# Distribution and Abundance Estimation of Sperm whales (*Physeter macrocephalus)* along the Hellenic Trench in Eastern Mediterranean.

**Kalliopi-Charitomeni Gkikopoulou**

**A thesis submitted to the University of St Andrews**

**for the degree of M.Phil**

**Supervised by Prof Jason Matthiopoulos**

**School of Biology**

University of
St Andrews

**Submitted:**

**September 21th 2012**

## 1. Candidate's declarations:

I, Kalliopi-Charitomeni Gkikopoulou, hereby certify that this thesis, which is approximately 18000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

I was admitted as a research student in January, 2010 and as a candidate for the degree of M.Phil in Jul, 2012; the higher study for which this is a record was carried out in the University of St Andrews between 2010 and 2012.

Date 19/09/12 signature of candidate .....

## 2. Supervisor's declaration:

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of MPHIL in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date 19/09/12 signature of supervisor .........

## 3. Permission for electronic publication:

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that my thesis will be electronically accessible for personal or research use unless exempt by award of an embargo as requested below, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. I have obtained any third-party copyright permissions that may be required in order to allow such access and migration, or have requested the appropriate embargo below.

The following is an agreed request by candidate and supervisor regarding the electronic publication of this thesis:

Embargo on both [all] of printed copy and electronic copy for the same fixed period of 3 years on the following ground(s):

publication would preclude future publication

Date 19/09/12 signature of candidate ......

signature of supervisor .........

"This thesis has been conducted in collaboration with the Pelagos Cetacean Research Institute (PCRI). The PCRI has accepted me as an intern and provided: 1) the research vessel NEREIS and the necessary scientific equipment and facilities 2) the scientific protocols and 3) their expertise and their long series of data collected within the framework of the research programme “Sperm whales of the Greek Seas”, independently of my thesis. Therefore, the copyright of the raw data of this thesis belongs to PCRI, who maintain the right of their scientific use and publication.

Part of my thesis (fieldwork and data preparation in PCRI’s headquarters) was supervised by Dr. A. Frantzis, scientific director of the PCRI."

*To the “first breath”*

*After each of our deepest dives*

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# 1. ABSTRACT

Sperm whales (*Physeter macrocephalus*) of the Hellenic Trench, Mediterranean Sea, illustrate a constant summer distribution and abundance. The sperm whale population of the Mediterranean Sea has been characterized as “Endangered” by the IUCN (2012) although areas of high occurrence should be under a wider conservation planning. Here, I modelled sperm whale distribution in the Hellenic Trench in order to quantify the distribution of the sperm whale along the Hellenic Trench. To do this a combined method of GAMs-GEEs were used to account for the autocorrelation existed in the data. Social groups and solitary or loosely aggregated males varied significantly in the habitat use within the study area, with males using habitat closer to the shore and social groups to present an affinity for higher Sea Surface Temperature (SST) and Sea Level Anomaly (SLA) values. The covariates remained in the model for the combined dataset (social groups-males) are depth, seabed steepness and distance from the shore, distance from 1km depth contour, SST and SLA. Point transects sampling was used for the abundance estimation of the summer sperm whale population from a combined acoustic and visual survey and an estimate of 27 [19.7, 32.08] individuals was derived with 95% CI. An acoustic detection function was modelled with a Generalized Linear Model (GLMs) with data derived from an experimental dataset. The detectability of sperm whales was influenced by group size, so stratification sampling was applied to take into account the bias introduced by the number of individuals in each group. An acoustic effective

range of 13 – 21 km was derived, with bigger sized groups being detected at greater distances than the smaller ones. The Hellenic Trench presents apparently an important area for the sperm whale sub-population of the Mediterranean Sea. The Hellenic Trench has been recommended to be an MPA for the protection of the sperm whale by ACCOBAMs (Agreement on the Conservation of Cetaceans of the Black Sea, Mediterranean Sea and contiguous Atlantic Area).

# 2. ACKNOWLEDGEMENTS

I would like to thank…

A random morning when I found on my table a piece of newspaper about the sperm whales of the southern Crete 12 years ago…

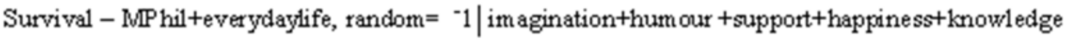
Dr Alexandros Frantzis for giving me the opportunity to come closer to one of my childhood dreams through his research program seven years ago. For trusting me his data, for his great enthusiasm and passion for what he is doing and for his strong believes to his values whatever the cost is.

Vasilis Valavanis for his assistance with the environmental data and GIS methods, and moreover for his constant optimism.

Professor Jason Matthiopoulos for his *great* supervision and friendship. His creative and continuous brain-storming of ideas and solutions to every question or concern. His support is always reassuring; for just being there and for being just tall enough. I am sure that I will find this again - not.

My lovely office-mates and friends Dr Debbie Russell – Heather Anderson - Esther Lane Jones

For making my mixed model



Survival – MPhil+everydaylife, random= ~1|imagination+humour +support+happiness+knowledge

to converge, with the random effects to play the most important role in the outcome of my model.

Dr Ricardo Antunes for inspiring me each time for sperm whale studies, especially when the motivation was running low.

Darren Kidney for his useful comments in the abundance estimation.

Dr Patric Miller and Dr Simone Panigada for their useful comments and corrections.

Trias for his blood connection and for making Scotland feel a bit like home & Pasxalia

And…

My imaginary friend “Procrastination” who was always by my side…

# 3. GENERAL INTRODUCTION

The Mediterranean Sperm whale sub-population has been characterized as "Endangered" by the International Union for Conservation of Nature (IUCN, 2012). The Pelagos Cetacean Research Institute in Greece (PCRI) has been conducting annual monitoring surveys since 1998 in the area of the Hellenic Trench (eastern Mediterranean basin), collecting data for the sperm whale population. The Hellenic Trench is a unique geomorphological structure in the eastern Mediterranean containing the deepest point of the Mediterranean, the Calypso Deep (5.267 m), an important area for sperm whale feeding and breeding. The continuously increasing number and intensity of risks (such as shipping) to whales (Panigada *et al.*, 2006) in this area increases the need for monitoring and conservation planning. Understanding the underlying processes that drive a species to exist in a specific area at a specific time along with the fluctuations in its population size are primary ecological objectives. Habitat and abundance studies have increased along with the emerging need for conservation planning and the efforts to protect biodiversity (Canadas *et al.*, 2005; Hooker *et al.*, 1999).

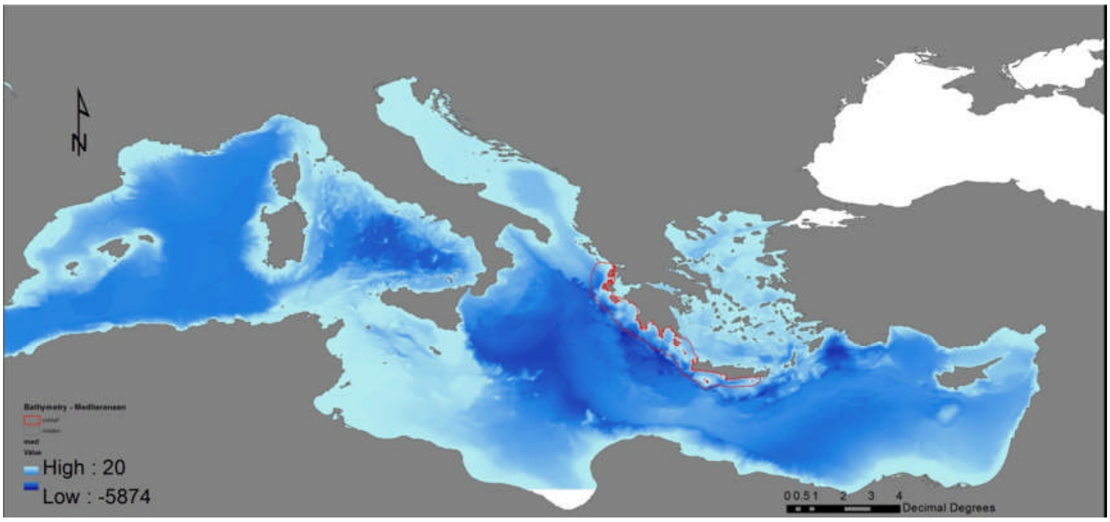


**Figure 1 Bathymetric map of Mediterranean Sea. The red polygon represents the study area along the Hellenic Trench in Eastern Mediterranean Sea.**

## 3.1 Habitat

A number of factors and their interactions determine the distribution of animals and their populations. Mostly the distribution of a species is influenced by the distribution of its prey but factors such as physiology, behaviour, predator avoidance and intra- and interspecific competition influence the distribution of animals. Furthermore, the species themselves influence and transform the environment in which they live; i.e. high densities of a species could deplete its prey resulting in a later shift in the predator’s distribution away from such areas of low prey availability. Therefore, (past and future) population dynamics should play an important role when attempting to understand the occurrence of animals in a specific area.

Habitat preference studies use environmental covariates to understand variations in the probability of occurrence or the density of animals in space, hence quantifying habitat preference.

Habitat preference can be defined as *usage over availability* (Aarts *et al*., 2008), though availability can be quantified at different ranges of spatial and temporal scales in which different scales different preference can be identified.

Because of the dynamic nature of some explanatory variables, distribution studies use proxy covariates. In Cetacean studies, it is difficult to use data regarding prey availability since they have an opportunistic diet, and their prey exists in the lower levels of the water column that are difficult to sample. So, environmental covariates such as depth, slope, distance from shore and other oceanographic data such as sea-surface temperature or chlorophyll-α are used as proxies in order to characterize the habitat of the species (Hastie *et al.*, 2005). The connection of those variables with the habitat of cetaceans is not direct but they indicate the area in the environmental space where its habitat exists.

Spatial and temporal autocorrelation, scale choice of environmental covariates, multi-collinearity between them and heterogeneous sampling effort are all aspects that could influence the findings of habitat selection studies and thus make comparative inferences problematic.

## 3.2 Abundance

Abundance is the number or the density of animals over units of time or space. Knowledge of animal abundance is needed in order abundance trends to be identified to assess conservation status and thus prioritize management actions (Buckland *et al.*, 2004).

The methodology of distance sampling has been introduced by Burnham and Anderson (1976) and since its proposal it has been widely used in abundance estimation studies for several wildlife populations (Buckland *et al.*, 2001). The success of the distance sampling methodology lies in the fact that, assuming its key assumptions hold, it provides more robust estimates, than older methodologies such as plot counts.

The methodology that has most widely been used for cetacean abundance studies is line-transect distance sampling employing whale blows, splashes or fin observations as cues. A different methodology similar to transect distance sampling is the method of point sampling is used mainly in vocal animals, such as birds, elephants, long-diving whales and primates, with most of the existing applications seen in birds (Buckland, 2006). A combination of methods could be used, dependent on the behaviour of the study animal (Hammond, 1995).

