.

Cañadas A., Sagarminaga R., De Stephanis R., Urquiola E. and Hammond, P.S. (2005). Habitat preference modelling as a conservation tool: proposals for marine protected areas for cetaceans in southern Spanish waters. *Aquatic Conservation*, 15: 495-521.

Carrillo M., Pérez-Vallazza C.P. and Alvarez-Vázquez R. (2010). Cetacean diversity and distribution off Tenerife (Canary Islands). *Marine Biodiversity Records*, 3: 1-9.

Carnabuci M., Schiavon G., Bellingeri M., Fossa F., Paoli C., Vassallo P and Gnone G. (2016). Connectivity in the network macrostructure of *Tursiops truncatus* in the Pelagos Sanctuary (NW Mediterranean Sea): does landscape matter? *Population Ecology*, 58(2).

Correia A.M., Tepsich P., Rosso M., Caldeira R. and Sousa-Pinto I. (2015). Cetacean occurrence and spatial distribution: Habitat modelling for offshore waters in the Portuguese EEZ (NE Atlantic). *Journal of Marine Systems*, 143: 73-85.

Correia A.M., Gil A., Valente R., Rosso M., Pierce G.J. and Sousa-Pinto I. (2019, in press) Distribution and habitat modelling for short-beaked common dolphins (*Delphinus delphis*) in Eastern North Atlantic Ocean. *Journal of the Marine Biological Association UK*, 1-15.

Corten A. (2014). EU–Mauritania fisheries partnership in need of more transparency. *Marine Policy* 49: 1-11.

Derville S., Torres L.G., Iovan C. and Garrigue C. (2018). Finding the right fit: Comparative cetacean distribution models using multiple data sources and statistical approaches. *Diversity and Distributions*, 24(11): 1657-1673.

Dionísio M.A. and Arriegas P.I. (2018). *Madeira Tore*. Report. https://www.cbd.int/marine/doc/submissions/2015-071/Portugal-submission-2-en.pdf

Elith J. and 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.

Elith J., Phillips S.J., Hastie T., Dudík M., Chee Y.E. and Yates C.J. (2011). A statistical explanation of Maxent for ecologists. *Diversity and Distribution*, 17: 43-47.

ESRI (2016). ArcGIS Desktop: Release 10.5. Redlands, CA: Environmental Systems Research Institute.

FAO (2013) *Marine protected areas - Country case studies on policy, governance and institutional issues.* In: FAO Fisheries and Aquaculture Technical Paper. Food and Agriculture Organization of the United Nations. Available at http://www.fao.org/publications/card/en/c/452ece29-e667-5f47-92a5-8a0e8eb5cbab/ [last access 8 November 2018].

Fernandez M., Yesson C., Gannier A., Miller P.I. and Azevedo J.M.N. (2018). A matter of timing: how temporal scale selection influences cetacean ecological niche modelling. *Marine Ecology Progress Series*, 595: 217-231.

Fernández R., MacLeod C.D., Pierce G.J., Covelo P., López A., Torres-Palenzuela J., *et al.* (2013). Inter-specific and seasonal comparison of the niches occupied by small cetaceans off north-west Iberia. *Continental Shelf Research*, 64: 88-98.

Fiedler P.C., Redfern J., Forneu K.A., Palacios D.M., Sheredy C., Rasmussen K., *et al.* (2018). Prediction of Large Whale Distributions: A Comparison of Presence–Absence and Presence-Only Modeling Techniques. *Frontiers in Marine Science*, 5: 419.

Frederiksen M., Edwards M., Richardson A.J., Halliday N.C. and 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(6):1259-1268.

Freitas L., Dinis A., Nicolau C., Ribeiro C. and Alves F. (2012). New records of cetacean species for Madeira Archipelago with an updated checklist. *Boletim do Museu Municipal do Funchal*, 62: 25-43.

Friedlaender A., Johnston D.W., Fraser W., Burns J., Halpin P. and Costa D. (2011). Ecological niche modeling of sympatric krill predators around Marguerite Bay, western Antartic Peninsula. *Deep-Sea Research*, 58: 1729-1740.

Fourcade Y., Engler J.O., Rödder D. and Secondi J. (2014). Mapping Species Distributions with MAXENT Using a Geographically Biased Sample of Presence Data: A Performance Assessment of Methods for Correcting Sampling Bias. *PLoS ONE*, 9(5): e97122.

García-Barón I., Authier M., Caballero A., Vázquez J. A., Santos M. B., Murcia J. L., and Louzao, M. (2019). Modelling the spatial abundance of a migratory predator: A call for transboundary marine protected areas. *Diversity and Distributions*, 25(3): 346-360.

García L.G., Pierce G.J., Autret E., Torres-Palenzuela J.M. (2018). Multi-scale habitat preference analyses for Azorean blue whales. *PLoS ONE,* 13(9): e0201786.

GEBCO (2017). Gridded bathymetry data. http://www.gebco.net/data_and_products/gridded_bathymetry_data/ (accessed September 2017).

Goetz S., Read F.L., Ferreira M., Portela J.M., Santos M.B., Vingada J., Siebert U., *et al.* (2015). Cetacean occurrence, habitat preferences and potential for cetacean - fishery interactions in Iberian Atlantic waters: results from cooperative research involving local stakeholders*. Aquatic Conservation: Marine and Freshwater Ecosystems,* 25:138-154.

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

Hazevoet C.J., Monteiro V., López P., Varo N., Torda G., Berrow S. and Gravanita B. (2010). Recent data on whales and dolphins (Mammalia: Cetacea) from the Cape Verde Islands, including records of four taxa new to the archipelago. *Zoologia Caboverdiana* 1: 75-99.

Heyning J.E. (1989). Cuvier's beaked whale *Ziphius cavirostris* g. Cuvier, 1823. In: Ridgway S.H., & Harrison R.S. (eds) *Handbook of marine mammals, Vol 4. River dolphins and the larger toothed whales*. Academic Press, London, pp. 289-308.

Hijmans R.J., Phillips S., Leathwick J. and Elith J. (2017). dismo: Species Distribution Modeling. https://CRAN.R-project.org/package=dismo.

Hooker S.K., Cañadas A., Hyrenbach K.D., Corrigan C., Polovina J.J. and Reeves, R.R. (2011). Making protected area networks effective for marine top predators. *Endangered Species Research*, 13: 203-218.

Jungblut S., Nachtsheim D.A., Boos K. and Joiris C.R. (2017). Biogeography of top predators – seabirds and cetaceans – along four latitudinal transects in the Atlantic Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography*, 141: 59-73.

Kiszka J., Van Canneyt O., Macleod K., Walker D. and Ridoux, V. (2007). Distribution, encounter rates and habitat characteristics of toothed cetaceans in the Bay of Biscay and adjacent waters from platform of opportunity data. *ICES Journal of Marine Science*, 64: 1033-1043.

Liu C., Berry P.M., Dawson T.P. and Pearson R.G. (2005). Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, 28: 385-393.

Marçalo A., Nicolau L., Giménez J., Ferreira M., Santos J., Araújo H., Silva A., *et al.* (2018). Feeding ecology of the common dolphin (*Delphinus delphis*) in Western Iberian waters: has the decline in sardine (*Sardina pilchardus*) affected dolphin diet? *Marine Biology*, 165:44.

Marubini F., Gimona A., Evans P.G.H., Wright P.J. and Pierce G.J. (2009). Habitat preferences and interannual variability in occurrence of the harbour porpoise *Phocoena phocoena* off northwest Scotland. *Marine Ecology Progress Series*, 381: 297-310.

Mason E. (2009). High-resolution Modelling of the Canary Basin Oceanic Circulation. Doctorate thesis (PhD) in Oceanography, Universidad de Las Palmas de Gran Canaria, Las Palmas de Gran Canaria. 245 pp.

Moura A.E., Sillero N. and Rodrigues A. (2012). Common dolphin (*Delphinus delphis*) habitat preferences using data from two platforms of opportunity. *Acta Oecologica*, 38: 24-32.

Moura A.E., Silva S.E., SPEA, Correia A.M., Sousa-Pinto I., Gil A., *et al.* (2017). In: Bencatel J., Álvares F., Moura A. E. and Barbosa, A. M. (eds.) (2017). *Atlas de Mamíferos de Portugal.* Universidade de Évora, Portugal.

Nagel P. and Gray T. (2012). Is the EU's Fisheries Partnership Agreement (FPA) with Mauritania a genuine partnership or exploitation by the EU? *Ocean & Coastal Management*, 56: 26-34.

NASA (2017). Ocean color data — MODISA. https://oceandata.sci.gsfc.nasa.gov/ (accessed September 2017).

Parsons E.C.M. (2016). Why IUCN should replace ‘Data Deficient’ conservation status with a precautionary ‘Assume Threatened’ status—A cetacean case study. *Frontiers in Marine Science*, 3: 193.

Passadore C., Möller L.M., Diaz-Aguirre F. and Parra G.J. (2018). Modelling dolphin distribution to inform future spatial conservation decisions in a marine protected area. *Scientific reports*, 8(1): 15659.

Pearce J.L. and Boyce M.S. (2006). Modelling distribution and abundance with presence only data. *Journal of Applied Ecology*, 43: 405-412.

Pennino M.G., Arcangeli A., Fonseca V.P., Campana I., Pierce G.J., Rotta A. and Bellido J.M. (2017). A spatially explicit risk assessment approach: Cetaceans and marine traffic in the Pelagos Sanctuary (Mediterranean Sea). *PLoS ONE*, 12: e0179686.

Phillips S.J. (2008). Transferability, sample selection bias and background data in presence-only modelling: a response to Peterson, *et al.* (2007). *Ecography*, 31: 272-278.

Phillips S.J., Anderson R.P. and Schapire R.E. (2006). Maximum entropy modelling of species geographic distributions. *Ecological Modelling*, 190: 231-259.

Phillips S.J., Anderson R.P., Dudík M., Schapire R.E. and Blair M.E. (2017). Opening the black box: an open-source release of Maxent. *Ecography*, 40: 887-893.

Phillips S.J., Dudík M., Elith J., Graham C.H., Lehmann A., Leathwick J. and Ferrier S (2009). Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological Applications*, 19: 181-197.

Pitcher T.J., Morato T., Hart P.J.B., Clark M.R., Haggan N. and Santos R.S. (2007). *Seamounts: Ecology, Fisheries & Conservation*. Fish and Aquatic Resources Series 12. Blackwell Publishing, Oxford, UK.

Praca E., Gannier A., Das K. and Laran S. (2009). Modelling the habitat suitability of cetaceans: example of the sperm whale in the northwestern Mediterranean Sea. *Deep-Sea Research I*, 56: 648-657.

Prieto R., Tobeña M. and Silva M.A. (2017). Habitat preferences of baleen whales in a mid-latitude habitat. *Deep-Sea Research II*, 141: 155-167.

Quian S.S. (2009). *Environmental and Ecological Statistics with R*. Chapman & Hall: Canada.

Raes N. and ter Steege H. (2007) A null-model for significance testing of presence-only species distribution models. *Ecography*, 30: 727-736.

R Development Core Team (2012). R: A Language and Environment for Statistical Computing. Vienna: R foundation for Statistical Computing. http://www.R-project.org/.

Redfern J.V., Ferguson M.C., Becker E.A., Hyrenbach K.D., Good C., Barlow J., *et al.* (2006). Techniques for cetacean-habitat modeling: a review. *Marine Ecology Progress Series*, 310: 271-295.

Redfern J.V., Moore T.J., Fiedler P.C., de Vos A., Brownell R.L., Forney K.A., *et al.* (2017). Predicting cetacean distributions in data-poor marine ecosystems. *Diversity and Distributions*, 23(4): 394-408.

Roberts J.J., Best B.D., Dunn D.C., Treml E.A. and Halpin P.N. (2010). Marine Geospatial Ecology Tools: An integrated framework for ecological geoprocessing with ArcGIS, Python, R, MATLAB, and C++. *Environmental Modelling & Software*, 25: 1197-1207.

Robinson I.S. (2010). *Discovering the Ocean from Space – The Unique Applications of Satellite Oceanography*. Springer: Bath, UK.

Sala I., Caldeira R.M.A., Estrada-Allis S.N., Froufe E. and Couvelard X. (2013). Lagrangian transport pathways in the northeast Atlantic and their environmental impact. *Limnology and Oceanography – Fluids & Environments*, 3: 40-60.

Santos M.B. and Pierce G.J. (2015). Marine mammals and good environmental status: science, policy and society; challenges and opportunities. *Hydrobiologia* 750: 13-41.

Sillero N. (2011). What does ecological modelling model? A proposed classification of ecological niche models based on their underlying methods. *Ecological Modelling*, 222: 1343-1346.

Silva M.A., Prieto R., Cascão I., Seabra M.I., Machete M., Baumgartner M.F. and Santos R.S. (2014). Spatial and temporal distribution of cetaceans in the mid-Atlantic waters around the Azores. *Marine Biology Research*, 10: 123-137.

Smith J.M., Pierce G.J., Zuur A.F., Martins H., Martins M.C., Porteiro F. and Rocha F. (2011). Patterns of investment in reproductive and somatic tissues in the loliginid squid *Loligo forbesii* and *Loligo vulgaris* in Iberian and Azorean waters. *Hydrobiologia*, 670: 201-221.

Storrie L., Lydersen C., Andersen M., Wynn R.B., and Kovacs K.M. (2018). Determining the species assemblage and habitat use of cetaceans in the Svalbard archipelago, based on recorded observations from 2002-2014. *Polar Research*, 37 (1): 1463065.

Tepsich P., Rosso M., Halpin P.N. and Moulins A. (2014). Habitat preferences of two deep-diving cetacean species in the northern Ligurian Sea. *Marine Ecology Progress Series*, 508: 247-260.

Thorne L.H., Johnston D.W., Urban D.L., Tyne J., Bejder L., Baird R.W., *et al*. (2012). Predictive modeling of spinner dolphin *Stenella longirostris* resting habitat in the Main Hawaiian Islands. *PLoS ONE*, 7: e43167.

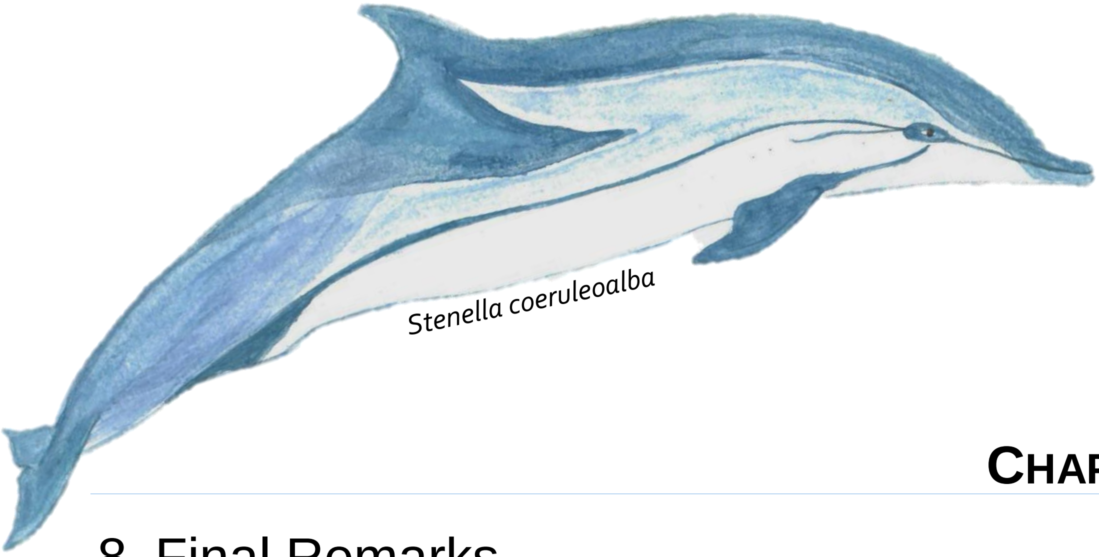
Tobeña M., Prieto R., Machete M. and Silva M.A. (2016). Modeling the potential distribution and richness of cetaceans in the Azores from fisheries observer program data. *Frontiers in Marine Science*, 3: 202.