Covariate models account for heterogeneity that could exist in the detection function, the bias is then taken into consideration with the use of covariates, such as observer id, group size, type of habitat etc. The use of logistic regression in such cases can model the detection probability as the response variable and quantify it as a function of covariates that potentially could influence detectability. Logistic regression has also been used in model-based approaches, such as density surfacing modeling (DSM), for estimating abundance and is useful when survey design violates assumptions such of equal coverage probability (Hedley, 2004). Generalised Linear Models (GLMs) and Generalized Additive Models (GAMs) have been used for spatial model based methods for estimating abundance (Hedley, 2004) of species populations (Canadas and Hammond, 2006; De Segura *et al.,* 2007; Hedley, 2004). Spatial modelling for abundance estimates is useful as it allows spatially explicit abundances to be estimated along with surrounding uncertainty, though the estimates can be biased if the spatial model is not accurate.

## 3.3 Habitat and abundance for social animals

An evolutionary force promoting group formation in animals is predation avoidance. Group size is often the most obvious feature of an animal society, in the case of social animals. However defining and distinguishing between animal groups is not straightforward because groups may mean different things to different observers and to the animals themselves (Whitehead, 2008). It is widely accepted that some social animals have culture, in the sense that individuals learn *information* or *behaviour* from their conspecifics (Whitehead *et al.,* 2004). Culture can generate geographic variation in phenotypic traits such as whale vocalization (Laland, 2004).

The characteristics and advantages of social living and group formation, such as social learning, memory and skills, could influence the habitat selection and usage of social species. Social learning is thought to facilitate the transfer of critical information such as the location of food, water and other resources in foraging environments. Spatial memory could lead social animals and groups in establishing preference for specific locations.

Habitat use skills could be developed by the animals to improve their environment use and could affect habitat selection. In the presence of habitat destruction, such habitat skills could be used by the animals to avoid increasing risks (anthropogenic or natural). An example of manmade a habitat degradation factor for marine mammals is collision risk.

## 3.4 Sperm whale natural history

The sperm whale (*Physeter macrocephalus*) has attracted so much attention because of its scientific interest (large brain size, highly developed social life) and its commercial significance during the whaling years. Sperm whale social structure has likely been the driving force behind the evolution of sexual behavioural dimorphism, signaling systems and cognition (Whitehead, 2008). Male sperm whales can weigh up to 60 tons and reach 18.5m in length. Females are smaller at 15 tons and 12.5m respectively. The major morphological difference between the two sexes is in the size of their rostrum which houses the spermaceti organ, the sperm whales' most distinguishing characteristic. Sperm whales are one of the most vocal species on the planet, producing loud clicks for approximately 70% of their time (Watwood 2006; Whitehead and Weilgart, 1989), for the purposes of prey echolocation while diving and intraspecific communication while on the surface.

Sperm whales form matrilineal groups, consisting of mature female sperm whales and immature animals. The composition of these groups is usually stable but there are several observations of individuals moving from one social unit to another. Males can leave their maternal group anytime between the ages of 3 and 15 (Whitehead, 2003), to become members of bachelor groups of immature males until they reach maturity. Mature bulls are solitary except when mating. In the Galapagos Islands, the population of males rises in the period of April and May where they associate with female groups during what is believed to be the breeding season for the area (Kasuya and Miyashita 1988).

## 3.5 Sperm whales in the Mediterranean

The Mediterranean sperm whale population is believed to be isolated from the Atlantic ones. Genetic approaches (Drouot et al., 2004) have the potential to provide an insight on the degree of isolation between these populations. Specifically, mitochondrial DNA studies suggest that the population of the Mediterranean is isolated from the one in the Atlantic (Engelhaupt et al. 2009). However, it is not certain whether males from the Atlantic enter the Mediterranean basin for mating. Another way of investigating the population differentiation is to study the vocalisation patterns at different regions of the species' range. In the Mediterranean Sea the vocalization of sperm whales for communication purposes (coda type) is mainly represented by "3+1" (Frantzis *et al.*, 1999; Pavan *et al.*, 2000), and less frequently by "2 +1" (Frantzis *et al,*. 1999).

A typical deep dive in the north-western Mediterranean Sea lasts for 40-50 min for the underwater feeding period and 9 min for the surface period (Praca *et al.,* 2008) with a typical dive depth of 400-1200m (Drouot *et al.,* 2004). In the eastern basin of the Hellenic Trench, the average dive is 50.6 min and the average surface time is 9.8 minutes. The highest proportion of sperm whale diet in the Hellenic Trench (Roberts, 2003) is represented by *Histioteuthis bonnellii.*

## 3.6 Mediterranean Sea – Hellenic Trench

The Mediterranean Sea is a semi-enclosed basin divided into the Western and Eastern basin, by the Sicilian Channel, which is characterized by warm, salty and nutrient-poor waters. The surface circulation of Mediterranean includes the intense and well defined Western and Eastern Alboran gyres, the Ierapetra eddy and the Pelops anticyclone, with other more intense and variable structures in the Ionian and Levantine basin (Pujol & Larnicol, 2005). The bathymetry of the Eastern basin is highly variable. The Hellenic Trench is its most districted feature extending from the Ionian Sea up to the Island of Rhodes.

The enclosed nature of the Mediterranean Sea makes it a vital maritime highway linking ship traffic to the Atlantic through the Strait of Gibraltar, to the Black Sea through the Turkish Straits, and to the Indian Ocean through the Suez Canal. Bordered by 22 countries, it is a basin of multiple seas each with its own unique marine biodiversity and risks, while it is considered among the world's busiest waterways accounting for 15% of global shipping activity. Overall vessel activity within the Mediterranean has been rising steadily over the past 10 years (Panigada *et al.*, 2006) and is projected to increase by a further 18% over the next 10 years.

The high marine traffic is the major anthropogenic factor for collisions with marine mammals and increased noise pollution.

## 3.7 Thesis objectives

In this study I aim to quantify the distribution and habitat preferences of sperm whales, and their absolute abundance in the area of the Hellenic Trench. In more detail the questions that were addressed in this study were 1) distribution of the Sperm whales across the Hellenic Trench, 2) what are the main factors that would drive the animals preferable habitat 3) whether mature males and social groups have the same habitat preference, 4) Estimating an acoustic detection function 5) Abundance estimate for the sperm whales in the Hellenic Trench.

## 3.8 References

Aarts, G., MacKenzie, M., McConnell, B., Fedak, M. & Matthiopoulos, J. (2008) Estimating space-use and habitat preference from wildlife telemetry data. Ecography, 31, 140-160.

Buckland, S.T. (2006) Point-transect surveys for songbirds: Robust methodologies. Auk, 123, 345-357.

Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., D.L., B. & Thomas, L. (2001) Introduction to Distance Sampling: Estimating abundance of biological populations. Oxford University Press, Oxford.

Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., D.L., Borchers. & Thomas, L. (2004) Advanced Distance Sampling. Oxdord University Press, Oxford.

Canadas, A. & Hammond, P.S. (2006) Model-based abundance estimates for bottlenose dolphins off southern Spain: implications for conservation and management. J. Cetacean Res. Manage., 8, 13-27.

Canadas, A., Sagarminaga, R., De Stephanis, R., Urquiola, E. & Hammond, P.S. (2005) Habitat preference modelling as a conservation took proposals for marine protected areas for cetaceans in southern Spanish waters. Aquatic Conservation-Marine and Freshwater Ecosystems, 15, 495-521.

De Segura, A.G., Hammond, P.S., Canadas, A. & Raga, J.A. (2007) Comparing cetacean abundance estimates derived from spatial models and design-based line transect methods. Marine Ecology-Progress Series, 329, 289-299.

Drouot, V., Goold, J.C. & Gannier, A. (2004) Regional diversity in the social vocalizations of sperm whale in the Mediterranean Sea. Revue D Ecologie-La Terre Et La Vie, 59, 545-558.

Engelhaupt, D., Hoelzel, A.R., Nicholson, C., Frantzis, A., Mesnick, S., Gero, S., Whitehead, H., Rendell, L., Miller, P., De Stefanis, R., Canadas, A., Airoldi, S. & Mignucci-Giannoni, A.A. (2009) Female philopatry in coastal basins and male dispersion across the North Atlantic in a highly mobile marine species, the sperm whale (Physeter macrocephalus). Molecular Ecology, 18, 4193-4205.

Frantzis, A., Swift, R., Gillespie, D., Mhhennett, C., Gordon, J. & Gialinakis, S. (1999) Sperm whale presence off South Crete, Eastern Mediterranean Sea In: European Research on Cetaceans ,13,214-217.

Gordon, J.C.D. (1987) Behaviour and ecology of sperm whales off Sri Lanka. . Ph.D. dissertation, University of Cambridge.

Hammond, P.S. (1995) Estimating the abundance of marine mammals: A North Atlantic perspective. Whales, Seals, Fish and Man, 4, 3-12.

Hastie, G.D., Swift, R.J., Slesser, G., Thompson, P.M. & Turrell, W.R. (2005) Environmental models for predicting oceanic dolphin habitat in the Northeast Atlantic. Ices Journal of Marine Science, 62, 760-770.

Hedley, S.L. & Buckland, S.T. (2004) Spatial models for line transect sampling. Journal of Agricultural Biological and Environmental Statistics, 9, 181-199.

Hooker, S.K., Whitehead, H. & Gowans, S. (1999) Marine protected area design and the spatial and temporal distribution of cetaceans in a submarine canyon. Conservation Biology, 13, 592-602.

Kasuya, T. &  Miyashita. (1988) Distribution of sperm whale stocks in the North Pacific. The Scientific. Reports of the Whales Research Institute. 1: 109-22.

Laland, K.N. (2004) Social learning strategies. Learning & Behavior, 32, 4-14.

Panigada, S., Pesante, G., Zanardelli, M., Capoulade, F., Gannier, A. & Weinrich, M.T. (2006) Mediterranean fin whales at risk from fatal ship strikes. Marine Pollution Bulletin, 52, 1287-1298.

Pavan, G., Hayward, T.J., Borsani, J.F., Priano, M., Manghi, M., Fossati, C. & Gordon, J. (2000) Time patterns of sperm whale codas recorded in the Mediterranean Sea 1985-1996. Journal of the Acoustical Society of America, 107, 3487-3495.

Praca, E. & Gannier, A. (2008) Ecological niches of three teutophageous odontocetes in the northwestern Mediterranean Sea. Ocean Science, 4, 49-59.

Pujol, M.I. & Larnicol, G. (2005) Mediterranean sea eddy kinetic energy variability from 11 years of altimetric data. Journal of Marine Systems, 58, 121-142.

Rendell, L. & and Whitehead, H. (2001) Culture in whales and dolphins. Behavioral and Brain Sciences, 24, 309-382.

Roberts, S.M. (2003) Examination of the stomach contents from a Mediterranean sperm whale found south of Crete, Greece. Journal of the Marine Biological Association of the United Kingdom, 83, 667-670.

Robinson, A.R., Golnaraghi, M., Leslie, W.G., Artegiani, A., Hecht, A., Lazzoni, E., Michelato, A., Sansone, E., Theocharis, A. & Unluata, U. (1991) The Eastern Mediterranean General-Circulation - Features, Structure and Variability. Dynamics of Atmospheres and Oceans, 15, 215-240.

Watkins, W.A. & Schevill, W.E. (1977) Sperm Whale Codas. Journal of the Acoustical Society of America, 62, 1485-1490.

Watwood, S.L., Miller, P.J.O., Johnson, M., Madsen, P.T. & Tyack, P.L. (2006) Deep-diving foraging behaviour of sperm whales (Physeter macrocephalus). Journal of Animal Ecology, 75, 814-825.

Whitehead, H. (2003) Sperm whales: Social Evolution in the Ocean. The University of Chicago Press, Chicago and London.

Whitehead, H. (2008) Analyzing Animal Societies: Quantitative Methods for Vertebrate Social Analysis. The University of Chicago Press, Chicago.

Whitehead, H., Rendell, L., Osborne, R.W. & Wursig, B. (2004) Culture and conservation of non-humans with reference to whales and dolphins: review and new directions. Biological Conservation, 120, 427-437.

# 4. DISTRIBUTION OF SPERM WHALES ALONG THE HELLENIC TRENCH

## 4.1 Introduction

The sperm whale (*Physeter macrocephalus*) is known to inhabit the waters of the Greek seas since 1992, from opportunistic sightings, while the first dedicated visual and passive acoustic survey conducted in 1998 (Frantzis *et al.,* 1999) confirmed the existence of a resident population. The distribution of the sperm whale Mediterranean subpopulation extends along the whole Mediterranean basin with some areas of higher occurrence, such as the Balearic Islands, the Strait of Gibraltar and the Hellenic Trench, and it is believed to be divided in two further subpopulations living in the western and eastern basins (Frantzis *et al.*, submitted; Frantzis *et al.,* 2011). Closed populations like the one of the Mediterranean basin are highly vulnerable to environmental changes and human activity, and in such cases long-term monitoring studies are necessary to give an insight to the population's status. The Mediterranean sperm whale subpopulation has recently been officially characterized as "Endangered" by the IUCN Red List of Threatened Species (2012). The threats that this population is facing include: shipping lanes, by-catch on fishing activities, and disturbance from underwater noise pollution. Data relating to sperm whale spatial occurrence are needed in order to implement conservation actions for minimizing the human impact on the subpopulation. Habitat and distribution studies are necessary in order to understand the underlying environmental function that determines the occurrence of a species in an area and constitute a useful first step towards conservation of wild populations.

Two of the most important terms in ecology “niche” and “habitat” are among the most vaguely defined (Whittaker, 1973; Kearney, 2006). From the *n*-dimensional hyper-space of Hutchinson (1957) to the “mechanistic” concept of Leibold (1995), habitat has been defined *as a description of a physical place, at a particular scale of space and time where an organism either actually or potentially lives* (Kearney *et al.,* 2006), a set of environmental conditions (Wakefield *et al.,* 2009) or regions in environmental space (Aarts *et al.,* 2008). Two more unclear, but related, terms are “selection” and “preference”. Johnson (1980) uses the term “Preference” to describe the likelihood of a resource being chosen if offered on an equal basis with others. This implies that preference is the ratio of usage over the availability of a resource or a habitat. In contrast, “selection” is defined as the internal behavioural process which leads to preference (Garshelis, 2000). Habitat preference is most recently used as the ratio between the usage (time that species spend in a given habitat) over its availability (Johnson, 1980). Conclusions reached from usage-availability studies depend on what habitats are available to the animal (Johnson, 1980; Aarts *et al*., 2008), something that observers may experience in a different way to each other and to the study animals.

In Marine mammals the techniques that have been used more for habitat selection are: the Generalized Linear Model (GLM) (Gordon *et al.*, 2000; Canadas *et al.*, 2002; 2005; Panigada *et al.,* 2005), the Generalised Additive Models (GAMs) (Scott-Hayward, 2006; Aarts *et al.*, 2008; Embling *et al.*, 2010), PCA (Principal Component Analysis) (Jaquet & Whitehead, 1996; Praca *et al.*, 2009), and models that use only presence data such as ENFA (Environmental Niche Factor Analysis) and MARS (multivariate adaptive regression splines) (Praca *et al.*, 2009). The GLM uses a link function to induce linearity between response and predictor variables and is more flexible than linear regression models. The GAM is an extension of the GLM which allows locally non-linear relationships between response and explanatory variables to be specified from the data. PCA and ENFA are types of multivariate analysis that are based in ordination, rearrangement of the factors (covariates in case of PCA or factors with ecological meaning in case of ENFA) in order similarities in the produced new combined factors to be found i.e. type of habitat. Models using presence-only data such as ENFA may be useful in identifying the environmental covariates to which animals responds but provide no information on the shape of that response (Wakefield *et al.*, 2009).

The first law of geography is that "*Everything is related to everything else, but near things are more related than distant things*" (Tobler, 1970). In macroecology studies where distributional analysis use environmental covariates to explain the occurrences of animals in their study area have to deal with the existed autocorrelation between the environmental covariates. Information about Spatial and temporal auto-correlation in data used for spatial modeling are common (Aarts *et al*., 2008), and studies treating auto-correlated observations as independent can underestimate the variances of the resulting parameter estimates and model predictions. Recent analysis tools, such as Generalized Estimating Equations (GEEs) (Liang and Zeger, 1986), are used to analyze the influence of autocorrelation of observations on logistic regression models.

### 4.1.1 Sperm whale habitat and distribution.

Distributional and habitat studies for sperm whales have been conducted in areas all over the world such as is the Gulf of Mexico (Baumgartner *et al.*,2001; Jaquet,1996; O'Hern and Biggs, 2009, Davis *et al.*, 2002; Scott-Hayward, 2006), Peru and Azores The first enquiries about sperm whale habitat studies were made during the whaling period when whalers were keen to find high density whale areas to increase their income where they associated the abundance of the animals with oceanographic processes (Jaquet, 1996). Sperm whales have been associated with down-welling areas, where the nutrients from surface sink together with cold water making deeper layers more nutrient-rich (Jaquet, 1996). Also studies in the Gulf of Mexico have found links between sperm whales, especially bachelor groups, with surface primary productivity on the scales of 1 to 2 weeks and spatial scales as small as 9km$^2$, the link was supported by the correlation of encounter rate to sea surface chlorophyll (O'Hern and Biggs, 2009). Other large oceanographic processes such as cyclonic eddies have been associated with the sperm whale presence (Davis *et al.*, 2002). Sperm whales are found in open waters but also along the continental self, while they approach waters close to the coast line when the sea bed is steep, as in areas of the coast of Peru and at the volcanic islands of Azores (reviewed from Whitehead 2003). Bathymetry plays an important role in the habitat of sperm whales (Collum and Fritts, 1985; Davis *et al.*, 2002; Baumguarten *et al.*, 2001; Scott-Hayward, 2006). Studies in the Gulf of Mexico have shown that sperm whales have a preference for waters around 1500 m depth (Davis *et al.*, 2002) and there is an affinity for waters above the continental slope (Baumguarten, 2001).

### 4.1.2 Sperm whales habitat and distribution in Mediterranean

Sperm whales inhabit the Mediterranean Sea both in the eastern and the western basin. Distribution studies have been conducted in the Mediterranean Sea (Gannier *et al.*, 2002) and more specific in areas as the Ligurian Sea (Laran *et al.*, 2002; Azzelino, 2008), in the straits of Gibraltar and the Spanish Mediterranean waters (Stephanis *et al.*, 2008; Canadas *et al.*, 2002), Corsica and Balearic Islands (Pirotta *et al.*, 2011; Praca *et al.*, 2009). Most of the studies have been conducted in the western Mediterranean basin whereas this analysis is the first for the distribution in the eastern basin. As in other studies, outside the Mediterranean basin sperm whales present an affinity for the continental slope (Gannier *et al.*, 2002; Azzelino *et al.*, 2008) with bathymetry appearing to have a big influence (Pirotta *et al.*, 2009; Canadas *et al.*, 2002; Stephanis *et al.*, 2008). Steepness of the sea bed has been considered important in different studies in the Mediterranean (Praca *et al.*, 2009; Gannier *et al.*, 2002; Azellino *et al.*, 2008). Variable indicators for high productivity areas have been associated with the animal’s distribution such as sea surface temperature (SST) (Pirotta *et al.*, 2011; Laran *et al.*, 2002; Azzelino *et al.*, 2008). SST has been proposed as a key indicator of habitat suitability of sperm whales, stripped dolphins and fin whales (Azzellino *et al.*, 2008), although the connection to sperm whale ecology cannot be direct, it can be used as an indicator for areas of high productivity in different levels of the water column.

Animals may show different habitat preferences in different geographical regions. Long-lived animals may select habitats based on what is currently available or on some long-term average or median condition (Arthur *et al.*, 1996). Sperm whales live up to 70 years (Whitehead, 2003) and are highly opportunistic; in different areas they may be feeding on different kind of prey than their main prey (mesopelagic squid). Depth, an important variable for sperm whale habitat, is likely to vary between regions depending on the distribution of their food resources (Jaquet and Gendront, 2002). The area of the Hellenic Trench characterized by its deep waters and the combination of the steep slopes and permanent oceanographic features is a unique environment for the sperm whales within the Mediterranean basin. Studies that examine distributional occurrence of species should take into account the information about the geographical position of the animals in order to make it region-specific and when the environmental explanatory variables fail to explain the environmental reasons driving this distribution a more detailed design and a more sophisticated choice of explanatory variables may need to be chosen for further analysis.

Sperm whales in the Mediterranean are influenced by a number of anthropogenic threats (Lewis *et al.*, 2007). Those threats are: collisions, noise and water pollution, interaction with fisheries using gear such as driftnets (Tudela *et al,*. 2005) and plastic floating objects in the sea that end up in the digestive system of the animals and become lethal. The Mediterranean Sea is characterized by high ship traffic and is amongst the world’s busiest waterways; every year 220,000 ships larger than 100 tons cross the basin and 2000 vessels navigate these waters daily (Panigada *et al.*, 2006). In the Greek waters, the proportion of stranded sperm whales carrying propeller marks is around 70% of total standings from the period 1999-2007 (Panigada *et al*., 2007) and around 56% from 2007-2009 (Frantzis, unpublished). Monitoring whale presence, distribution and identifying high density areas is important in order to apply suggestions such as decreasing ship speed while crossing through high whale presence areas (Panigada *et al.,* 2007) or moving ferry routes outside of proposed marine protected areas, like in the area of Gibraltar (Canadas *et al,*. 2004).