UNEP-WCMC, IUCN & NGS (2018). *Protected Planet Report 2018*. UNEP-WCMC, IUCN and NGS: Cambridge UK; Gland, Switzerland; and Washington, D.C., USA.

Valente R., Correia A.M., Gil Á., González L.G. and Sousa-Pinto I. (2019). Baleen whales in Macaronesia: occurrence patterns revealed through a bibliographic review. Mammal Review, 49(2): 129-151.

Viddi F.A., Hucke-Gaete R., Torres-Florez J.P. and Ribeiro S. (2010). Spatial and seasonal variability in cetacean distribution in the fjords of northern Patagonia, Chile. ICES Journal of Marine Science, 67: 959-970.

Zuur A.F., Ieno E. and Elphick C. (2010). A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution*, 1: 3-14.



CHAPTER VIII.

8 Final Remarks

8.1 General Discussion and Conclusion

8.1.1 Achievements, Limitations and Advantages

With this work, we successfully achieved the goals we proposed to better understand cetacean occurrence, distribution patterns and habitat preferences in the eastern north Atlantic. First, we provided an open-source effort-related dataset on cetacean occurrence (CETUS dataset), available through OBIS and EMODnet, using the recent biodiversity data standards, the OBIS-ENV-DATA format (De Pooter *et al.*, 2017). The dataset contains 12405 events, 9440 occurrences (2833 of cetacean species) on 44 taxa (30 of the infraorder Cetacea) and 86022 measurements or facts, spanning the period from 2012 to 2017 and including nearly 500 days of on-effort surveys. The records are mostly concentrated in high-seas which are data-poor areas where less research effort has been previously undertaken. CETUS dataset can be used as a model for similar visual line transect data collections, hence this is also a contribution to the increasing research in data science. In fact, large collections of open-source standardized data are important tools for both researchers and decision-makers, which in turn, improves knowledge over larger scales and supports conservation measures and management plans (Halpin *et al.*, 2006).

During the CETUS campaigns, the survey effort was very heterogeneous, as it depended on the schedules and routines of the cargo ship company and also on the weather conditions. Weather conditions highly affect the capacity of the observer to spot the animals. Although survey effort is interrupted with very poor weather conditions, the effect of detectability factors is significant when building ecological niche models. Thus, variables related with detectability (e.g., wave height, wind speed, visibility range) should be considered as predictors when fitting the models.

Effort was not highly correlated with encounter rates, which suggests that relative abundances provided are reliable and not significantly biased by heterogeneous effort coverage. Unfortunately, CETUS campaigns were mostly concentrated between July and October, due to the winter weather conditions that often inhibit cetacean monitoring. This impedes seasonal assessments and it has been proven that habitat preferences may change according to seasons (Fernández *et al.*, 2013). Hence, the results here presented are mostly representative of the summer season.

The limitations of this work need to be recognized and we have to be realistic of how much we can explain of cetacean distribution in the eastern north Atlantic. In other words, we are modelling highly mobile species occurring over a very wide area. We are probably not including all explanatory variables in the analysis (or even the most appropriate) and sightings data represent only a sample of cetacean occurrence as they spend a great amount of time underwater. Moreover, we are including individuals that use the area differently (e.g., transient or resident populations, resting or foraging individuals) or that belong to different sub-populations in the same analysis. Hence, we were cautious in drawing conclusions and providing recommendations by avoiding overinterpretation or erroneous interpretations.

Nevertheless, in order to provide reliable and useful information, different analyses were undertaken: we combined descriptive techniques, ecological niche modelling and spatial mapping, and tested several explanatory variables at different spatial and temporal scales. Also, we worked at a wide latitudinal and longitudinal range to have a good representation of the eastern north Atlantic. Redfern *et al.* (2017) proved that using datasets from multiple ecosystems, i.e. with a wide range of spatial and temporal variability, improves transferability/extrapolation of model results.

Often, studies on distribution patterns are made at smaller scales than the occurrence range of the species. It is also frequent to define management units, for political and/or practical rather than scientific reasons, as the part of the population that falls within the waters of a state (Santos & Pierce, 2015). However, at sea, there are no physical barriers and cetaceans are highly mobile species. So, effective conservation of species is only achieved when considering their entire range of occurrence, which calls for transboundary protected areas (Alves *et al.*, 2018; García-Barón, 2019). Realizing that to protect populations occurring in one's nation requires looking beyond its administrative boundaries is key to hold responsibility for the management of waters beyond its jurisdiction. For example, management of northwestern African waters, an important area for eastern north Atlantic cetacean stocks with a high cetacean biodiversity, is also the responsibility of European states - especially so, given that conservation issues in the area are mostly related with inefficient fishery agreements with Europe (Corten, 2014; FAO, 2013; Nagel & Gray, 2011).

8.1.2 Species distribution, habitat preferences and priority areas

This work reports a great biodiversity of cetacean species in high-seas, and specifically in areas beyond national jurisdiction. It also pointed out that most likely diversity in these waters has been underestimated, as number of reported species increases with survey effort (only reaching a plateau after ~ 2000 km surveyed for each 100 km^2). Besides, there were highly suitable habitats for several species in international waters, and, for example, it was the area where *Z. cavirostris* was most frequently encountered. Evidently, management in high-seas is challenging, and while national waters are under several agreements and regulations (e.g., Habitats Directive and MSFD), management of areas beyond national jurisdiction demands international articulation and agreements. Unsurprisingly, the 16.8% of protected areas in national waters contrasts with 1.2% of protected areas in international waters (UNEP-WCMC *et al.*, 2018). And, from this low percentage, a very small number of protected areas has management plans defined or reported (www.protectedplanet.net).

In high-seas, the most important features shaping cetacean niches were seamounts, which assumed a higher importance for more oceanic species such as *Z. cavirostris* and *B. acutorostrata*. Depth was also a very important variable shaping cetacean distribution, mostly for *D. delphis*, *T. truncatus* and *Globicephala* sp. However, besides static topographic variables, oceanographic variables were significant, with the sea-surface temperature being the most relevant predictor for *S. frontalis*, *S. coeruleoalba* and *P. macrocephalus*. In the area, sea-surface temperature has a very marked latitudinal gradient, shifting from colder northern waters to warmer southern waters. Moreover, the strong, highly productive upwelling systems, occurring mostly in the mainland coasts (Iberian Peninsula and northwestern Africa) are also characterized by lower temperatures (Mason, 2009).

Seasonality of oceanographic processes such as the upwelling and current systems have an evident influence in determining the distribution of cetaceans as they determine the locations of the feeding and nursery grounds and migratory corridors. As such, conservation of these species has to account for this dynamism. In fact, dynamic MPAs have been suggested as the most efficient way to ensure cetacean protection (Hooker *et al.*, 2011) because: their distribution is conditioned by dynamic oceanographic processes; habitat preferences may change seasonally (Fernández *et al.*, 2013); and to capture the entire life cycle of the species, e.g., the long seasonal migrations between feeding and breeding grounds (García-Barón, 2019; Valente *et al.*, 2019).

The most frequently sighted species in the area was the common dolphin (*D. delphis*), and predictions of a fairly restricted niche indicated that the species is more of an ecological specialist than generalist.

Predicted suitable habitats for the eight most frequently sighted species can be summarized as follows:

- **Common dolphins** preferred areas located north, in coastal shallow waters mostly in the Iberian Peninsula and the Azores archipelago.
- ***Stenella* sp.** presented a wide ecological niche, distributing mostly in oceanic waters and in the archipelagos, being almost completely absent along the Iberian coastline.
- **Bottlenose dolphin** (*T. Tursiops*) suitable habitat was associated with continental platforms, thus extending further offshore in mainland (Iberia and northwest Africa) where these are wider than in the islands. However, there were also very important

areas for the species in offshore waters, where suitable habitat was mostly in the seamounts.

- **Cuvier's beaked whales** (*Z. cavirostris*) preferred northern oceanic waters with seamounts shaping its ecological niche. However, unlike bottlenose dolphins, the species preferred the areas at the base of the seamounts (probably related with distribution of its main prey, squids, in benthic habitats; Tepsich *et al.*, 2014).
- **Pilot whales** (*Globicephala* sp.) and **sperm whales** (*P. macrocephalus*) occurred further south with suitable habitats in the northwestern Africa and Cape Verde standing out.
- **Minke whales** (*B. acutorostrata*) preferred northern oceanic waters, in areas closer to the seamounts.

8.1.3 Recommendations for future research and conservation efforts

Conservation or further dedicated monitoring efforts in the eastern north Atlantic should start in priority areas where suitable habitats were predicted with higher probability. Studies should focus on the population structure and connectivity between regions, species movements and effects of anthropogenic impacts at the basin-scale. As for ecological niche modelling, improvements would be achievable through the testing of other variables, scales or techniques, the use of a bigger dataset with a higher effort coverage (spatially and temporally), the ensemble of several models or coupling results from low and high-resolution, and broad and narrow-scale models. Such endeavours may be crucial for an integrated conservation of cetacean populations across the eastern north Atlantic.

Efficient conservation measures and management plans need to be: long-term and cost effective to ensure continuity; transboundary to guarantee protection of the entire range of species occurrence; and dynamic given changes in habitat preferences and influence of oceanographic processes. Such requirements pose some challenges for cetacean conservation and overall marine management which, in summary, are:

1. Get data in remote data-poor areas. Programmes such as CETUS, that ally the cost-effective use of OPOs with dedicated monitoring (i.e., dedicated observers onboard performing a standard cetacean monitoring protocol), may be the solution to ensure continuity and long-term monitoring in areas, where, often, even baseline data is lacking. Also, the adoption of cheap non-invasive and new techniques such as sophisticated multi-scale ecological niche modelling, use of automated vehicles, photo-ID and environmental DNA (Bohorquez *et al.*, 2019). Finally, insufficient data or data gaps must not lead to inaction, and conservation measures should be based on

available information, and focus on attain more knowledge where it is lacking (Santos & Pierce, 2015).

2. Protection of international waters (in areas beyond national jurisdiction). This will require international agreements and highly depends on the worldwide capacity to work in cooperation towards conservation and management goals.
3. Achieve year-round monitoring. A cost-effective less weather-dependent solution is needed (at least for the winter season and specially for the less accessible high-seas). The key may be in passive monitoring techniques less dependent on factors affecting the human capacity to detect animals: for example, acoustic monitoring from fixed stations (e.g., offshore oceanographic buoys) or the use of autonomous vehicles (although these are still, to some degree, weather dependent).

These challenges should be addressed in the face of technologic developments, new analytical and data collection methods, the increasing access to open-source datasets, and the international cooperation and data-sharing.

The CETUS project is ongoing and will continue to provide effort-related data on cetacean occurrence feeding the open-source CETUS dataset, which will allow the improvement of the ecological niche models. Further analysis using the dataset should address spatial marine planning and model cetacean distribution under future climate change scenarios. Climate change is arguably the biggest challenge we face and addressing its effects on ocean health is priority in conservation. It is urgent to understand how biodiversity will respond, and objectively, how ecological niches, occurrence ranges and distribution patterns will change (increase, decrease or shift). As better informed we are, the better “equipped” to respond and find solutions to ensure efficient management strategies and biodiversity conservation.

8.2 References

Alves F., Ferreira R., Fernandes M., Halicka Z., Dias L. and Dinis A. (2018). Analysis of occurrence patterns and biological factors of cetaceans based on data from platforms of opportunity: Madeira Island as a case study. *Marine Ecology*, e12499.

Bohorquez J.J., Dvarskas A. and Pikitch E.K. (2019). Filling the Data Gap – A Pressing Need for Advancing MPA Sustainable Finance. *Frontiers in Marine Science*, 6: 45.

Corten A. (2014). EU–Mauritania fisheries partnership in need of more transparency. *Marine Policy,* 49: 1-11.

De Pooter D., Appeltans W., Bailly N., Bristol S., Deneudt K., Eliezer M., *et al.* (2017). Toward a new data standard for combined marine biological and environmental datasets-expanding OBIS beyond species occurrences. *Biodiversity Data Journal*, 5: e10989.

García-Barón I., Authier M., Caballero A., Vázquez J. A., Santos M. B., Murcia J. L., and Louzao, M. (2019). Modelling the spatial abundance of a migratory predator: A call for transboundary marine protected areas. *Diversity and Distributions*, 25(3): 346-360.

FAO (2013) *Marine protected areas - Country case studies on policy, governance and institutional issues*. In: FAO Fisheries and Aquaculture Technical Paper. Food and Agriculture Organization of the United Nations. Available at http://www.fao.org/publications/card/en/c/452ece29-e667-5f47-92a5-8a0e8eb5cbab/ (accessed November 2018).

Halpin P.N., Read A.J., Best B.D., Hyrenbach K.D., Fujioka E., Coyne M.S., *et al.* (2006). OBIS-SEAMAP: developing a biogeographic research data commons for the ecological studies of marine mammals, seabirds, and sea turtles. *Marine Ecology Progress Series,* 316: 239-246.

Hooker S.K., Cañadas A., Hyrenbach K.D., Corrigan C., Polovina J.J. and Reeves, R.R. (2011). Making protected area networks effective for marine top predators. *Endangered Species Research*, 13: 203-218.

Fernández R., MacLeod C.D., Pierce G.J., Covelo P., López A., Torres-Palenzuela J., *et al.* (2013). Inter-specific and seasonal comparison of the niches occupied by small cetaceans off north-west Iberia. *Continental Shelf Research*, 64: 88-98.

Mason E. (2009). High-resolution Modelling of the Canary Basin Oceanic Circulation. Doctorate thesis (PhD) in Oceanography. Universidad de Las Palmas de Gran Canaria, Las Palmas de Gran Canaria. 245 pp.

Nagel P. and Gray T. (2012). Is the EU's Fisheries Partnership Agreement (FPA) with Mauritania a genuine partnership or exploitation by the EU? *Ocean & Coastal Management*, 56: 26-34.

Redfern J.V., Moore T.J., Fiedler P.C., de Vos A., Brownell R.L., Forney K.A., *et al.* (2017). Predicting cetacean distributions in data-poor marine ecosystems. *Diversity and Distributions*, 23(4): 394-408.