### 4.1.3 Objectives

The main objective of the current analysis is to quantify the distribution of sperm whales along the Hellenic Trench and identify the environmental variables that influence it. It is hoped that the identification of distribution hot-spots of the study animal will strengthen the decisions for conservation plans and improve management efforts.

## 4.2 Materials and Methods

### 4.2.1 Study area

The study area is located in the Eastern Mediterranean Basin along the Hellenic Trench (39°N, 20°E to 34°N, 26°E) including the southeast Ionian Sea, the northern Libyan Sea, and the northwest Levantine Sea (Fig. 1). The topography is dominated by the Hellenic Trench resulting in extreme changes in bathymetry, including the deepest point in the Mediterranean (5012 m depth in the area of Pylos). Permanent oceanographic features consist of cyclonic and anticyclonic sub-basin scale gyres, the Pelops Anticyclone (PA), the West Cretan Cyclone (WC), the Ierapetra Eddie (IE), and the Rhodes Gyre (RG) to be clearly defined in an east to west direction (Robinson *et al.*,1991). These features have been observed over long periods and with consistent spatial structures. A particularly stable structure in the Mediterranean circulation is the Rhodes gyre (Pujol and Larnicol, 2005), characterized by a weak variability (Larnicol *et al.*, 2002), which is partially forced from the deep bathymetry east of Rhodes Island. Another well-defined large cyclonic feature occurs in the eastern Ionian Sea (Robinson *et al*,. 1991; Pujol and Larnicol, 2005). In the Ionian Sea annual variations are largely modulated by the Ierapetra Eddy (IE), located southeast of Crete, which reaches its maximum intensity in late summer. IE is the only clear seasonal signal in the Levantine basin with amplitude of about 20cm (Larnicol *et al*,. 2002). The topography with variable slopes and coastline is one of the factors creating this highly dynamic oceanographic environment, while the strong orography of Crete may also play an important role in the oceanography of the area.

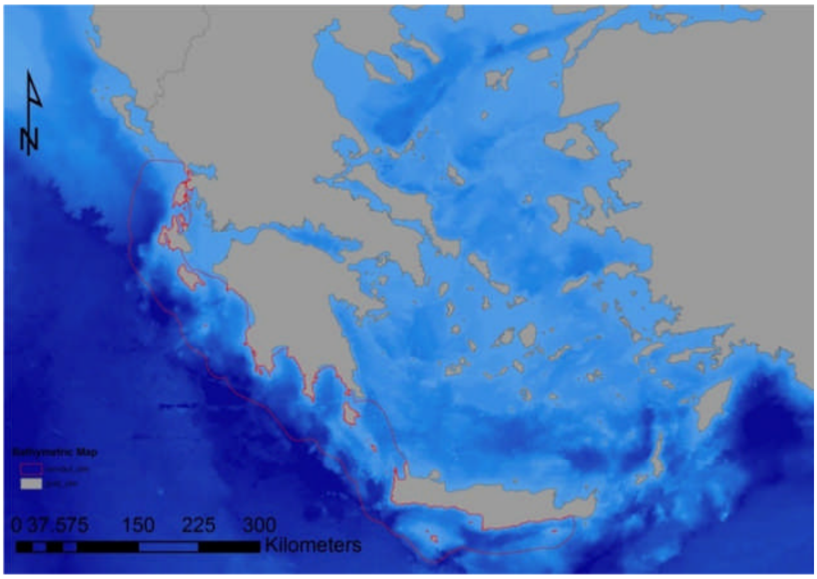


**Figure 2 Bathymetric map of the Greek waters. Bigger depths are represented with darker blue, while the study area is indicated with the red polygon.**

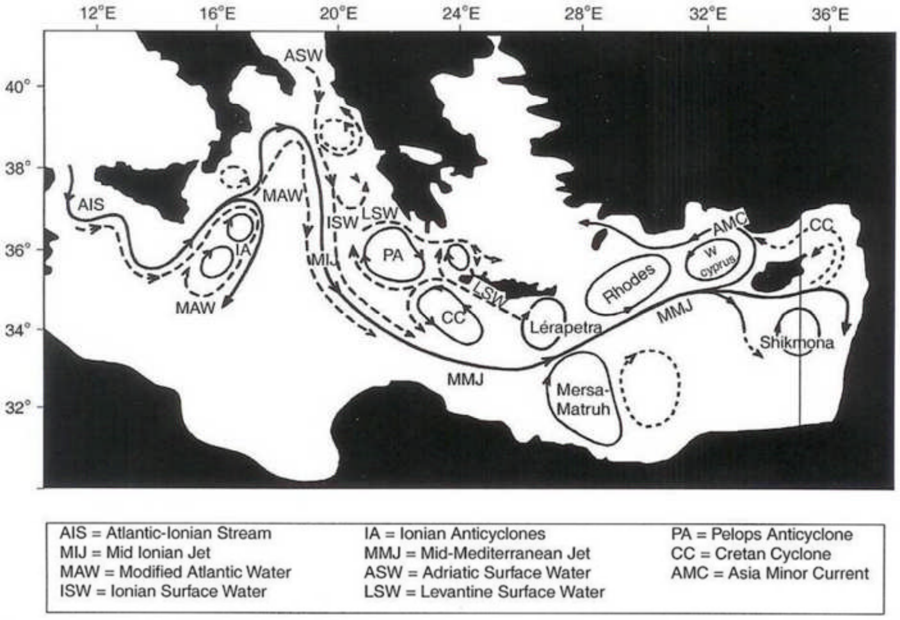


**Figure 3 Circulation features in the eastern Mediterranean (Reproduced from Robinson et al. 2001)**

### 4.2.2 Data collection

A 12 year time series of boat-based research surveys were carried out in the area of the Hellenic Trench in the summer months of 1998-2009. The duration of each survey period varied between 2 and 10 weeks (table 1).

| 1998 | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | 10 | 10 | 4 | 4 | 4 | 2 | 7 | 8 | 5 | 6 | 5 |

**Table 1 Yearly effort (measured in weeks).**

For all surveys, both acoustic and visual assessments of closed design type (searching effort stops when animals are detected) were conducted: Upon detection of animals the necessary time was taken to conduct photo–identification, photogrammetry, sloughed skin collection and underwater observation.  Transects were not homogeneously spread throughout the study area, until 2002 effort was focused in the SW Crete, whereas from 2002 onwards the survey was extended up to the Ionian Sea and transects conducted in such a way in order to allow bigger area coverage along the whole length of the study area.  A variety of vessels were used each year up to 2005, when a two engine motor vessel, Niriis, was used for the purposes of the project, for the period 2005-2009 (Table 2).

| **Year** | **Vessel Type** |
|---|---|
| 1998-1999 | Sailing boat 11m |
| 2000-2001 | Fishing boat 16m |
| 2002 | Sailing boat 11m |
| 2003 | Sailing boat |
| 2004 | Sailing boat 13.5m |
| 2005-2009 | Motor vessel-Niriis 16m |

**Table 2 Survey's vessel information**

A hydrophone array which contained two omnidirectional Benthos AQ-4 elements with 30-dB gain preamplifiers was used. The elements were mounted 3 m apart along the axis of a 10m oil-filled polyurethane tube. The frequency response of the elements was flat ± 1.5 dB and flat ± 2.0 dB for the 1 Hz to 15 kHz and 15-25 kHz bandwidths, respectively. The array was towed 100m behind the vessel or was sink into a vertical position 100m below the stern, when the vessel was not in motion. *Logger 2000*, from the International Foundation of Animal Welfare (IFAW; www/ifaw.org) was used for visual and navigation data recording. A Garmin Geographical Position System was connected to the *Logger 2000* and the vessel's positioning was being recorded regularly. An echosounder Simrad ES60 Series was used from years 2005 until 2009. Specifications of the echosounder are: a 38kHZ split-beam transducer frequency with a 10 deg circular beamwidth and a maximum pulse power input of 1500W with maximum range of 2500 m depth.

The acoustic effort began as soon as the hydrophone was deployed while an acoustic listening station conducted every 15 min, for 60 sec each. On acoustic detection of sperm whales the searching effort would stop and tracking effort would begin in order to locate the animals. Upon first visual observation of the encounter, the animal or group of animals were followed for 3 to 7 hours. While following the animals, observers scanned the horizon 360° for any distant member of the group to be detected. Along with the acoustic effort while being in searching sailing mode continuous visual scanning from two dedicated observers was conducted in order to spot other cetaceans or sperm whales that might have been on the surface. An area of 180° in front of the vessel was scanned with naked eyes and binoculars from the 2 observers.

The data used for the analysis where: i) the position of acoustic listening stations without detection for the absences and ii) the positions with visual contact with the animals for the presences.

### 4.2.3 Environmental covariates

The environmental covariates used in the analysis were divided in temporal and fixed in time.

#### Temporal

Temporal environmental covariates comprised sea surface temperature (SST) (range 19.3 to 29.75), sea surface chlorophyll-α concentration (CHL) (range 0.04 to 0.22), sea level anomaly (SLA) (range from 13.48 to 13.54), sea surface current direction (SSC) (vectors that assist to the identification of gyres) and the central points of oceanographic features such as gyres, cyclonic and anticyclonic.

Satellite imagery on SST was obtained from the Advanced Very High Resolution Radiometer (AVHRR). A time-series of weekly AVHRR SST imagery was downloaded from the Deutschen Zentrum fur Luft und Raumfahrt (DLR-GermanAcrospace Agency) online satellite data archive using DLR’s Graphical Interface to the Intelligent Satellite Data Information System (GISIS). Satellite images on CHL were obtained from the Sea-viewing Wide Field-of-view Sensor (SeaWiFS) through the National Aeronautics and Space

Administration's Distributed Active Archive Centre (NASA-DAAC) using DAAC's online data dissemination interactive hierarchical system.

SLA and SSC data were obtained from the Centre National d'Etudes Spatiales (CNES) Archiving, Validation and Interpretation of Satellite Oceanographic data archive (AVISO) that includes satellite altimetry products of merged data from Jason-1, Envisat, ERS-2, GFO and Topex/Poseidon altimetry satellite sensors.

All satellite datasets were processed under a common georeference system and were converted to regular grids using Environmental Systems Research Institute (ESRI) Arc/Info Workstation. Data were downloaded from archives in various distribution formats (e.g. TIFF, HDF, netCDF) and specific Arc/Info Arc Macro Language (AML) routines were developed for the processing of these data formats to a common Arc/Info grid format.

The extraction of the epicentre of the gyres was carried out through on-screen digitizing procedure and placed in an Arc/Info coverage of point topology. Specifically, weekly SLA and SSD grids were plotted in Arc/Info ArcEdit module and the central points of the geostrophic currents were saved in a vector-point coverage. Distance from the shore was calculated for each point, used in the analysis, with the command near, where the Euclidian distance for each point was calculated to the closest shoreline. The isobath of 1000m was created using ArcGIS latticecontour function, and the distance from the isobath was calculated for each point with the command near. Values from each fixed variable raster dataset (depth, slope, aspect) were extracted to the points dataset used for the analysis, using the command extract Values to Points in ArcMap 10.0.

## Fixed covariates

The set of time-invariant covariates comprised bathymetry (range 30 to 3215m), slope (range 0.05 to 45.7 degrees) (gradient of bathymetry), aspect (range 0.05 to 359.95 degrees), distance from the shore (range 1.3 to 30.8 km), distance from the 1000 contour (range 0 to 23km) and distance from the centre of gyres (range 7.3 to 48.3 km).

A merged dataset for the bathymetry layer was used, which comprised the bathymetry point measurements derived from the PCRI and the raster dataset downloaded from the European Marine Observations and Data Network (EMODnet, http://www.emodnet-hydrography.eku). Bathymetry data in the form of an ASCII file were downloaded and imported in GIS for the creation of a raster dataset. Central cell values were extracted from the raster dataset to point format using the command gridpoint in ArcInfo. Overlapping points between the two datasets were identified, EMODnet points deleted and replaced by the PCRI surveyed depth points. Thus, a merged point dataset was created from the two different depth sources and an interpolation procedure was followed using the topogrid command, for the creation of the merged raster dataset. The interpolation process included a boundary coverage and the coastline (bathymetry = 0). A new bathymetric raster was created with resolution 0.00416667 degrees or 250 meters projected (UTM Zone 34 N). Slope (the maximum rate of change of bathymetry in each cell), was created from the bathymetric raster dataset in ArcInfo Workstation with the command slope. Slope units were in degrees and the values ranged between 0° - 90°. Aspect, the compass orientation of the slope, was also derived from the bathymetry raster dataset using the command aspect with values ranging from 0° to 360°.

### 4.2.4 Data Analysis

Both the acoustic and tracking data were used for the analysis. Locations of acoustic stations that did not detect whales were used for the absence dataset and the sets of locations recorded while tracking the whales were used for the presence dataset.

Absence and presence data were highly serially autocorrelated both spatially and temporally due to the way they were collected. Absence data that corresponded to the acoustic listening stations were conducted every 15 min and presence data corresponding to the animal follows with sampling rate every 1 min. The acoustic detection radius (estimated in chapter 5 to be approximately 15 km) was bigger than the distance that the vessel had travelled between two successive acoustic listening stations. When each listening station was conducted, a mean vessel speed of 9 nm per hour and a travel period of 15 minutes created an overlap of the detection area that was sampled in the sequential listening stations. Similarly presence dataset points from the same encounter were auto-correlated. Point locations derived from follows of the animals were dependent upon previous points where animals were observed. As a result of the auto-correlation in the data the effective sample size is smaller than the number of points collected, and positively correlated residuals were less variable than those that would have been derived from independent data. To account for this auto-correlation, the sequences of both absence and presence temporal sequences were blocked. Our analysis accounted for the within-block autocorrelation and thus assumed independence between the different blocks.

### 4.2.4 Modelling

A modelling framework introduced by Pirotta *et al.* (2011) was followed for this analysis. A coupled use of Generalized Additive Models (GAMs) and Generalized Estimating Equations (GEEs) was used. GAMs which are a broader class of Generalized Linear Models (GLMs) are useful for interpreting non-linear ecological responses since they are able to fit non-parametric functions in estimating the relationship between response and predictor variables without imposing limitations on the form of the underlying relationships (Hastie *et al,.* 2005). Non-parametric functions allow for non-linearity in the model with the use of smoothing extensions. The non-linearity when applied to continuous variables allows the identification of non-linear covariate effects in exponential family of models and other likelihood-based regression models (Hastie and Tibshirani, 1986).

A binomial-response GAM was used with logit link function.

$$g(E[Y_i]) = g(\mu_i) = \eta_i = \log\left(\frac{p_i}{1-p_i}\right) = \beta_o + \sum_{j=1}^{p} f_j(X_j) \tag{1.1}$$

Where *i* is the model specified, $g(\mu_i)$ is the link function, $\eta_i$ is the additive predictor and the $f_i$ are smooth arbitrary functions.

While GAMs are very useful for modelling animal distribution and their interaction with the environment, they assume independence in the model errors. For model errors to be independent observations must be uncorrelated, something that is not true for our data. Generalized Estimating Equations account for existing auto-correlation in the data. GEEs were introduced from Liang and Zeger (1986) for analysing longitudinal data (repeated observations of an outcome). The existing correlation of repeated observations is being considered through a correlation structure matrix. The independence assumption between single observations is now relaxed and independence is assumed only between different blocks of observations (Liang and Zeger, 1986). The block-related variance covariance structure then takes the form

$$R = \begin{pmatrix} R_1 & 0 & 0 & 0 & 0 \\ 0 & R_2 & 0 & 0 & 0 \\ 0 & 0 & R_3 & 0 & 0 \\ 0 & 0 & 0 & R_4 & 0 \\ 0 & 0 & 0 & 0 & R_n \end{pmatrix} \tag{1.2}$$

Where $R_1 \ldots R_n$ are sub-matrices of correlations corresponding to each block, and have dimension equal to that of the size of the block the matrix represents (i.e. $n_i$ x $n_i$). The zeros represent zero correlation between blocks of data. There are different ways to define the patterns of correlation in the sub-matrices: Independence, Exchangeable, Unstructured, Auto-regressive, M-dependent and fixed.

A simple working independence (WI) structure, $R_i$, was used for model fitting, as suggested by Pan (2001) when the underlying correlation structure is unknown. The WI structure treats the within-block observations as independent, having first deflated the within- block sample size, resulting in correct estimation of parameters and standard errors.

### 4.2.5 Model Selection

There are several model selection criteria for likelihood-based models, such as Akaike's Information Criterion (AIC) and cross-validation. Quasi-likelihood under the independence model criterion (QIC), an extension of the AIC, has been introduced by Pan (2001) for model selection in non-likelihood based methods such as GEEs. Like AIC, it balances the model fit with model complexity for the most parsimonious model to be selected. The QIC allows the use of any general working correlation structure to estimate the parameters in GEE, and works well in variable selection and selecting the working correlation matrix (Pan 2001). An approximation of the QIC ($QIC_u$) (Hardin and Hilbe, 2003) was used for the identification of the variables to be retained in the final model.

$$QIC_u = -2Q + 2p \qquad (1.3)$$

where $p$ is the number of model parameters and $Q$ is the quasi-likelihood is

$$Q = y_i \log\left(\frac{\hat{p}}{1-\hat{p}}\right) + \log(1-\hat{p}) \qquad (1.4)$$

A series of all possible combinations of models were tested and the model with the lowest $QIC_u$ was selected each time. Nine variables were used, a combination of 19683 models were tested for the final model selection, with an automated code comparing the $QIC_u$ between the models. For all the models, each variable was allowed to take part as bs (basis for a polynomial spline) smoothed form, as a linear, or to be omitted from the model. The R library yags allowed us the extraction of the $QIC_u$ score for each model. The model with the lowest $QIC_u$ score was considered the best one describing the sperm whales' presence in the study area. The model selection procedure was applied to the whole set of data to determine the important environmental variables explaining the distribution of the species in the area. Latitude and Longitude were not considered for the model selection - since effort was not uniform in the study area among the different years, they were used to control for local variations in effort. A further reduction of the variables remaining in the model was possible under the Wald test statistic. The Wald test is a way of testing the significance of particular explanatory variables in a model.

Single animals and social groups were modelled separately for the second stage of the analysis, in order to identify any difference in their habitat use. Single animals were always mature males while social groups were most often female social units. Social groups which included mature males were also part of the social group dataset (visual observations from this area might suggest that mature males join a female social group for mating).

### 4.2.6 Model evaluation

Each explanatory variable and its contribution were assessed with the qqplot2 command in R. The estimated relationship between explanatory variable and the response was plotted on the link scale along with the associated uncertainty from the standard errors estimated through the GEEs. Whale presence was then explained based on the shape and significance of each variable.

Confusion matrices can be used in binary models in order to assess their performance (goodness-of-fit and predictive ability with new data). A confusion matrix is a table which reports the number of false positives, false negatives, true positives and true negatives (Figure 4). Four outcomes of a predicted model can be explained from the confusion matrix: i) Sensitivity, true positive portion, ii) Specificity, true negative portion iii) false positive portion iv) false negative portion. A model is perfect when specificity and sensitivity are 100% each. Specificity is defined as the ratio of $\frac{c}{a+c}$ and Specificity as $\frac{d}{b+d}$.

In order to build the confusion matrix a cut-off probability value must be chosen, above which prediction probability is considered to be presence.

| | | True Values | |
|---|---|---|---|
| | | Absence | Presence |
| Predicted Values | Absence | a | b |
| | Presence | c | d |

**Figure 4 Confusion Matrix a) is the number of correct predictions for absences b) is the number of incorrect predictions of absences, c) is the number of incorrect predictions of presences and d) is the number of correct predictions of presences.**

A Receiver Operating Characteristic (ROC) curve was used in order to choose a threshold value for the creation of the confusion matrix. In a ROC curve the true positive rate (Sensitivity) is plotted as a function of the false positive rate

(100-Specificity) for different cut-off points of a parameter. Each point on the ROC curve represents a sensitivity/specificity pair corresponding to a particular decision threshold. The area under the curve (AUC) measures the discriminating ability of a binary classification model. The AUC takes values from 0 to 1, the larger the AUC the higher the likelihood that an actual positive case will be assigned a higher probability of being positive than an actual negative case. In our case, the binary classification is presences/absences of whales. The threshold value for the creation of the confusion matrix was calculated from the point where the distance between the ROC curve and the 45° diagonal was maximised, calculating the distance between each point of the ROC plot and the diagonal. Then the confusion matrix could be created in order to assess the model performance.

## 4.3 Results

The final dataset used for the analysis consisted of 6759 points of which 3588 were absences and 3171 presences. Presence points were blocked into 178 cases of encounters during 12 years of summer-survey periods. The total survey length over all years was 36299 km of which 15973 km was acoustic effort and 2787 km was spent in visual contact with the animals.