Santos M.B. and Pierce G.J. (2015). Marine mammals and good environmental status: science, policy and society; challenges and opportunities. *Hydrobiologia* 750: 13-41.

Tepsich P., Rosso M., Halpin P.N. and Moulins A. (2014). Habitat preferences of two deep-diving cetacean species in the northern Ligurian Sea. *Marine Ecology Progress Series*, 508: 247-260.

UNEP-WCMC, IUCN & NGS (2018). *Protected Planet Report 2018*. UNEP-WCMC, IUCN and NGS: Cambridge UK; Gland, Switzerland; and Washington, D.C., USA.

Valente R., Correia A.M., Gil Á., González L.G. and Sousa-Pinto I. (2019). Baleen whales in Macaronesia: occurrence patterns revealed through a bibliographic review. *Mammal Review*, 49(2): 129-151.

ATTACHMENTS.

Chapter I

Attachment 1. Scientific outputs of the PhD.

Scientific papers

Correia A., Tepsich P., Rosso M., Caldeira R. and Pinto I. (2015). Cetacean occurrence and spatial distribution: habitat modelling for offshore waters in the Portuguese EEZ (NE Atlantic). *Journal of Marine Systems*, 143: 73-85. doi: 10.1016/j.jmarsys.2014.10.016.

Valente R., **Correia A.M.**, Gil A., Gonzalez-Garcia L. and Sousa-Pinto I. (2019). Baleen whales in Macaronesia: insights on occurrence patterns through a bibliographic review. *Mammal Review*, 49(2): 129-151. doi: 10.1111/mam.12148.

Gil A., **Correia A.M.** and Sousa-Pinto I. (2019). Records of harbour porpoise (*Phocoena phocoena*) in the mouth of the Douro River (Northern Portugal) with presence of an anomalous white individual. *Marine Biodiversity Records*, 12:1. doi: 10.1186/s41200-018-0160-3.

Correia A.M., Gil A., Valente R., Rosso M., Pierce G.J. and Sousa-Pinto I. (2019). Distribution and habitat modelling for short-beaked common dolphins (*Delphinus delphis*) in Eastern North Atlantic Ocean. *Journal of the Marine Biological Association UK*, 1-15. doi: 10.1017/S0025315419000249.

Correia A.M., Gil A., Valente R., Rosso M., Pierce G.J. and Sousa-Pinto I. (2019). A dataset of cetacean occurrences in the Eastern North Atlantic. *Scientific data*, 6: 177. doi: 10.1038/s41597-019-0187-2

Correia A.M., Gil A., Valente R., Rosso M., Pierce G.J. and Sousa-Pinto I. (in prep – to be submitted as a review). Cetacean in the north-western continental Africa: occurrence, species richness and distributional range from Morocco to Liberia.

Correia A.M., Gil A., Valente R., Rosso M., Sousa-Pinto I. and Pierce G.J. (submitted to the journal Diversity and Distributions) Analyzing distribution patterns of cetacean species at the basin scale – connecting continents with the Macaronesian archipelagos in the eastern North Atlantic.

Correia A.M., Guedes D.S., Gil A., Valente R., Rosso M., Sousa-Pinto I., Sillero N. and Pierce GJ (in prep – to be submitted as original research). Predicting cetacean distribution in the eastern north Atlantic (Iberian Peninsula and African coasts and Macaronesia) to support marine management.

Book chapters

Carvalho A., Gil A., Dinis A., **Correia A.M.**, Moura A.E., Cecchetti A., *et al.* (2017). Mamíferos marinhos. in: Bencatel J., Álvares F., Moura A.E. and Barbosa A.M. (eds.) (2017). Atlas de Mamíferos de Portugal. Universidade de Évora, Portugal.

Conference abstract

Correia A., Tepsich P., Rosso M., Caldeira R. and Pinto I. (2015). Cetacean occurrence and spatial distribution in the Portuguese Economic Exclusive Zone (NE Atlantic). Abstract book: 29th Annual Conference of the European Cetacean Society – Panel Presentation. BICREF. St. Julians, Malta.

Gil A., González L. and **Correia A.** (2015). Rutas migratorias de las yubartas en el Atlántico: ¿dónde y cuándo?. Abstract book: VII Congreso de la Sociedad Española de Cetáceos – Panel Presentation. Vigo, Spain.

Correia A., Pierce G., Rosso M. and Pinto I. (2015). Cetacean monitoring program in Macaronesia – contributing to the MSFD goals. Abstract book: ICES Annual Science Conference 2015 – Oral Presentation. Copenhagen, Denmark.

Correia A., Rosso M., Pierce G. and Pinto I. (2015). Evidences of hotspots for beaked whales in the Canary Basin, Atlantic Ocean. Abstract book: 2015 World Whale Conference – Oral Presentation. Azores, Portugal.

Gil A., **Correia A.**, and Pinto I. (2015). The power of citizen and science to monitor cetacean occurrence and spatial distribution in Macaronesia offshore waters. COMMOCEAN 2016 – Panel Presentation. Bruges, Brussels.

Correia A., Rosso M., Pierce G. and Pinto I. (2016). Cetaceans distribution in the Macaronesia: from Portugal to Cape Verde. Abstract book: 30th Annual Conference of the European Cetacean Society – Oral Presentation. Madeira, Portugal.

Correia A., Valente R., Gil A., Rosso M., Pierce G. and Pinto I. (2017). Five-year spatial and temporal distribution of cetaceans in the Macaronesian waters - Analyzing habitat preferences in the area. 31st Annual Conference of the European Cetacean Society – Panel Presentation. Middelfart, Denmark.

Gil A., **Correia A.**, Valente R., Rosso M. and Pinto I. (2017). Citizen science CETUS Program: an efficient tool for monitoring cetacean occurrence and spatial distribution in Macaronesia offshore waters. 31st Annual Conference of the European Cetacean Society – Panel Presentation. Middelfart, Denmark.

Valente R., **Correia A.**, Gil A., González, L. and Pinto I. (2017). Understanding Migratory Routes in the Macaronesia – Assessing Range of Physiographic Variables for Baleen Whales. 31st Annual Conference of the European Cetacean Society – Panel Presentation. Middelfart, Denmark.

Gil Á., **Correia A.M.**, Correia M. and Sousa-Pinto I. Sailing for Ocean Literacy. European Marine Science Educator Association (EMSEA) – Panel presentation. Malta, Italy.

Sambolino A., Alves F., **Correia A.**, Ferreira R., Carvalho P., Silva G. and Dinis A. (2017). Monitoring cetaceans in the Madeira Archipelago from a ferry along a fixed transect: a preliminary analysis of summer-autumn 2016. 31st Annual Conference of the European Cetacean Society – Panel Presentation. Middelfart, Denmark.

Gil A., **Correia A.** and Sousa-Pinto I. (2018). Encounters with an anomalously white harbour porpoise (Phocoena phocoena) – first record of porpoises in the mouth of Douro River (Porto, Portugal). Poster presentation. 32nd European Cetacean Society Annual Conference. La Spezia, Italy.

Silva B.S., Gil Á., **Correia A.M.** and Sousa-Pinto I (2019). The occurrence of the porpoise (Phocoena phocoena) in the mouth of the Douro river: state of conservation and anthropogenic threats. Young Investigation Conference of University of Porto (IJUP) - Panel Presentation. Porto, Portugal.

Correia A.M., Guedes D.S., Gil A., Valente R., Rosso M., Sousa-Pinto I, *et al.* (2019). Predicting cetacean distribution in the eastern north Atlantic (Iberian Peninsula and African coasts and Macaronesia) to support marine management. To be presented at the WMMC/ECS conference, Barcelona (December 2019).

Gil A., **Correia A.M.**, Valente R. and Sousa-Pinto I. (2019). Spatio-temporal distribution and habitat preferences of cetaceans in the Continental Economic Exclusive Zone of Portugal. To be presented at the WMMC/ECS conference, Barcelona (December 2019).

Valente R., **Correia A.M.**, Gil A., González L. and Sousa-Pinto I. (2019). Where and when are you going? Unravelling distribution patterns of baleen whales through a bibliographic review. To be presented at the WMMC/ECS conference, Barcelona (December 2019).

Reports

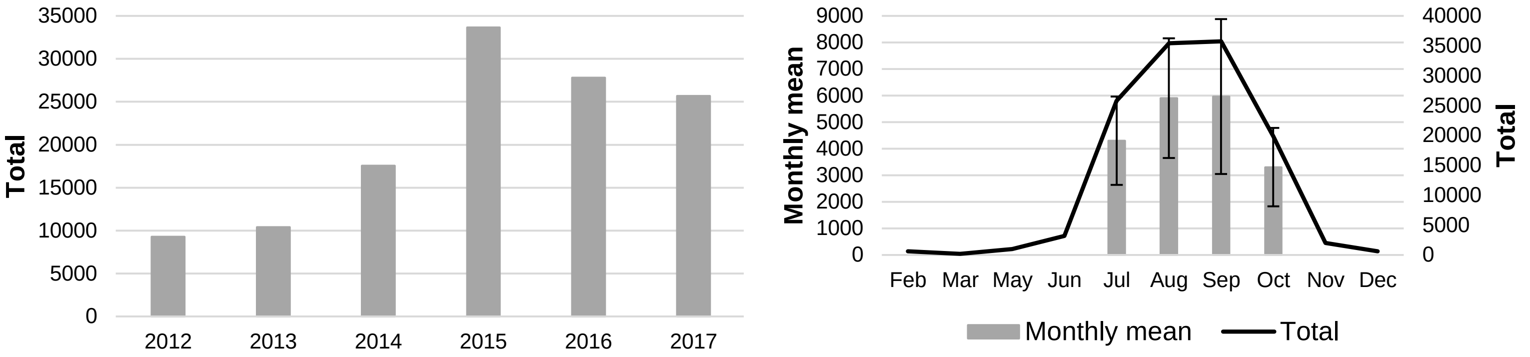
Contribution to the Report of the Working Group on Marine Mammal Ecology (WGMME) (ICES 2016).http://www.ices.dk/sites/pub/Publication%20Reports/Expert%20Group%20Report/acom/2016/WGMME/wgmme_2016.pdf

Contribution to the Intersessional Working Group on Research and Conservation Actions Undertaken in the Extended Agreement Area (ASCOBANS 2017). http://www.ascobans.org/sites/default/files/document/AC22_5.5.a_ExtensionArea_WGReport.pdf

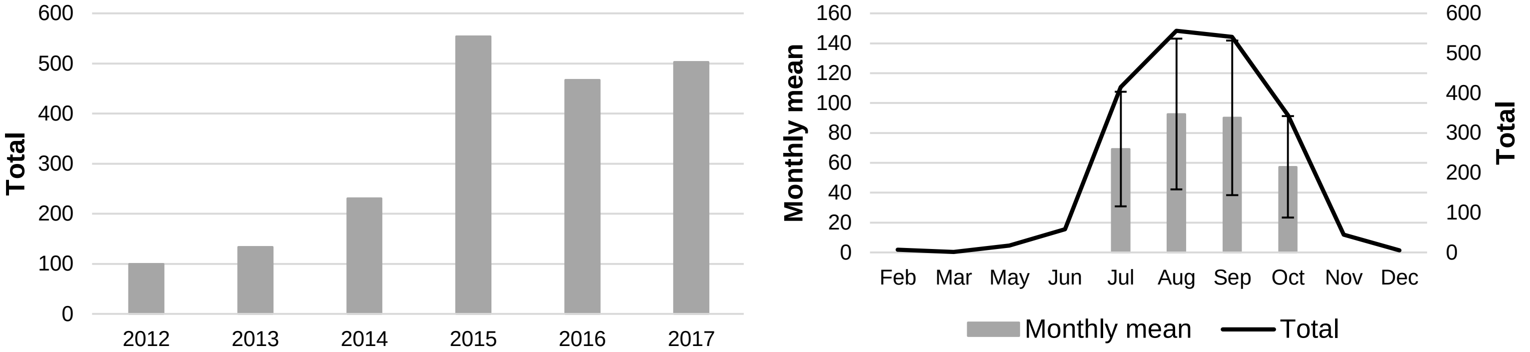
ATTACHMENTS.

Chapter III

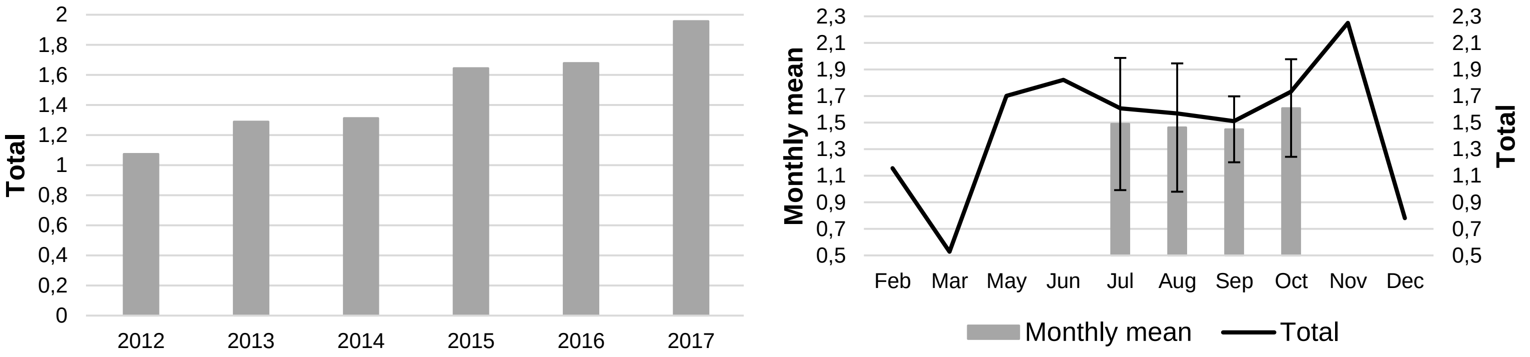
EFFORT (KM)



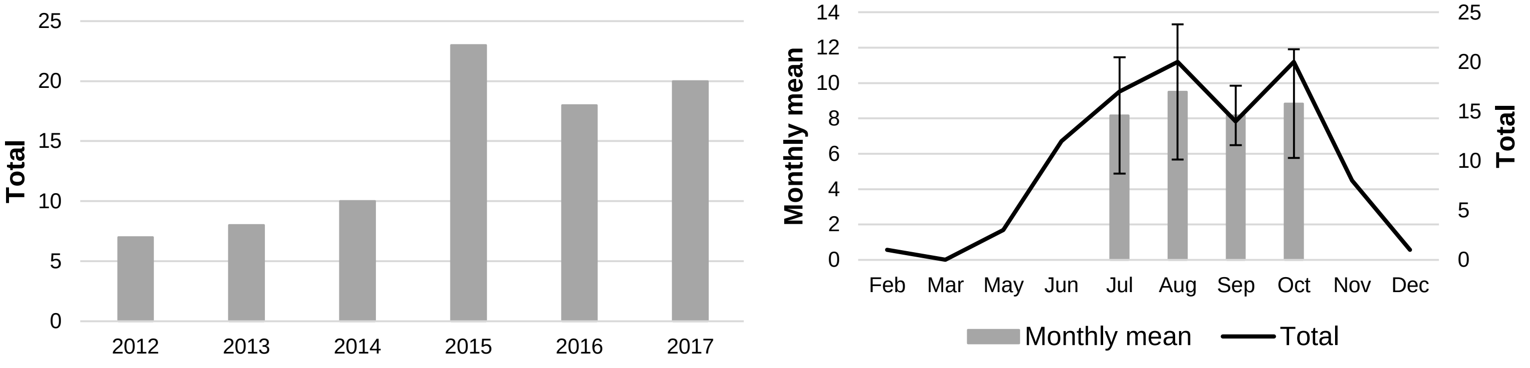
NUMBER OF SIGHTINGS



ENCOUNTER RATE (SIGHTINGS / 100 KM)



NUMBER OF SPECIES



Supplementary file 1. Temporal distribution of sampling effort, number of sightings recorded, encounter rate and number of cetacean species seen. Total values are presented by year (left panel) and month (right panel), over the whole study period. On the right panel, monthly means over the six years of surveys are presented along with the standard deviation. Monthly means are only presented for those months surveyed, at least, in three years of the study period. Total encounter rates were calculated as the number of cetacean sightings (for the all species total) divided by total survey effort (x 100).