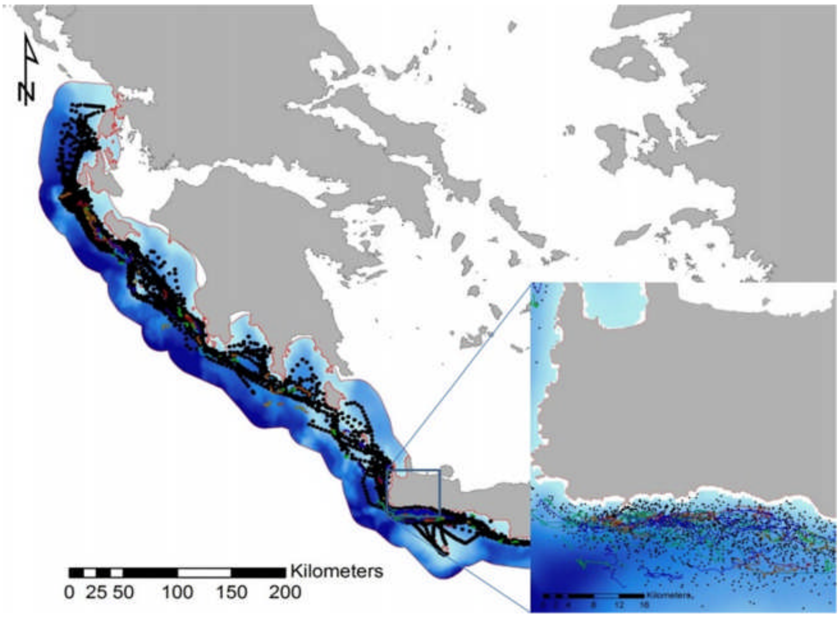


**Figure 5 Acoustic effort and follows of sperm whales denoted with points and lines correspondingly.**

### 4.3.1 Model variables for the entire data set

The model with the smallest QIC (QIC=8523) was used to model the presence of sperm whales in the study area. In addition to latitude and longitude, the following variables remained in the model: depth, aspect, slope, distance from 1000 contour, SST and SLA. The Wald’s statistic suggests that slope and aspect were not significant (*p* values of 0.136 and 0.130 respectively) so they were omitted from the final model, explaining the whales’ presence in the area. The variable SST remained with a linear form and latitude, longitude, depth, SLA and distance from 1000 contour as “bs” smoothed. The estimate of the dispersion parameter was $\varphi=1.02$, so there is no indication of over/under dispersion in the data. Depth and distance from the 1000 contour where the most important variables, SLA and SST follow. Whale presence increased with increasing depth values up to 1500m where the presence probability shows a plateau, while it decreases for bigger depths. Even if the slope remained in the model as “bs” smoothed with a small peak at 18°, the shape of the curve points out that a linear form would also be possible. Whale presence increased with increasing SST linearly, whereas SLA participates non-linearly with a peak at negative SLA values around -7 cm and another smaller one at 9 cm. Whale presence is higher at locations in proximity to the 1000 contour and away from it is decreasing while shows another smaller peak at distances 1700 m of the contour.

| **Covariates** | | Latitude | Longitude | Depth | Slope | Aspect | Distance from shore | Sst | Chl | Sla | Distance from 1kmc | Distance from gyres | Depth | Slope | Aspect | Distance from shore | Sst | Chl | Sla | Distance from 1kmc | Distance from gyres | **QIC** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Type** | | bs | | | | | | | | | | | linear | | | | | | | | | |
| **Model** | **1** | x | x | | | | | | | | | | | | | | | | | | | 9407.364 |
| | **2** | x | x | | | | | | | | | | x | x | x | x | x | x | x | x | x | 9235.358 |
| | **3** | x | x | | x | x | x | | | | | | | | | | | | x | x | | 9100.732 |
| | **4** | x | x | | | x | x | x | x | x | x | x | | | | | | | | | | 9003.56 |
| | **5** | x | x | x | x | x | x | x | x | | | | | | | | | | | | | 8993.832 |
| | **6** | x | x | | | | | | | | | | x | x | x | x | x | x | x | x | | 8990.44 |
| | **7** | x | x | x | x | x | x | | | | | | | | | | x | x | | | x | 8882.659 |
| | **8** | x | x | x | x | x | x | x | x | x | x | | | | | | | | | | | 8746.456 |
| | **9** | x | x | x | | x | x | x | | | | | | | | | | | x | x | | 8733.631 |
| | **10** | x | x | x | x | | | | | x | x | | | | x | | x | | | | | 8522.75 |

**Table 3 Different models for the full dataset which were taken under account in the model selection.**

| **Model coefficients for full dataset** | | |
|---|---|---|
| | **Estimate** | **Std err** |
| (intercept) | -11.6936 | 4.8573 |
| **bs(lat) 1** | 3.2671 | 3.6696 |
| **bs(lat) 2** | -4.8705 | 3.56 |
| **bs(lon) 1** | -6.3291 | 2.8697 |
| **bs(lon) 2** | -0.4029 | 4.2199 |
| **bs(lon) 3** | -5.5296 | 3.3196 |
| **bs(depth) 1** | 10.8005 | 3.5361 |
| **bs(depth) 2** | 2.8862 | 1.7789 |
| **sst** | 0.2431 | 0.1097 |
| **bs(sla) 1** | 7.9685 | 3.4992 |
| **bs(sla) 2** | -1.138 | 1.7932 |
| **bs(sla) 3** | 4.6317 | 2.6738 |
| **bs(sla) 4** | 0.4364 | 2.1808 |
| **bs(X1kmc) 1** | -4.2414 | 1.0374 |
| **bs(X1kmc) 2** | 1.39 | 2.099 |
| **bs(X1kmc) 3** | -1.5753 | 3.2387 |

**Table 4 Model coefficients of the final model for the full dataset.**

| Analysis of '***Wald statistic***' Table | | | |
|---|---|---|---|
| **Covariate** | **df** | **$\chi^2$** | **p(>\| Chi\|)** |
| bs(lat) | 2 | 1.6628 | 0.4354404 |
| bs(lon) | 2 | 3.8064 | 0.1490895 |
| bs(depth) | 2 | 15.2387 | 0.0004908 |
| bs(slope) | 2 | 7.6176 | 0.0221753 |
| sst | 1 | 3.9763 | 0.0461461 |
| bs(sla) | 4 | 15.7659 | 0.0033499 |
| bs(X1kmc) | 3 | 16.5231 | 0.0008857 |

**Table 5 Wald statistic table of the final model used for the full dataset.**

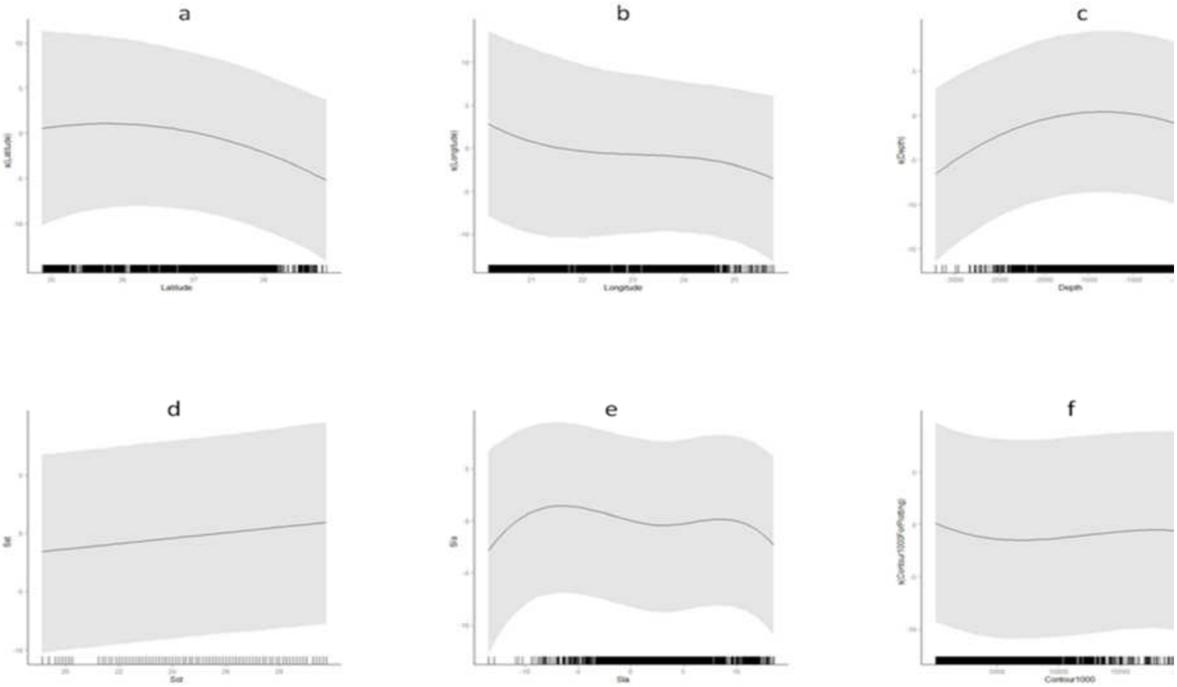
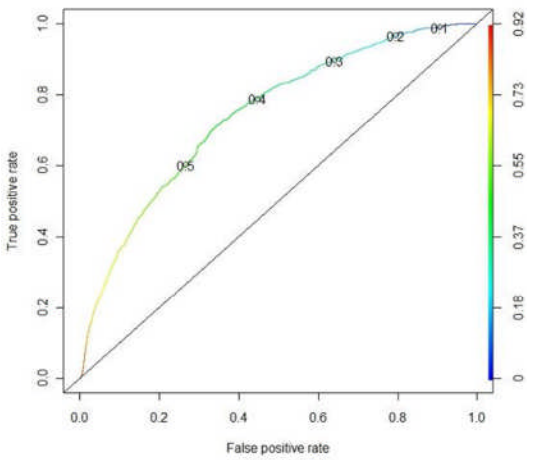


**Figure 6 Covariates remained in the final model for the whole dataset a) Latitude b) Longitude c) Depth d) SST e) SLA and f) Distance from 1km contour.**

A confusion matrix with a probability threshold value at 0.259 (derived from the ROC curve) shows that the model predicts correctly 58% (sd=0.006) and 42% of the absences. The area under the ROC curve (AUC) with value 0.743 .



| | | True Values | |
|---|---|---|---|
| | | Absence | Presence |
| Predicted Values | Absence | 0.421 | 0.389 |
| | Presence | 0.031 | 0.157 |