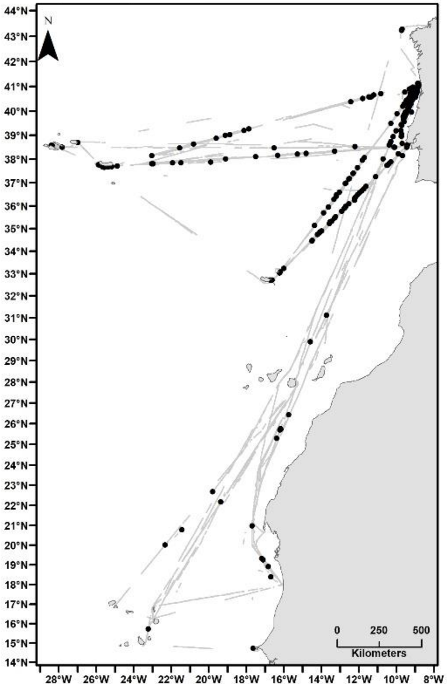
Supplementary file 2.

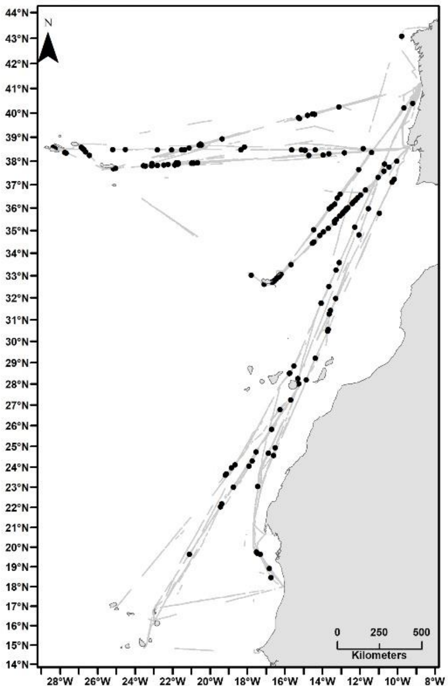
Supplementary file 2.1. Encounter rates by taxa per year and in total and relative contribution in percentage of each taxa to the total number of cetacean sightings. Encounter rates are calculated as the number of sightings of each taxon per 100 km of survey effort. Total encounter rates are the total number of sightings over the whole study period per 100 km of total survey effort, while mean is the yearly average. SD – standard deviation.

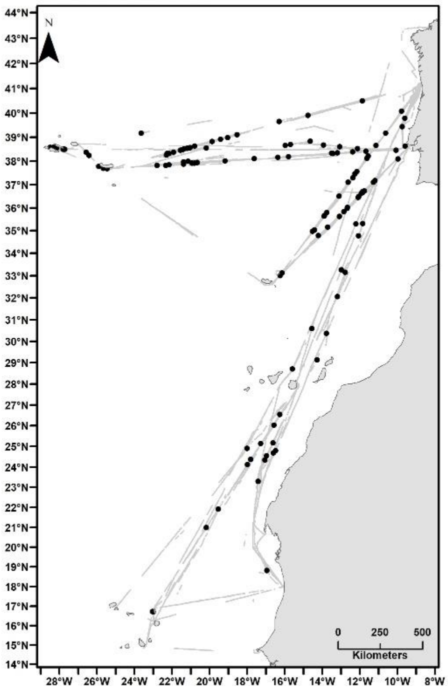
Taxa	2012	2013	2014	2015	2016	2017	Total	Mean ± SD	Relativ. contr. (%)
NI dolphin	0.226	0.356	0.330	0.392	0.468	0.595	0.427	0.394 ± 0.115	26.70
Delphinus delphis	0.150	0.212	0.199	0.163	0.216	0.272	0.206	0.202 ± 0.040	12.87
NI baleen whale	0.032	0.019	0.119	0.220	0.155	0.152	0.146	0.116 ± 0.071	9.15
Stenella frontalis	0.054	0.077	0.028	0.152	0.151	0.202	0.131	0.111 ± 0.062	8.20
NI cetacean	0.054	0.115	0.074	0.119	0.144	0.187	0.127	0.115 ± 0.044	7.94
Stenella coeruleoalba		0.029	0.210	0.062	0.129	0.074	0.093	0.084 ± 0.069	5.83
Physeter macrocephalus	0.075	0.067	0.063	0.131	0.083	0.090	0.092	0.085 ± 0.022	5.78
NI beaked whale	0.021	0.125	0.057	0.071	0.079	0.082	0.074	0.073 ± 0.031	4.63
Tursiops truncatus	0.193	0.135	0.040	0.039	0.068	0.066	0.071	0.090 ± 0.056	4.42
Balaenoptera acutorostrata	0.161	0.096	0.057	0.048	0.029	0.058	0.059	0.075 ± 0.044	3.72
Ziphius cavirostris	0.086	0.048	0.085	0.036	0.018	0.023	0.041	0.049 ± 0.027	2.56
***Globicephala* sp.**	0.021			0.033	0.050	0.047	0.031	0.025 ± 0.020	1.96
Balaenoptera physalus			0.023	0.051	0.018	0.004	0.022	0.016 ± 0.018	1.36
Stenella clymene				0.027		0.012	0.010	0.006 ± 0.010	0.60
Pseudorca crassidens				0.012	0.011	0.008	0.007	0.005 ± 0.005	0.45
Megaptera novaeangliae				0.009	0.004	0.016	0.006	0.005 ± 0.006	0.40
Stenella longirostris						0.016	0.005	0.003 ± 0.006	0.30
Mesoplodon densirostris			0.023	0.003			0.004	0.004 ± 0.008	0.25
***Kogia* sp.**		0.010				0.012	0.003	0.004 ± 0.005	0.20
Balaenoptera edeni			0.006	0.006		0.004	0.003	0.003 ± 0.003	0.20
Grampus griseus				0.006	0.004	0.004	0.003	0.002 ± 0.002	0.20
Hyperoodon ampullatus				0.003	0.007	0.004	0.003	0.002 ± 0.003	0.20
Steno bredanensis				0.006	0.007		0.003	0.002 ± 0.003	0.20
Orcinus orca				0.006	0.007		0.003	0.002 ± 0.003	0.20
Stenella attenuata					0.004	0.012	0.003	0.003 ± 0.004	0.20
Balaenoptera borealis				0.006		0.004	0.002	0.002 ± 0.002	0.15
Phocoena phocoena				0.003	0.004	0.004	0.002	0.002 ± 0.002	0.15
Peponocephala electra				0.003	0.004		0.002	0.001 ± 0.002	0.10
Balaenoptera musculus				0.006			0.002	0.001 ± 0.002	0.10
Lagenodelphis hosei				0.003			0.001	0.000 ± 0.001	0.05
Sightings with associated taxa									
Delphinus delphis / Stenella frontalis				0.009			0.002	0.001 ± 0.003	0.15
Globicephala* sp. / *Tursiops truncatus				0.003	0.004		0.002	0.001 ± 0.002	0.10
NI dolphin / NI baleen whale				0.003		0.004	0.002	0.001 ± 0.002	0.10
***Globicephala* sp. / NI dolphin**					0.007		0.002	0.001 ± 0.003	0.10
***Balaenoptera acutorostrata* / NI baleen whale**					0.004		0.001	0.001 ± 0.001	0.05
***Delphinus delphis* / NI baleen whale**					0.004		0.001	0.001 ± 0.001	0.05
NI beaked whale / NI cetacean						0.004	0.001	0.001 ± 0.001	0.05
Delphinus delphis / Stenella coeruleoalba					0.004		0.001	0.001 ± 0.001	0.05
Stenella coeruleoalba / Stenella frontalis						0.004	0.001	0.001 ± 0.001	0.05
Stenella coeruleoalba / Tursiops truncatus				0.003			0.001	0.000 ± 0.001	0.05
Delphinus delphis / Tursiops truncatus				0.003			0.001	0.000 ± 0.001	0.05
Globicephala* sp. / *Stenella clymene				0.003			0.001	0.000 ± 0.001	0.05
***Physeter macrocephalus* / NI dolphin**				0.003			0.002	0.000 ± 0.001	0.05
Total	1.075	1.289	1.314	1.646	1.680	1.958	1.599	1.494 ± 0.295	100.00
No. of species	7	8	10	23	18	20	26	14.333 ± 6.236	
Total effort (km)	9305	10392	17578	33656	27801	25695	124428	20737 ± 9028.394	

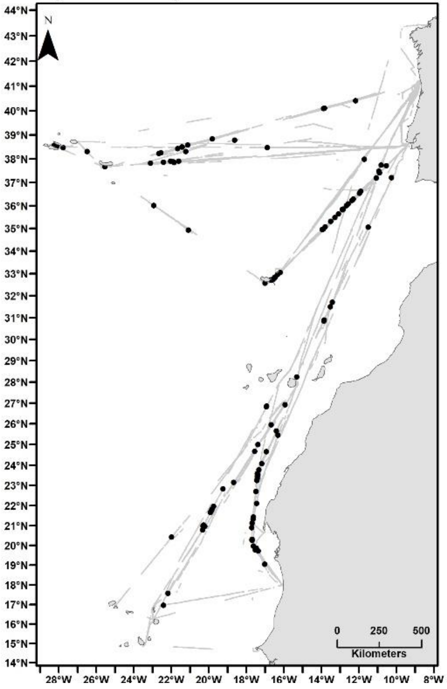
Supplementary file 2.2. Total encounter rates by taxa for each sub-region in the analysis. Encounter rates are calculated as the total number of sightings of each taxon per 100 km of survey effort over the whole study period, in each of the sub-regions (see figure 3). The highest encounter rates per taxa are marked in bold. The highest encounter rate of identified species (at least, to genus level) by sub-region is marked with ^{1}. EEZ – Exclusive Economic Zone; IP – Iberian Peninsula; Az – Azores archipelago; Mad – Madeira archipelago; NWA – Northwest Africa; CI – Canary Islands archipelago; CV – Cape Verde archipelago; IW – International waters.

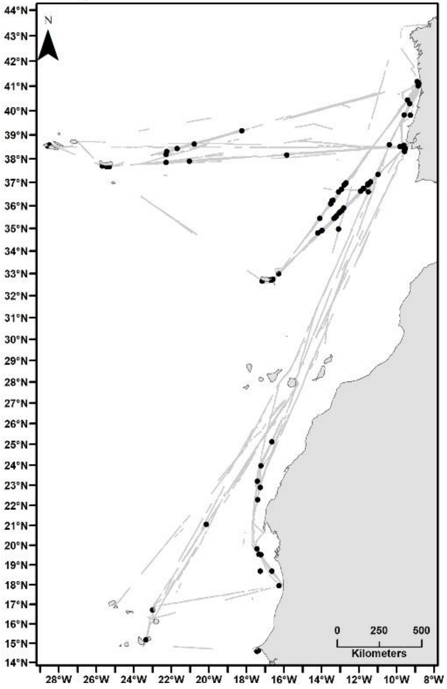
Taxa	IP EEZ	AZ EEZ	MAD EEZ	NWA EEZ	CI EEZ	CV EEZ	IW
NI dolphin	0.394	**0.796**	0.261	0.553	0.356	0.646	0.331
Delphinus delphis	**0.360**^{1}	0.352^{1}	0.085	0.077	0.034	0.040	0.089
NI baleen whale	0.202	**0.251**	0.074	0.091	0.136		0.110
Stenella frontalis	0.069	0.268	0.176^{1}	0.133	**0.271**^{1}	0.020	0.131^{1}
NI cetacean	0.109	0.168	0.085	0.224	0.119	**0.323**	0.076
Stenella coeruleoalba	0.076	**0.260**	0.059	0.105	0.085	0.020	0.081
Physeter macrocephalus	0.038	0.159	0.106	**0.273**^{1}	0.136	0.040	0.042
NI beaked whale	0.076		0.080	**0.091**	0.085	0.040	0.051
Tursiops truncatus	0.067	**0.193**	0.096	0.091		0.020	0.059
Balaenoptera acutorostrata	0.078	0.042	**0.090**	0.014	0.034	0.040	0.047
Ziphius cavirostris	0.049	0.042	0.032	0.014	0.017	0.040	**0.055**
***Globicephala* sp.**	0.016		0.037	0.119	0.034	**0.121**^{1}	
Balaenoptera physalus	0.033	**0.042**	0.005	0.028		0.020	0.004
Stenella clymene		0.008		**0.049**	0.017		0.013
Pseudorca crassidens	0.009	0.008		0.007	**0.017**	0.020	0.004
Megaptera novaeangliae		0.017		**0.035**			0.004
Stenella longirostris	0.004				**0.034**		0.008
Mesoplodon densirostris	**0.009**		0.005				
***Kogia* sp.**	0.002	**0.008**		0.007			0.004
Balaenoptera edeni	0.002		0.005	0.007	**0.017**		
Grampus griseus	0.002			**0.007**		0.020	0.004
Hyperoodon ampullatus	0.002		0.011		**0.017**		
Steno bredanensis				0.014		**0.040**	
Orcinus orca				**0.021**		0.020	
Stenella attenuata		0.008		**0.014**			0.004
Balaenoptera borealis	**0.004**						**0.004**
Phocoena phocoena	0.004					**0.020**	
Peponocephala electra				**0.014**			
Balaenoptera musculus	0.002			**0.007**			
Lagenodelphis hosei				**0.007**			
Sightings with associated taxa							
Delphinus delphis / Stenella frontalis	0.002			**0.014**			
Globicephala* sp. / *Tursiops truncatus		**0.008**		0.007			
NI dolphin / NI baleen whale	0.002				**0.017**		
***Globicephala* sp. / NI dolphin**	0.002					**0.020**	
***Balaenoptera acutorostrata* / NI baleen whale**					**0.017**		
***Delphinus delphis* / NI baleen whale**	**0.002**						
NI beaked whale / NI cetacean				**0.007**			
Delphinus delphis / Stenella coeruleoalba	**0.002**						
Stenella coeruleoalba / Stenella frontalis		**0.008**					
Stenella coeruleoalba / Tursiops truncatus						**0.020**	
Delphinus delphis / Tursiops truncatus	**0.002**						
Globicephala* sp. / *Stenella clymene				**0.007**			
***Physeter macrocephalus* / NI dolphin**				**0.007**			
Total	1.622	2.639	1.208	2.045	1.440	1.533	1.123
No. of species	19	14	12	21	12	14	16
Total effort (km)	44955	11938	18794	14282	5901	4957	23600

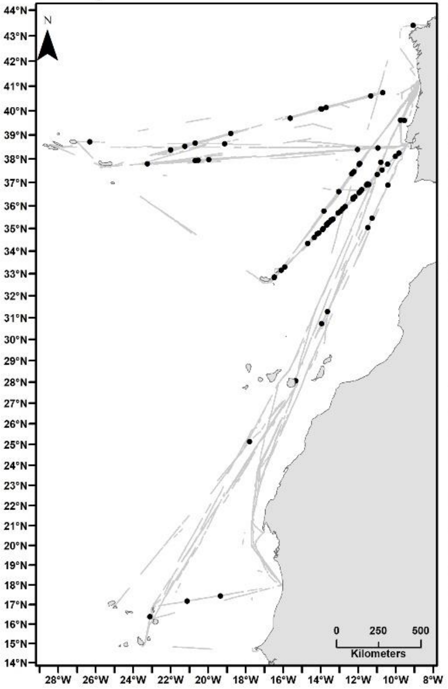

Delphinus delphis

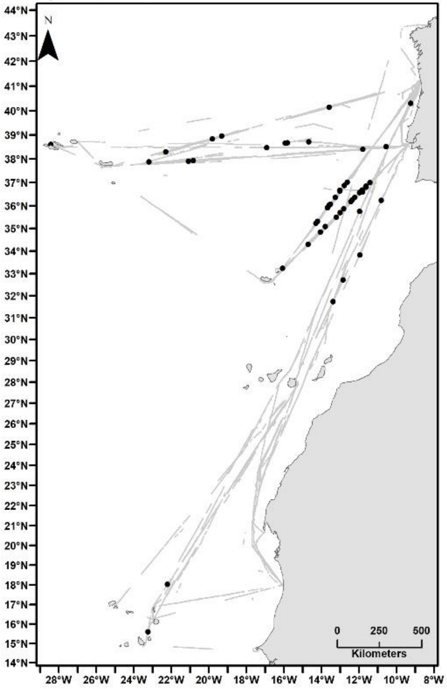

Stenella frontalis

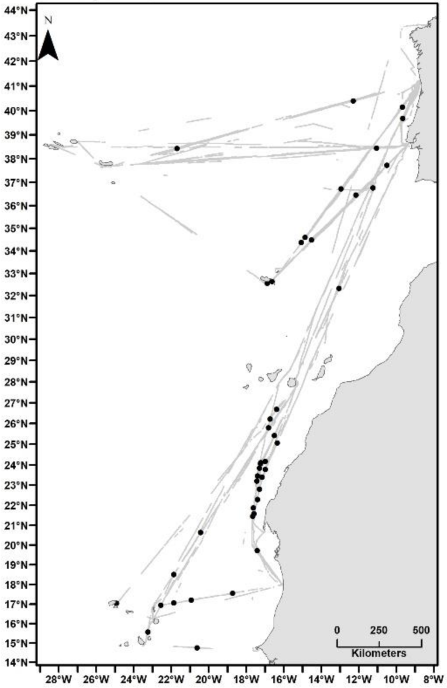

Stenella coeruleoalba


Physeter macrocephalus


Tursiops truncatus


Balaenoptera acutorostrata


Ziphius cavirostris


Globicephala sp.

Supplementary file 3. Spatial distribution of recorded sightings for the eight most frequently sighted species. Those occurrences records where a species was associated with other taxa were included in the analysis (i.e., as occurrence records for all species in the mixed group).

Supplementary file 4.

Supplementary file 4.1. Statistics for the depth at the location of the occurrences for the most frequently sighted species and for the surveyed area. Those occurrences records where a species was associated with other taxa were included in the analysis (i.e., as occurrence records for all species in the mixed group).

	Depth (m)						
Taxa	**No. sightings**	**Min**	**Q1**	**Median**	**Q3**	**Max**	**Mean ± SD**
Delphinus delphis	262	13	113	330	3735	5455	1698.55 ± 2003.90
Stenella frontalis	167	100	2358.5	3651	4714	5490	3406.69 ± 1491.52
Stenella coeruleoalba	119	69	2845	4313	4849	5610	3644.64 ± 1602.06
Physeter macrocephalus	116	89	1298.25	3262	4111	5715	2888.73 ± 1650. 82
Tursiops truncatus	92	33	275.25	1377	4711.75	5215	2278.40 ± 2093.09
Balaenoptera acutorostrata	75	180	3465.5	4391	4890.5	5364	4001.68 ± 1212.01
Ziphius cavirostris	51	111	3796	4560	4915	5445	4234.00 ± 1152.05
***Globicephala* sp.**	44	414	1168	1748.5	3534.75	5290	2312.14 ± 1436.20
Surveyed area	No. points: 23430	0	2783	4140.5	4868	5835	3526.91 ± 1682.91

Supplementary file 4.2. Statistics for distance to coast at the location of the occurrences for the most frequently sighted species and for the surveyed area. Those occurrences records where a species was associated with other taxa were included in the analysis (i.e., as occurrence records for all species in the mixed group).

	Distance to coast (km)						
Taxa	**No. sightings**	**Min**	**Q1**	**Median**	**Q3**	**Max**	**Mean ± SD**
Delphinus delphis	262	1	25.25	56	267	835	149.16 ± 186.89
Stenella frontalis	167	3	69.5	276	440.5	779	270.59 ± 201.34
Stenella coeruleoalba	119	3	144.5	329	435.5	766	315.34 ± 199.19
Physeter macrocephalus	116	4	85.25	210	360	738	227.12 ± 171.95
Tursiops truncatus	92	1	9.75	93.5	365.75	791	195.35 ± 203.78
Balaenoptera acutorostrata	75	7	208	332	437	728	316.33 ± 167.30
Ziphius cavirostris	51	7	295.5	375	477	737	375.15 ± 166.93
***Globicephala* sp.**	44	4	82.75	145	235	442	169.29 ± 114.55
Surveyed area	No. points: 23430	0	129.68	325.25	434.51	871.22	301 ± 190.25

Supplementary file 4.3. Statistics for latitude at the location of the occurrences for the most frequently sighted species and for the surveyed area. Those occurrences records where a species was associated with other taxa were included in the analysis (i.e., as occurrence records for all species in the mixed group).

	Latitude (decimal)						
Taxa	**No. sightings**	**Min**	**Q1**	**Median**	**Q3**	**Max**	**Mean ± SD**
Delphinus delphis	262	14.73	37.68	38.68	40.51	43.29	37.73 ± 4.72
Stenella frontalis	167	18.46	32.71	35.97	38.29	43.09	34.17 ± 5.42
Stenella coeruleoalba	119	16.70	35.10	37.86	38.51	40.50	35.25 ± 5.50
Physeter macrocephalus	116	16.96	23.23	32.73	37.26	40.42	30.44 ± 7.15
Tursiops truncatus	92	14.57	32.72	36.23	38.17	41.19	33.59 ± 7.10
Balaenoptera acutorostrata	75	16.37	35.18	36.64	38.04	43.41	35.76 ± 4.73
Ziphius cavirostris	51	15.59	35.73	36.58	37.92	40.31	35.87 ± 4.26
***Globicephala* sp.**	44	14.74	21.54	24.13	34.41	40.40	26.88 ± 7.79
Surveyed area	No. points: 23430	14.56	34.71	36.47	38.47	43.44	34.43 ± 6.42

Supplementary file 4.4. Statistics for longitude at the location of the occurrences for the most frequently sighted species and for the surveyed area. Those occurrences records where a species was associated with other taxa were included in the analysis (i.e., as occurrence records for all species in the mixed group).

		Longitude (decimal)					
Taxa	**No. sightings**	**Min**	**Q1**	**Median**	**Q3**	**Max**	**Mean ± SD**
Delphinus delphis	262	-28.41	-16.37	-10.46	-9.30	-8.79	-13.70 ± 5.91
Stenella frontalis	167	-28.35	-19.27	-15.75	-13.32	-9.20	-16.69 ± 4.49
Stenella coeruleoalba	119	-28.49	-21.07	-16.24	-12.66	-9.59	-16.92 ± 4.99
Physeter macrocephalus	116	-28.26	-28.26	-19.86	-13.94	-10.25	-17.37 ± 3.87
Tursiops truncatus	92	-28.59	-17.33	-14.14	-11.76	-8.82	-15.49 ± 5.00
Balaenoptera acutorostrata	75	-26.32	-15.48	-13.43	-11.84	-9.07	-14.22 ± 3.74
Ziphius cavirostris	51	-28.42	-15.27	-13.26	-11.97	-9.24	-14.63 ± 3.98
***Globicephala* sp.**	44	-24.91	-17.63	-17.08	-15.02	-9.66	-16.81 ± 3.60
Surveyed area	No. points: 23861	-28.62	-17.53	-13.88	-12.28	-8.33	-15.16 ± 4.21

Supplementary file 5. Kruskal-Wallis and Mann-Whitney statistical tests for comparison of the PC scores obtained from the PCA applied to the most frequently sighted species.

Supplementary file 5.1. Kruskal-Wallis and Mann-Whitney statistical tests for overall and pairwise comparison of the PC1 scores for the most frequently sighted species. Bonferroni correction was applied to the Mann-Whitney pairwise comparisons. P-values of less than 0.002 (significance level adjusted for pairwise comparisons) were regarded as statistically significant, and are marked in bold. Those occurrences records where a species was associated with other taxa were included in the analysis (i.e., as occurrence records for all species in the mixed group). df – degrees of freedom.

PC1 ~ Species | Kruskal-Wallis chi-squared = 136.78, df = 7, p < 0.001								
Mann-Whitney	***Delphinus delphis***	***Stenella frontalis***	***Stenella coeruleoalba***	***Physeter macrocephalus***	***Tursiops truncatus***	***Balaenoptera acutorostrata***	***Ziphius cavirostris***	***Globicephala* sp.**
Delphinus delphis		**<0.001**	**<0.001**	1.000	1.000	**<0.001**	**<0.001**	1.000
Stenella frontalis			1.000	0.034	0.007	0.292	0.007	**<0.001**
Stenella coeruleoalba				**<0.001**	**0.001**	1.000	1.000	**<0.001**
Physeter macrocephalus					1.000	**<0.001**	**<0.001**	0.818
Tursiops truncatus						**<0.001**	**<0.001**	1.000
Balaenoptera acutorostrata							1.000	**<0.001**
Ziphius cavirostris								**<0.001**

Supplementary file 5.2. Kruskal-Wallis and Mann-Whitney statistical tests for overall and pairwise comparison of the PC2 scores for the most frequently sighted species. Bonferroni correction was applied to the Mann-Whitney pairwise comparisons. P-values of less than 0.002 (significance level adjusted for pairwise comparisons) were regarded as statistically significant, and are marked in bold. Those occurrences records where a species was associated with other taxa were included in the analysis (i.e., as occurrence records for all species in the mixed group). df – degrees of freedom.

PC2 ~ Species | Kruskal-Wallis chi-squared = 192.36, df = 7, p < 0.001								
Mann-Whitney	***Delphinus delphis***	***Stenella frontalis***	***Stenella coeruleoalba***	***Physeter macrocephalus***	***Tursiops truncatus***	***Balaenoptera acutorostrata***	***Ziphius cavirostris***	***Globicephala* sp.**
Delphinus delphis		**<0.001**	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**0.0016**	**<0.001**
Stenella frontalis			1.000	0.008	0.477	**<0.001**	0.014	0.008
Stenella coeruleoalba				0.009	1.000	0.825	0.004	0.007
Physeter macrocephalus					**<0.001**	**<0.001**	**<0.001**	1.000
Tursiops truncatus						1.000	1.000	**0.0018**
Balaenoptera acutorostrata							1.000	**<0.001**
Ziphius cavirostris								**<0.001**

ATTACHMENTS.

Original Excel files upon request (anamafaldacorreia@gmail.com)

Chapter IV

SUPPLEMENTARY FILE 1 Table summarizing sources for cetacean records within north-western Africa.