**Figure 7 Confusion matrix % for the full dataset & Roc plot with area of 0.743**

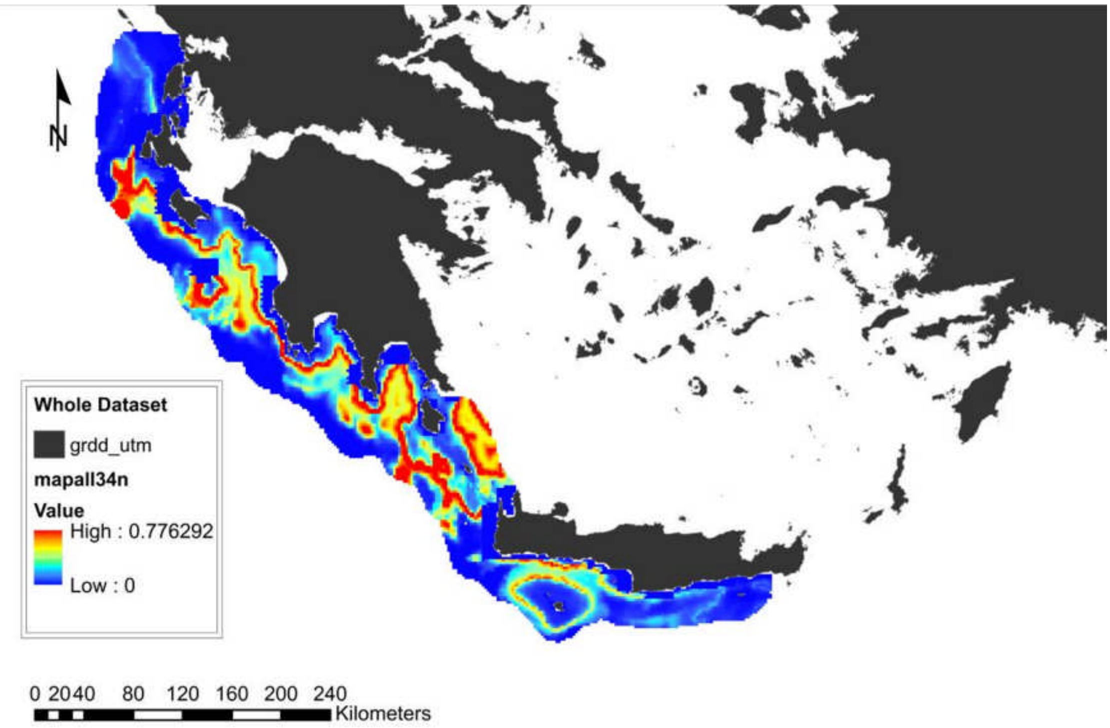


**Figure 8 Model prediction for the full dataset (social unites and solitary males) along the Hellenic Trench. Colors from blue to red represent the increasing probability of sperm whale presence.**

### 4.3.2 Model Variables for the different encounter types (males vs. social groups)

Only the fixed variables were used for the males' sperm whale habitat modelling, while for the social group dataset all the available variables. The presence of males was not constant between years so model including temporal variables did not converge.

### Dataset contained mature males

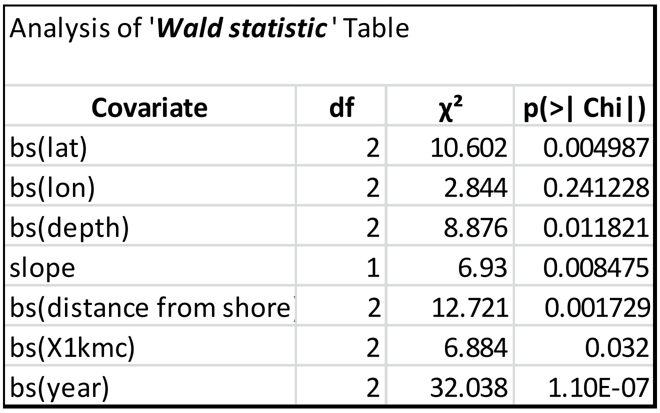
In the period 1998-2008, 77 mature male encounters were observed, of which 45 were loose male aggregations and 32 cases of solitary mature males. In 2009, no solitary males were observed. The variables that were retained in the model for the mature males were: depth, slope distance from shore, distance from 1km depth contour and year, in addition to latitude and longitude. All the variables remained as "bs" smoothed except slope for which model selection chose a linear form. Year had the higher coefficient range with a peak in the year 2002.

| **Covariates** | | Latitude | Longitude | Depth | Slope | Aspect | Distance from shore | Distance from 1kmc | Year | Depth | Slope | Aspect | Distance from shore | Distance from 1kmc | Year | **QIC** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Type** | | bs | | | | | | | | linear | | | | | | |
| **Model** | **1** | x | x | | x | x | | | | x | | | x | x | | 10987.98 |
| | **2** | x | x | | | | | | | x | x | x | x | x | | 10907.75 |
| | **3** | x | x | | | | | | | x | x | | x | x | x | 10776.03 |
| | **4** | x | x | | | | | | | x | x | x | x | x | x | 10689.7 |
| | **5** | x | x | | | | | | | x | x | x | x | | x | 10647.54 |
| | **6** | x | x | x | | | | | x | | | | | | | 9578.104 |
| | **7** | x | x | | x | x | x | x | x | | | | | | | 9492.128 |
| | **8** | x | x | | | | x | x | x | | | | | | | 9374.785 |
| | **9** | x | x | x | x | x | x | x | x | | | | | | | 9373.592 |
| | **10** | x | x | x | | x | x | x | x | | x | | | | | 9338.712 |

**Table 6 Different models for the male dataset which were taken under account in the model selection.**

| **Model coefficients for males dataset** | | |
|---|---|---|
| | **Estimate** | **Std err** |
| (intercept) | -3.6334 | 7.5137 |
| **bs(lat) 1** | 4.8179 | 4.0971 |
| **bs(lat) 2** | -10.9402 | 9.2734 |
| **bs(lon) 1** | -5.0975 | 8.2723 |
| **bs(lon) 2** | -6.34 | 7.0898 |
| **bs(depth) 1** | 11.6593 | 5.7731 |
| **bs(depth) 2** | 4.081 | 3.0504 |
| **slope** | 0.0384 | 0.0238 |
| **bs(distance)** | 2.9931 | 2.238 |
| **bs(distance)** | -6.9484 | 2.1675 |
| **bs(X1kmc) 1** | -3.3209 | 1.6481 |
| **bs(X1kmc) 2** | 7.592 | 2.5453 |
| **bs(year) 1** | 4.864 | 1.529 |
| **bs(year) 2** | -2.943 | 0.8585 |

**Table 7 Coefficients of the final model for the males' dataset**



Analysis of '***Wald statistic***' Table

| Covariate | df | χ² | p(>\| Chi\|) |
|---|---|---|---|
| bs(lat) | 2 | 10.602 | 0.004987 |
| bs(lon) | 2 | 2.844 | 0.241228 |
| bs(depth) | 2 | 8.876 | 0.011821 |
| slope | 1 | 6.93 | 0.008475 |
| bs(distance from shore | 2 | 12.721 | 0.001729 |
| bs(X1kmc) | 2 | 6.884 | 0.032 |
| bs(year) | 2 | 32.038 | 1.10E-07 |

**Table 8 Wald statistic table of the final model used for the males dataset**

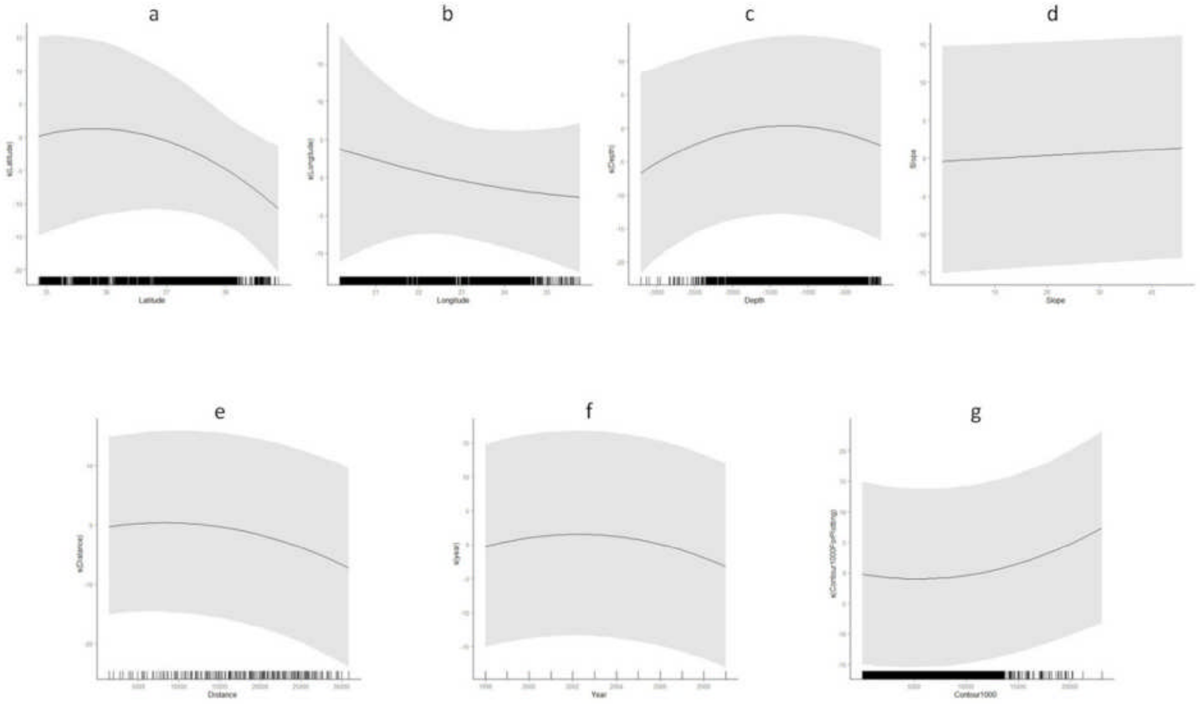
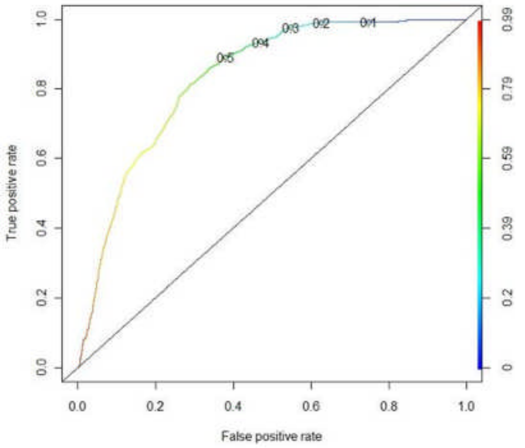


**Figure 9 Covariates of the final model for the males' dataset a) Latitude, b) Longitude, c) Depth, d) Slope, e) Distance from shore, f) Year g) Distance from 1km contour.**

The threshold value for the creation of the confusion matrix was chosen at 0.370, under the suggestion of the ROC curve. The confusion matrix shows that the model predicted 6568 correct and 2049 incorrect. The accuracy rate was at 0.762 (sd = 0.004) and the error rate at 0.238. The area under the curve gives a probability of 0.831 .