Sources	Type of data	Study area	Period of Survey	No of species	No of records	MFS species
Baines & Reichelt 2014	OPOs (seismic surveys)	Mauritania (off Banc d'Arguin)	Nov 2012 to Jan 2013	15 (15)	236s	*Physeter macrocephalus*
Bamy *et al.* 2010	Inventory (cetacean species)	Guinea	1953-2008	9 (9)	4s + 10ns	-
Benchoucha *et al.* 2018	Stranding (*Kogia sima*)	Morocco	2015	1 (1)	1ns	-
Boisseau *et al.* 2007	Dedicated surveys (vessel) + Acoustic (*Phocoena phocoena*)	Morocco to Mauritania (Cadiz to Cap Blanc)	Apr to May 2005	1 (1)	7s + 50ns	-
Boisseau *et al.* 2010	Dedicated surveys (vessel)	Morocco to South Adriatic*	Morocco block: Apr 2005	6 (6)	16s	*Delphinus delphis*
Bowman Bishaw Goham 2003	Report (EIA)	Mauritania (off Nouakchott)	Mar 2003	9 (9)	18s	*Delphinus delphis*
Camphuysen 2000	Report (cruise)	Morocco to Senegal (Cadiz to Dakar)	Jan 2000	12 (12)	68s	*Delphinus delphis*
Camphuysen *et al.* 2012	Report (cruise)	Mauritania (Cap Blanc to Nouakchott)	Nov to Dec 2012	10 (10)	63s	*Physeter macrocephalus*
Djiba *et al.* 2015	OPOs (fishing research vessel) + short review (cetacean species)	Morocco to Guinea (Tangier to Conakry)	Surveys: Oct to Dec 2011 May to Jul 2012 May 2013	18 (18)	286s + 19ns	*Delphinus delphis*
Expósito & Qnimba 2010	Ship trip	Morocco (Almuñecar to Las Palmas)	Sep 2010	4 (3)	5s	*Stenella frontalis*
Gray & Van Waerebeek 2011	Dolphin behaviour to 3D seismic explorations (*Stenella attenuata*)	Liberia	18th March 2009	1 (1)	10s	-
Hammond & Lockyer 1988	Review (*Orcinus orca*)	Northeast Atlantic Equator to British Isles (+ Mediterranean and Baltic seas)*	1950-1988	1 (1)	37s[a] + 9ns	-
Jefferson *et al.* 1997	Review (cetacean species)	West Africa (Strait Gibraltar to Congo River + Madeira, Canaries and Cape Verde)*	-	15 (3)	3s[b] + 10ns	-
Jung *et al.* 2016	Stranding (*Balaenoptera omurai*)	Mauritania (Chott Boul)	Nov 2013	1 (1)	1ns	-
Leeney *et al.* 2016	Dedicated surveys (vessel) + review (*Tursiops truncatus* and *Sousa teuszii*)	Guinea-Bissau	Survey: Feb to Apr 2008 Review: 1761 to 2009	10 (10)	150s	*Tursiops truncatus*
Masski & de Stéphanis 2015	Strandings database	Morocco	1980 to 2009	27 (0)	Qualitative only	
Moore *et al.* 2010	Interview survey (fishing effort and by-catch)	Africa, Asia, Caribbean**	2007 to 2008	3 (0)	Qualitative only	
Mullié *et al.* 2013	Strandings	Mauritania (Coastline)	Nov 2012 Jan to Feb 2013 Apr to May 2013	5 (5)	28ns	-
Notarbartolo di Sciara *et al.* 1997	Report on birth at sea (*Pseudorca crassidens*)	Morocco (Atlantic coast)	1995	1 (1)	2s	-
Pinela *et al.* 2010	Samples collected opportunistically (stable isotope analyses)	Mauritania (Coastline)	1990 to 2008	11 (11)	231ns	-
Perrin & Van Waerebeek 2012	Review (small cetaceans)	West Africa (Morocco to South Africa + Macaronesia)*	-	25 (0)	Qualitative only	
Robards and Reeves 2011	Review (marine mammal consumption)	Worldwide	1970 to 2009	12 (0)	Qualitative only	
Robineau & Vely 1998	Beach monitoring (census) + Review (cetacean species)	Mauritania (Cap Blanc to Nouakchott)	Monitoring: Nov 1994 Jan to Feb, Jun and Nov, 1995	29 (20)	90s + 283ns[c]	*Tursiops truncatus*
Rojo-Nieto *et al.* 2011	Strandings	Morocco (Mediterranean coast)***	1991 to 2008 (plus other isolated years)	10 (10)	422ns	*Stenella coeruleoalba*
Russel *et al.* 2018	OPOs (geophysical survey)	Mauritania (Coastline)	Sep to Dec 2012	16 (16)	228s	*Globicephala macrorhynchus*
Tsai and Mead 2018	Stranding (*Caperea marginata*)	The Gambia	Jan 1995	1 (1)	1ns	-
Tudela *et al.* 2005	Fishing monitoring (by-catches)	Morocco (Mediterranean coast)	Dec 2002 to Sep 2003	2 (2)	237ns	*Stenella coeruleoalba*
Tulp & Leopold 2004	Report (census)	Mauritania	Apr 2004	6 (6)	13s	*Delphinus delphis* and *Tursiops truncatus*
Van Waerebeek *et al.* 1999	Review: strandings + captures + by-catches + sightings (*Balaenoptera acutorostrata*)	NE Atlantic (Galicia to Gabon + Macaronesia)*	1905 to 1998	1 (1)	4s[d] + 6ns	-
Van Waerebeek *et al.* 2000	Report (UNEP-CMS)	Senegal to Guinea-Bissau	Up to 1998	25 (23)	95s + 551ns	*Tursiops truncatus*
Van Waerebeek *et al.* 2013	OPOs (fishing and oceanographic research – *Megaptera novaeangliae*)	Morocco to Guinea (Tangier to Conakry)	Oct to Dec 2011	1 (1)	21s	-
Van Waerebeek *et al.* 2017	Dedicated and opportunistic beach and at-sea surveys (*Sousa teuszii*)	Guinea, Nigeria, Togo and Cameroon	Guinea: 2012	2 (2)	2s + 6ns[e]	-

Weir & Collins 2015	Review (*Sousa teuszii*)	West Africa (Strait Gibraltar to Angola)*		1 (1)	307s + 32ns	-
Weir & Pierce 2013	Review (human activities impacting cetaceans)	Mauritania to Angola*	1942 to 2005	16 (16)	264ns[f]	-
Weir *et al.* 2014	Review (*Stenella clymene*)	Mauritania to Angola*	1846 to 1995	1 (1)	17s	-
Weir 2016	Dedicated surveys (*Sousa tseuszii*)	Senegal (Saloum Delta)	Oct to Nov 2015	1 (1)	30s[g]	-

No – number; s – sightings records; ns – non-sightings records; OPOs – Observation Platforms of Opportunity; MFS – Most frequently sighted; EIA – Environmental Impact Assessment. Number of species with quantitative data within brackets. Number of species included only confirmed identifications to the species level of records positively reported within the study range (Morocco to Liberia). Number of records are sightings + non-sightings of both confirmed and unconfirmed identifications positively reported within the study range. For the MFS species only confirmed identifications to the species level of sightings records positively confirmed within the study range were considered. Whenever studies were dedicated to one single species, this is identified in the type of data.

* Only compiled records from Morocco to Liberia; ** Only Sierra Leone considered; *** Only strandings in coast of Morocco considered. [a] Three sightings correspond to groups that were sighted several times in the area; [b] One sighting corresponds to one individual, the "Senegal dolphin", sighted several times in the area; [c] There were three more non-sightings reports: 15 captures + strandings; 17 specimens; 19 captures + strandings. Captures are mostly/all cited in Weir & Pierce 2013 and the 17 specimens are under the IFAN collection (presented by Van Waerebeek *et al.* 2000). There is no confirmation on the exact number of strandings, hence, these three records were not considered from the source Robineau & Vely 1998; [d] One sighting corresponds to one individual sighted several times in the area. [e] Although Van Waerebeek et al. (2017) focus only on *Sousa teuszii*, they also mention (secondary source) a stranding of *Tursiops truncatus*, here included. [f] There was one more by-catch event of 10 animals of *Tursiops truncatus* + *Delphinus* sp., but since there was no specification of numbers of each species, these were not considered. [g] There are 3 re-sightings included.

SUPPLEMENTARY FILE 2

Table summarizing the number of sightings records with unconfirmed identifications found in the literature for each cetacean species in north-western Africa.

Species	Sightings	Non-sightings	Group size (min-max)	Unconfirmed location	No of sources
Phocoena phocoena	1	-	-	-	2 (2,13)
Delphinus capensis	-	15	-	1ns	1 (30)
Delphinus delphis	11	3	-	4ns	2 (9,30)
Delphinus delphis/Stenella coeruleoalba	-	75	-	-	1 (27)
Delphinus sp.	5	252[a]	30-85	2s (-) + 1ns	9 (7,13,17,20,23,28,30,34,35)
Delphinus sp./_Stenella coeruleoalba_	1	-	2	-	1 (7)
Globicephala macrorhynchus	-	1	-	-	1 (30)
Globicephala melas	1	-	-	-	1 (23)
Globicephala sp.	6	9	3-20	-	7 (9,10,15,17,28,30,34)
Grampus griseus	1	1	3	-	2 (9,23)
Kogia sp.	Qualitative only				1(17)
Lagenodelphis hosei	1	-	-	-	1 (13)
Lagenorhynchus sp.	-	1	-	-	1 (13)
Stenella attenuata	1	-	4	-	1 (30)
Stenella clymene	-	3	-	8s (-)	2 (30,35)
Stenella frontalis	2	-	5	-	1 (9)
Stenella sp.	2	5	1,150	-	5 (7,8,17,130,34)
Tursiops truncatus	7	1	1-1	-	2 (9,30)
Tursiops truncatus/Sousa teuszii	-	1	-	-	2 (30,34)
Balaenoptera borealis	1	1	2	-	2 (9,18)
Balaenoptera edeni	3	-	1-1	-	1 (8)
Balaenoptera edeni/borealis	62	-	1-25	-	3 (1,8,9)
Balaenoptera musculus	2	1	1,1	-	2 (8,9)
Balaenoptera edeni/borealis/musculus	5	-	1-3	-	2 (7,8)
Balaenoptera physalus	-	1	-	-	1 (30)
Megaptera novaeangliae	4	1	1-5	-	3 (9,18,31)
NI baleen whale	119	12	1-25	1ns	8 (1,2,9,15,18,23,25,30)
NI beaked whale	2	-	1,3	-	2 (8,9)
NI dolphin	171	188	1-35	4ns	10 (1,5,7,8,9,22,24,25,30,34)
NI cetacean (whale)	20	6	1-2	-	8 (1,5,6,7,8,18,30,34)
NI cetacean	18	3	1-1	-	4 (9,18,28,30)

No – number; s – sightings records; ns – non-sightings records; NI – Non-identified. Records compiled as with unconfirmed location are those probably within the study range (Morocco to Liberia) but without sufficient data for a positive confirmation, the group size (min-max) of these sightings records is within brackets. The sources reporting on the species are within brackets and identified by the numbers: : 1. Baines and Reichelt 2014; 2. Bamy et al. 2010; 3. Benchoucha et al. 2018; 4. Boisseau et al. 2007; 5. Boisseau et al. 2010; 6. Bowman Bishaw Goham 2003; 7. Camphuysen 2000; 8. Camphuysen et al. 2012; 9. Djiba et al. 2015; 10. Expósito and Qnimba 2010; 11. Gray and Van Waerebeek 2011; 12. Hammond and Lockyer 1988; 13. Jefferson et al. 1997; 14. Jung et al. 2016; 15. Leeney et al. 2016; 16. Massky and de Stéphanis 2015; 17. Moore et al. 2010; 18. Mullié et al. 2013; 19. Notarbartolo di Sciara et al. 1997; 20. Pinela et al. 2010; 21. Perrin and Van Waerebeek 2012; 22. Robards and Reeves 2011; 23. Robineau and Vely 1998; 24. Rojo-Nieto et al. 2011; 25. Russel et al. 2018; 26. Tsai and Mead 2018; 27. Tudela et al. 2005; 28. Tulp and Leopold 2004; 29. Van Waerebeek et al. 1999; 30. Van Waerebeek et al. 2000; 31. Van Waerebeek et al. 2013; 32. Van Waerebeek et al. 2017; 33. Weir and Collins 2015; 34. Weir and Pierce 2013; 35. Weir et al. 2014; 36. Weir 2016.

[a] There was one more by-catch event of 10 animals of *Tursiops truncatus* + *Delphinus* sp., but since there was no specification of numbers of each species, these were not considered.

Supplementary file 3 and Supplementary file 4 are Excel files provided upon request.

SUPPLEMENTARY FILE 5

Table with cetacean sightings and encounter rates per taxa collected during the CETUS project by Exclusive Economic Zones (EEZ).

Species	**Western Saharan EEZ**	**Mauritanian EEZ**	**Senegalese EEZ**	**TOTAL**
	Total sightings / Sightings on effort (ER)	**Total sightings / Sightings on effort (ER)**	**Total sightings / Sightings on effort (ER)**	**Total sightings / Sightings on effort (ER)**
Stenella frontalis	24 / 12 (0.18)	4 / 3 (0.09)	-	28 / 15 (0.15)
Delphinus sp.	8 / 7 (0.11)	6 / 3 (0.09)	1/1 (0.92)	15 / 11 (0.11)
Stenella coeruleoalba	13 / 8 (0.12)	1 / 1 (0.03)	-	14 / 9 (0.09)
Tursiops truncatus	5 / 2 (0.03)	4 / 4 (0.13)	2/2 (1.83)	11 / 8 (0.08)
Stenella clymene	9 / 7 (0.11)	-	-	9 / 7 (0.07)
Lagenodelphis hosei	-	-	1/1 (0.92)	1 / 1 (0.01)
Steno bredanensis	1 / 1 (0.02)	-	-	1 / 1 (0.01)
NI dolphin	52 / 37 (0.56)	22 / 15 (0.47)	-	74 / 52 (0.53)
Physeter macrocephalus	29 / 19 (0.29)	13 / 7 (0.22)	-	42 / 26 (0.26)
Globicephala sp.	12 / 10 (0.15)	3 / 2 (0.06)	-	15 / 12 (0.12)
Orcinus orca	4 / 3 (0.05)	-	-	4 / 3 (0.03)
Peponocephala electra	-	3 / 2 (0.06)	1/0 (-)	4 / 2 (0.02)
Grampus griseus	1 / 0 (-)	-	-	1 / 0 (-)
NI beaked whale	10 / 7 (0.11)	1 / 1 (0.03)	-	11 / 8 (0.08)
Megaptera novaeangliae	2 / 1 (0.02)	3 / 3 (0.09)	-	5 / 4 (0.04)
Balaenoptera acutorostrata	3 / 1 (0.02)	1 / 1 (0.03)	-	4 / 2 (0.02)
Balaenoptera physalus	2 / 2 (0.03)	-	2/2 (1.83)	4 / 4 (0.04)
Balaenoptera edeni	-	1 / 1 (0.03)	-	1 / 1 (0.01)
Balaenoptera musculus	-	1 / 1 (0.03)	-	1 / 1 (0.01)
NI baleen whale	20 / 5 (0.08)	4 / 2 (0.06)	-	24 / 7 (0.07)
NI cetacean	38 / 16 (0.24)	16 / 7 (0.22)	-	54 / 23 (0.23)
Globicephala sp. / _Tursiops truncatus_	2 / 1 (0.02)	2 / 0 (-)	-	4 / 1 (0.01)
Delphinus sp. / _Stenella frontalis_	1 / 1 (0.02)	1 / 1 (0.03)	-	2 / 2 (0.02)
Globicephala sp. /_Stenella clymene_	1 / 1 (0.02)	-	-	1 / 1 (0.01)
Physeter macrocephalus / NI dolphin	1 / 1 (0.02)	-	-	1 / 1 (0.01)
Stenella frontalis / NI dolphin	-	1 / 0 (-)	-	1 / 0 (-)
TOTAL SIGHTINGS AND ER	238 / 142 (2.16)	87 / 54 (1.71)	7 / 6 (5.50)	332 / 202 (2.05)
TOTAL EFFORT (km)	6562	3161	109	9832

ER – Encounter rate presented in sightings per 100 kilometres; NI – Non-identified.

ATTACHMENTS.

Chapter VII

Supplemental file 1. Variables tested as predictors for statistical modelling and its characteristics.

Model	Variables	Source	Reference	Product name	Name used in the analysis	Spatial resolution	Temporal resolution	Unit	Model technique
Detectability	Sea-state	Sea-surveys	-	-	sea	-	-	Douglas scale	GAM
	Wind-state	Sea-surveys	-	-	wind	-	-	Beaufort scale	GAM
	Visibility	Sea-surveys	-	-	vis	-	-	1-10 scale[a]	GAM
Spatiotemporal	Latitude	Sea-surveys (GPS)	-	-	lat	-	~10 seconds	Decimal degrees	GAM; MaxEnnt
	Day of the year	Date of survey	-	-	day	-	Daily	Day	GAM
Environmental	Depth	GEBCO	GEBCO, 2017	bathy_30arc_second	depth	30 sec	-	Meters (m)	GAM; MaxEnt
	Slope	GEBCO	GEBCO, 2017	-	slope	30 sec	-	Degrees (º)	GAM; MaxEnt
	Distance to coast	-	-	-	dist_coast	-	-	Kilometres (km)	GAM; MaxEnt
	Distance to seamounts	GEBCO	GEBCO, 2017	-	dist_sm	-	-	Kilometres (km)	GAM; MaxEnt
	Sea-surface temperature	MODIS Aqua	NASA, 2017	sst4_4_sst4	sst(4km/9km; 8d/m)	4 km / 9 km	8 day / monthly	Celsius (ºC)	GAM; MaxEnt
	Chlorophyll-a	MODIS Aqua	NASA, 2017	CHL_chlor_a	chl(4km/9km; 8d/m)	4 km / 9 km	8 day / monthly	Density (mg m^{-3})	GAM; MaxEnt
	Mean sea level anomalies	AVISO	AVISO, 2017	MSLA_h_DT_all_sat_merged_0.25	msla(8d/m)	0.25 degree	8 day / monthly	Centimetres (cm)	GAM

[a] Visibility scale: 5 - 1 to 2 km; 6 - 2 to 4 km; 7 - 4 to 10 km; 8 - 10 to 20 km; 9 - 20 to 50 km; 10 - > 50 km. Below 5 (1 km of visibility), the survey stopped (off effort).

Supplemental file 2. Best GAM model results produced for cetacean species: a) *D. delphis,* b) *S. frontalis,* c) *S. coeruleoalba,* d) *Globicephala* sp., e) *P. macrocephalus* and f) *B. acutorostrata.* res – residuals from GAM model distance to coast ~ depth. Results extracted from R studio.

***D. delphis*:**

```
Family: binomial
Link function: logit

Formula:
data$PA ~ s(day, k = 4) + s(wind, k = 4) + s(depth, k = 4) + s(slope, k
= 4) + s(res, k = 4) + s(dist_sm, k = 4) + s(sst_4km_8d, k = 4) +
s(chl_9km_m, k = 4) + s(msla_8d,k = 4)

Parametric coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -4.70645    0.09223  -51.03   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Approximate significance of smooth terms:
                edf Ref.df Chi.sq  p-value
s(day)        2.816  2.977  25.84 8.74e-06 ***
s(wind)       2.862  2.985  54.95 8.84e-12 ***
s(depth)      2.011  2.401 156.60 < 2e-16 ***
s(slope)      2.949  2.998  14.47 0.002004 **
s(res)        2.938  2.997  27.57 3.16e-06 ***
s(dist_sm)    2.815  2.975  16.59 0.002279 **
s(sst_4km_8d) 2.798  2.971  48.99 1.37e-10 ***
s(chl_9km_m)  1.871  2.053  33.10 4.52e-07 ***
s(msla_8d)    2.934  2.996  16.98 0.000538 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

R-sq.(adj) =  0.0827   Deviance explained = 20.2%
UBRE = -0.82968  Scale est. = 1         n = 19798
```

***S. frontalis*:**

```
Family: binomial
Link function: logit

Formula:
data$PA ~ s(day, k = 4) + s(lat, k = 4) + s(sea, k = 4) + s(depth, k =
4) + s(slope, k = 4) + s(dist_sm, k = 4) + s(sst_4km_8d, k = 4) +
chl_4km_m + s(msla_8d, k = 4)

Parametric coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -4.4645     0.1143  -39.07   <2e-16 ***
CHL_4km_mo   -0.5666     0.3113   -1.82   0.0687 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Approximate significance of smooth terms:
                edf Ref.df Chi.sq  p-value
s(day)        2.545  2.877 12.410  0.00910 **
s(lat)        1.036  1.070  7.086  0.00996 **
s(sea)        2.950  2.998 25.455  1.23e-05 ***
s(depth)      2.702  2.934 31.861  6.43e-07 ***
s(slope)      2.843  2.980 11.662  0.01413 *
s(dist_sm)    2.612  2.886 19.656  0.00016 ***
s(sst_4km_8d) 2.389  2.732 35.570  6.60e-08 ***
s(msla_8d)    2.828  2.978  7.800  0.03814 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-sq.(adj) =  0.0137   Deviance explained = 7.45%
UBRE = -0.84701  Scale est. = 1         n = 19589
```

***S. coeruleoalba*:**

```
Family: binomial
Link function: logit

Formula:
data$PA ~ s(day, k = 4) + s(lat, k = 4) + wind + s(vis, k = 4) + slope
+ s(dist_sm, k = 4) + s(sst_4km_m, k = 4) + chl_4km_8d + s(msla_m, k =
4)

Parametric coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -3.56027    0.26256 -13.560  < 2e-16 ***
wind        -0.43665    0.07726  -5.652 1.59e-08 ***
slope        0.04164    0.01746   2.384  0.01712 *
chl_4km_8d  -1.62099    0.62668  -2.587  0.00969 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Approximate significance of smooth terms:
               edf Ref.df Chi.sq  p-value
s(day)       2.280  2.643  5.980 0.074361 .
s(lat)       2.220  2.611  7.541 0.059456 .
s(vis)       2.974  2.999 39.950 1.16e-08 ***
s(dist_sm)   2.838  2.979 37.375 4.38e-08 ***
s(sst_4km_m) 2.415  2.769  5.660 0.080420 .
s(msla_m)    2.815  2.973 18.818 0.000298 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-sq.(adj) =  0.0177   Deviance explained = 10.1%
UBRE = -0.88912  Scale est. = 1         n = 17790
```

***T. truncatus*:**

```
Family: binomial
Link function: logit

Formula:
data$PA ~ s(lat, k = 4) + s(wind, k = 4) + s(depth, k = 4) + s(dist_sm,
k = 4) + chl_4km_m + s(msla_m, k = 4)

Parametric coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -5.7526     0.1427  -40.32   <2e-16 ***
chl_4km_m    -0.2204     0.1198   -1.84   0.0658 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Approximate significance of smooth terms:
               edf Ref.df Chi.sq  p-value
s(lat)       2.964  2.999  23.51 2.59e-05 ***
s(wind)      2.942  2.997  52.97 1.89e-11 ***
s(depth)     2.914  2.994  73.87 6.92e-16 ***
s(dist_sm)   2.918  2.995  14.55  0.00197 **
s(msla_m)    2.565  2.874  15.51  0.00117 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-sq.(adj) =  0.0157   Deviance explained = 12.1%
UBRE = -0.93665  Scale est. = 1         n = 22423
```

***Z. cavirostris*:**

```
Family: binomial
Link function: logit

Formula:
```

```
data$PA ~ s(day, k = 4) + s(lat, k = 4) + wind + s(depth, k = 4) +
sst_9km_8d + s(msla_m, k = 4)

Parametric coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -10.2708    2.4018  -4.276 1.90e-05 ***
Wind         -0.7624    0.1072  -7.115 1.12e-12 ***
sst_9km_8d    0.2599    0.1041   2.498   0.0125 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Approximate significance of smooth terms:
                edf Ref.df Chi.sq  p-value
s(day)        1.855  2.276 17.430 0.000899 ***
s(lat)        2.946  2.997 16.923 0.000721 ***
s(depth)      1.296  1.523  6.466 0.026855 *
s(msla_m)     2.724  2.945 24.884  1.6e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

R-sq.(adj) =  0.0162   Deviance explained = 13.9%
UBRE = -0.94585  Scale est. = 1         n = 21143
```

Globicephala sp.:

```
Family: binomial
Link function: logit

Formula:
data$PA ~ s(lat, k = 4) + wind + s(depth, k = 4) + s(sst_4km_m, k = 4)
+ s(chl_4km_8d, k = 4) + s(msla_8d, k = 4)

Parametric coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -6.0895     0.5535 -11.001  < 2e-16 ***
Wind         -0.7167     0.1516  -4.727 2.27e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Approximate significance of smooth terms:
                edf Ref.df Chi.sq  p-value
s(lat)        1.885  2.280 47.164 5.57e-10 ***
s(depth)      2.807  2.971 23.779 2.63e-05 ***
s(sst_4km_m)  2.817  2.966 19.655 0.000189 ***
s(chl_4km_8d) 2.845  2.976  9.405 0.032921 *
s(msla_8d)    2.618  2.896  8.780 0.034038 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

R-sq.(adj) =  0.0461   Deviance explained = 26.8%
UBRE = -0.96708  Scale est. = 1         n = 17737
```

P. macrocephalus:

```
Family: binomial
Link function: logit

Formula:
data$PA ~ s(lat, k = 4) + s(wind, k = 4) + vis + s(res, k = 4) +
s(dist_sm, k = 4) + s(sst_4km_m, k = 4) + s(chl_9km_8d, k = 4) +
s(msla_8d, k = 4)

Parametric coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -7.19381    0.77912  -9.233   <2e-16 ***
Vis          0.21747    0.09304   2.337   0.0194 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Approximate significance of smooth terms:
```

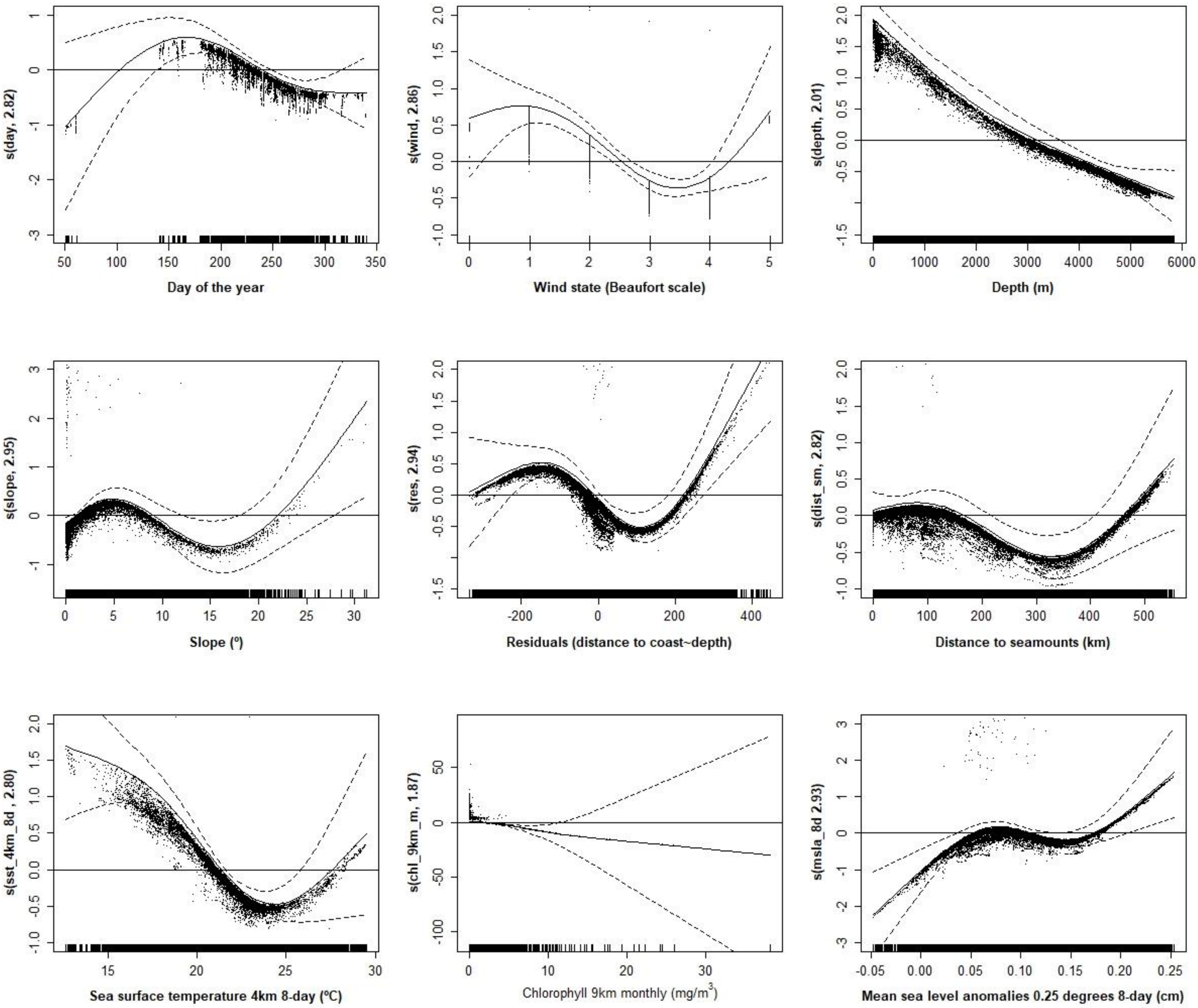
```
                edf Ref.df Chi.sq  p-value    
s(lat)        2.870  2.985  81.45  < 2e-16 ***
s(wind)       1.723  2.072  46.24 1.43e-10 ***
s(res)        1.637  2.009  14.36 0.000774 ***
s(dist_sm)    2.921  2.994  59.77 9.59e-13 ***
s(sst_4km_m)  2.725  2.948  23.59 0.000143 ***
s(chl_9km_8d) 2.940  2.996  26.42 7.83e-06 ***
s(msla_8d)    2.698  2.933  14.29 0.003125 ** 
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-sq.(adj) =  0.093   Deviance explained = 17.8%
UBRE = -0.90779  Scale est. = 1         n = 18893
```

__B. acutorostrata__**:**

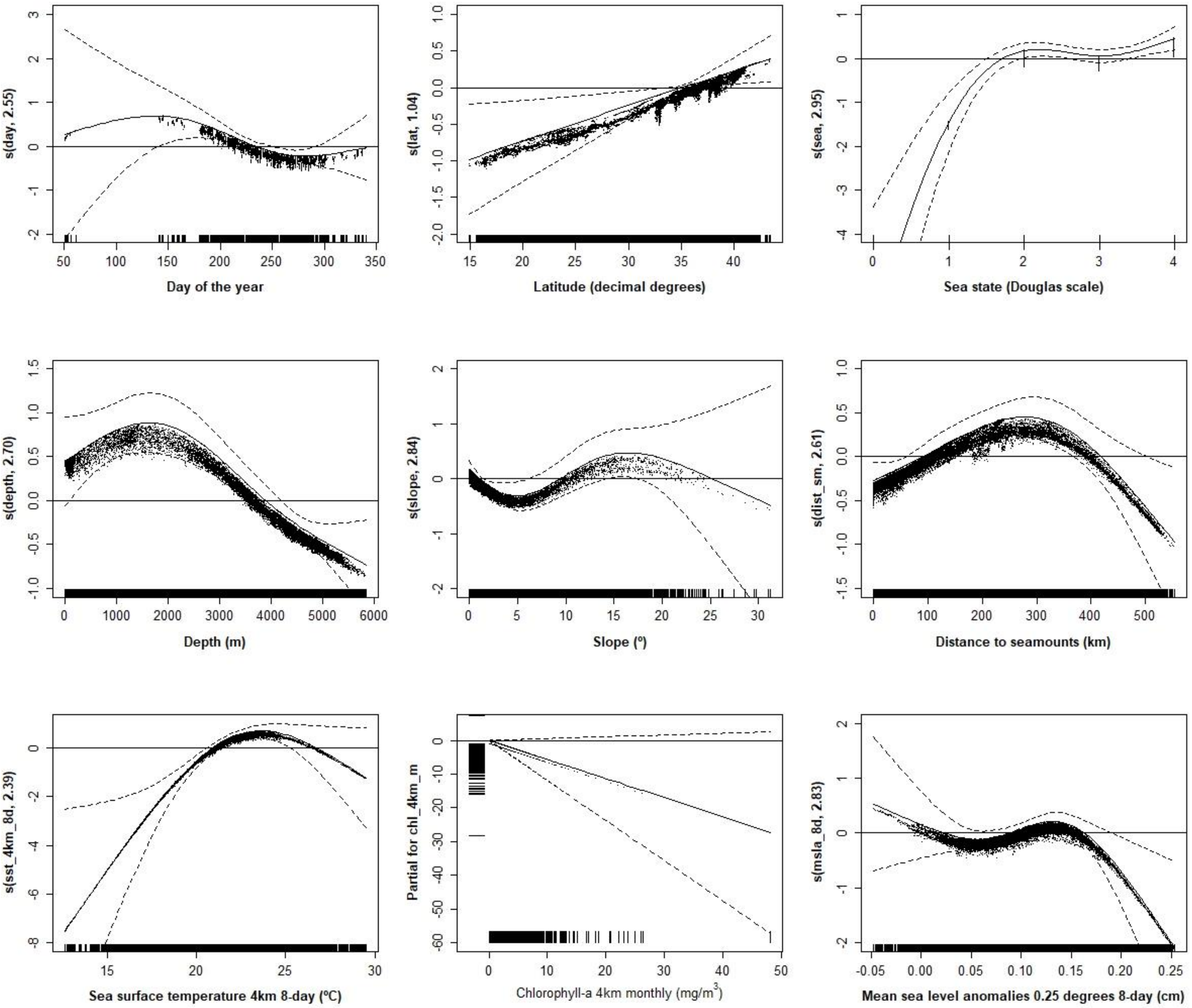
```
Family: binomial 
Link function: logit 

Formula:
data$PA ~ s(day, k = 4) + s(wind, k = 4) + s(vis, k = 4) + s(depth, k = 
4) + dist_sm + s(msla_8d, k = 4)

Parametric coefficients:
              Estimate Std. Error z value Pr(>|z|)    
(Intercept) -5.7948083  0.1923604  -30.12   <2e-16 ***
dist_sm     -0.0020430  0.0008108   -2.52   0.0117 *  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Approximate significance of smooth terms:
                edf Ref.df Chi.sq  p-value    
s(day)        2.880  2.984 11.352  0.00909 ** 
s(wind)       2.734  2.949 25.679 1.21e-05 ***
s(vis)        2.435  2.773  7.642  0.02985 *  
s(depth_m)    2.641  2.900 11.487  0.00861 ** 
s(msla_8d)    2.134  2.547  9.530  0.01425 *  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

R-sq.(adj) =  0.00623   Deviance explained = 7.48%
UBRE = -0.94701  Scale est. = 1         n = 23494
```

Supplemental file 3. GAM predicted splines of the response variable species presence as a function of the explanatory variables for the environmental model produced for cetacean species: a) *D. delphis,* b) *S. frontalis,* c) *S. coeruleoalba,* d) *Globicephala* sp., e) *P. macrocephalus* and f) *B. acutorostrata*. res – residuals from GAM model distance to coast ~ depth. The degrees of freedom are in brackets on the y-axis. Tick marks above the x-axis indicate the distribution of observations. Dashed lines delimit the 95% confidence intervals of the spline functions and dots on the graph area represent the residuals. The abbreviations used for the variables are defined in supplemental file 1.

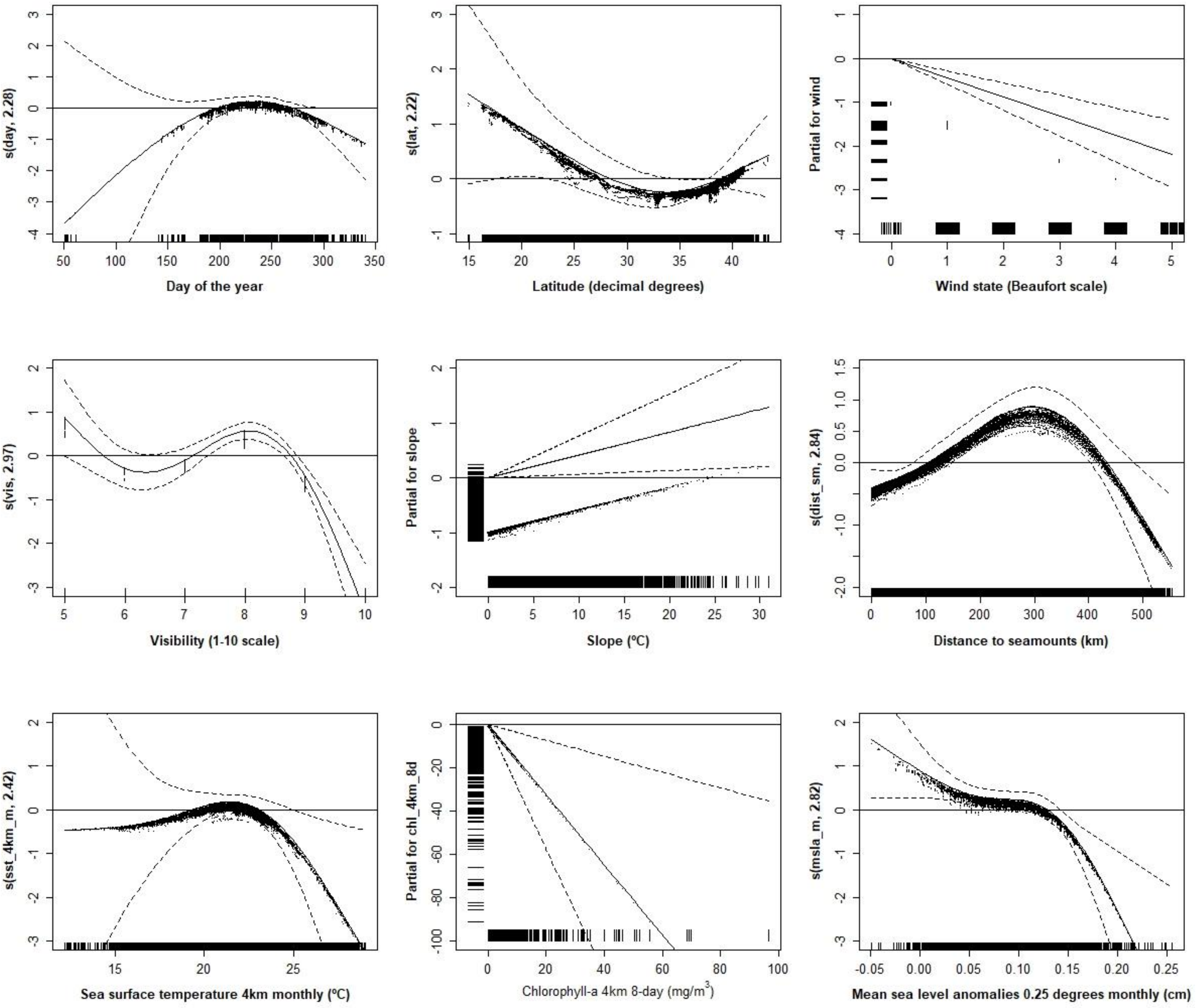
a) *D. delphis*



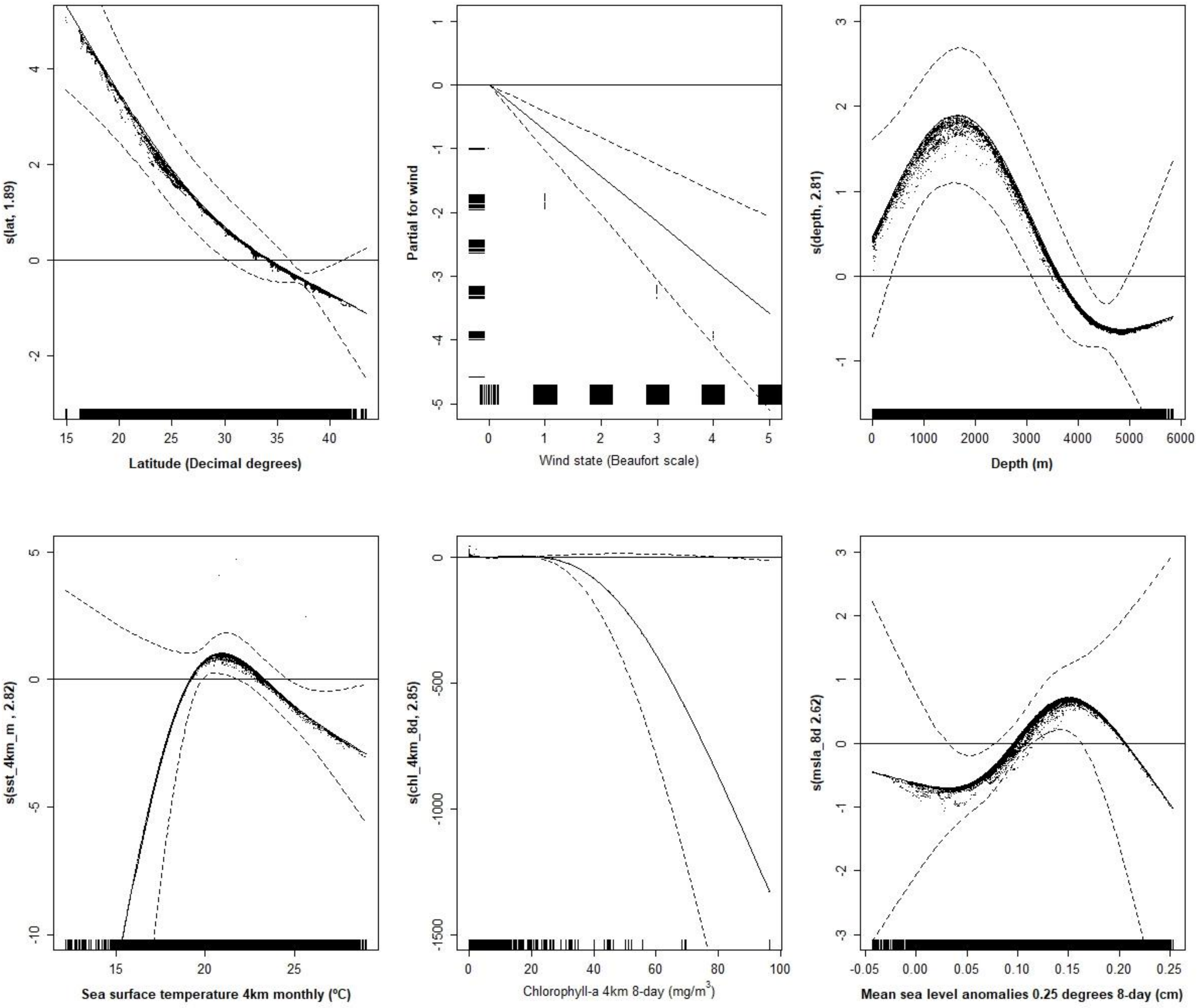
b) *S. frontalis*



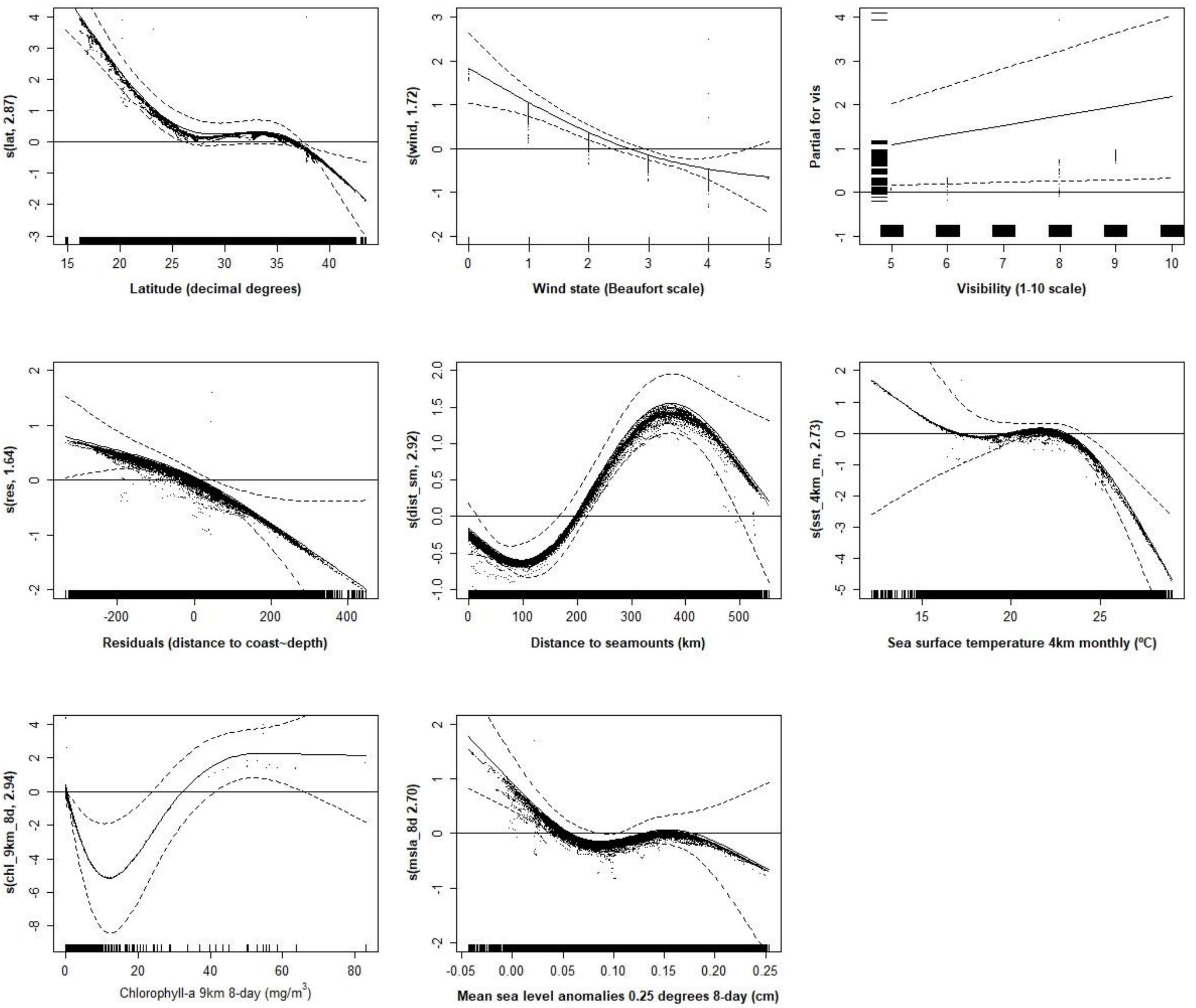
c) *S. coeruleoalba*



d) *Globicephala sp.*



e) *P. macrocephalus*



f) *B. acutorostrata*