| | | True Values | |
|---|---|---|---|
| | | Absence | Presence |
| Predicted Values | Absence | 0.55182 | 0.20599 |
| | Presence | 0.0318 | 0.2104 |

**Figure 10 Confusion matrix % for the males dataset and Roc plot with area 0.831.**

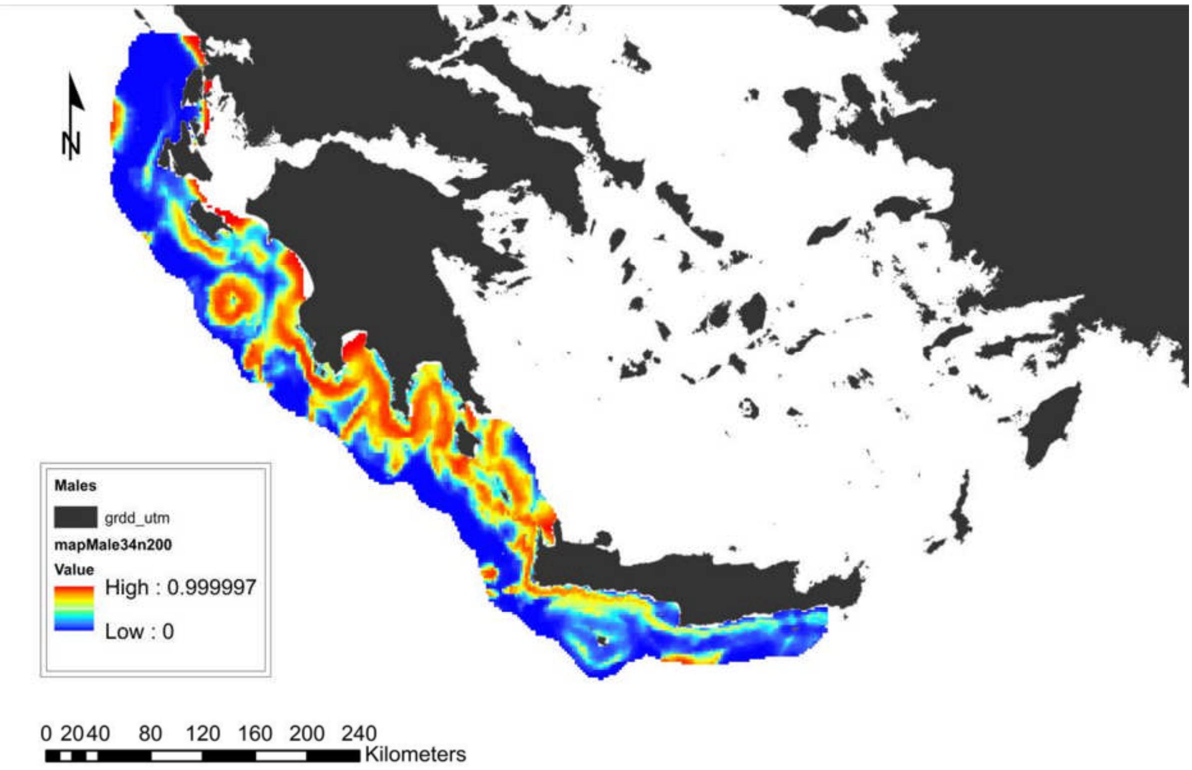


**Figure 11 Model predictions for the males' dataset along the Hellenic Trench. Colors from blue to red represent the increasing probability of male sperm whale presence.**

### Dataset containing social units (with or without males)

The dataset used for the social unit distribution contained 98 encounter cases of social groups of which 14 cases were with a mature male addition. The model with the smallest $QIC_u$=11445 included the variables: depth, slope, SST, CHL, SLA and distance from 1km contour plus latitude and longitude, which best describes the presence of social groups. The variable CHL was further removed based on the Wald's statistic. All the variables remained in the model as bs-smoothed form except SST which remained linear. Social group presence was best described by depth, slope and SLA variable. Depth showed an increasing relationship up to 1500m and then the presence probability decreased sharply.

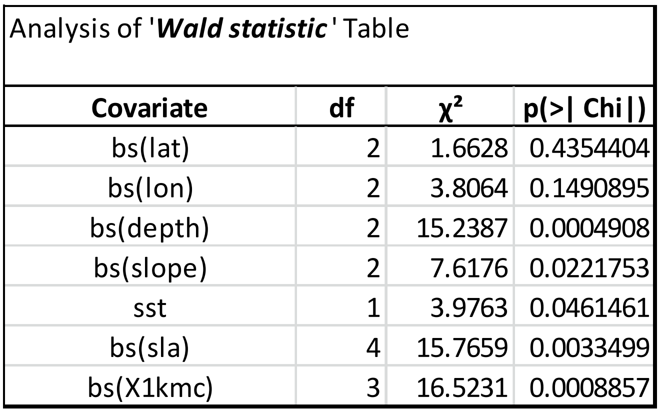
Slope displays a peak at 15°. For Sea level anomaly (SLA) the higher presence probability occurs at around -5 cm of sea level anomaly. Distance from 1km contour had high coefficient values at closer distances to the isobath.

| **Covariates** | | Latitude | Longitude | Depth | Slope | Aspect | Distance from shore | Sst | Chl | Sla | Distance from 1kmc | Distance from gyres | Depth | Slope | Aspect | Distance from shore | Sst | Chl | Sla | Distance from 1kmc | Distance from gyres | **QIC** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Type | | bs | | | | | | | | | | | linear | | | | | | | | | |
| **Model** | **1** | x | x | | | x | x | x | | | | x | | x | | | | x | | | | 13007.52 |
| | **2** | x | x | | | x | x | x | | | | | x | x | | | | x | | | | 12955.16 |
| | **3** | x | x | | | | | | | | | | | | | | | | | | | 12782.5 |
| | **4** | x | x | | | x | x | x | | x | | | x | x | | | | x | | | | 12600.34 |
| | **5** | x | x | | | | | | | | | | x | x | x | x | x | x | x | x | x | 12380.75 |
| | **6** | x | x | x | | | | | x | x | x | x | x | x | x | x | | | | | | 12150.65 |
| | **7** | x | x | | x | x | x | x | x | x | x | | | | | | | | | | | 12010.87 |
| | **8** | x | x | x | x | x | x | x | x | x | x | x | | | | | | | | | | 11721.41 |
| | **9** | x | x | x | x | x | x | x | x | x | x | | | | | | | | | | | 11605.96 |
| | **10** | x | x | x | x | | | | x | x | x | | | | | | x | | | | | 11445.11 |

**Table 9 Different models for the social unite dataset which were taken under account in the model selection.**

| **Model coefficients for social group dataset** | | |
|---|---|---|
| | **Estimate** | **Std err** |
| (intercept) | -12.126 | 6.3042 |
| **bs(lat) 1** | -1.5129 | 3.0503 |
| **bs(lat) 2** | -6.82 | 4.4326 |
| **bs(lon) 1** | -4.4535 | 3.8556 |
| **bs(lon) 2** | -6.6248 | 4.3594 |
| **bs(depth) 1** | 10.5317 | 4.0478 |
| **bs(depth) 2** | 2.0094 | 2.0395 |
| **bs(slope)1** | 2.6613 | 1.0819 |
| **bs(slope)1** | -6.1336 | 2.2541 |
| **sst** | 0.2781 | 0.1371 |
| **bs(sla) 1** | 12.1026 | 4.1059 |
| **bs(sla) 2** | 0.2464 | 1.776 |
| **bs(sla) 3** | 7.9205 | 3.1092 |
| **bs(sla) 4** | 2.4261 | 2.4449 |
| **bs(X1kmc) 1** | -4.4483 | 1.3418 |
| **bs(X1kmc) 2** | 1.5352 | 2.3757 |
| **bs(X1kmc) 3** | -2.2319 | 3.5071 |

**Table 10 Coefficients of the final model for the social unite dataset**



| Analysis of '***Wald statistic***' Table | | | |
|---|---|---|---|
| **Covariate** | **df** | **χ²** | **p(>\| Chi\|)** |
| bs(lat) | 2 | 1.6628 | 0.4354404 |
| bs(lon) | 2 | 3.8064 | 0.1490895 |
| bs(depth) | 2 | 15.2387 | 0.0004908 |
| bs(slope) | 2 | 7.6176 | 0.0221753 |
| sst | 1 | 3.9763 | 0.0461461 |
| bs(sla) | 4 | 15.7659 | 0.0033499 |
| bs(X1kmc) | 3 | 16.5231 | 0.0008857 |

**Table 11 Wald statistic table of the final model used for the social unites' dataset.**

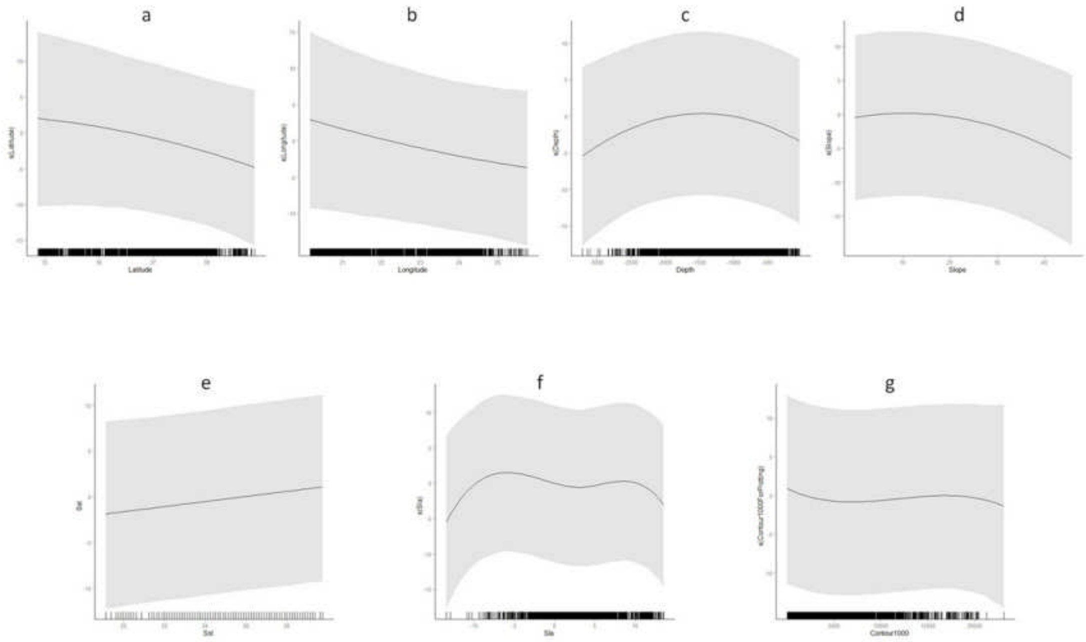
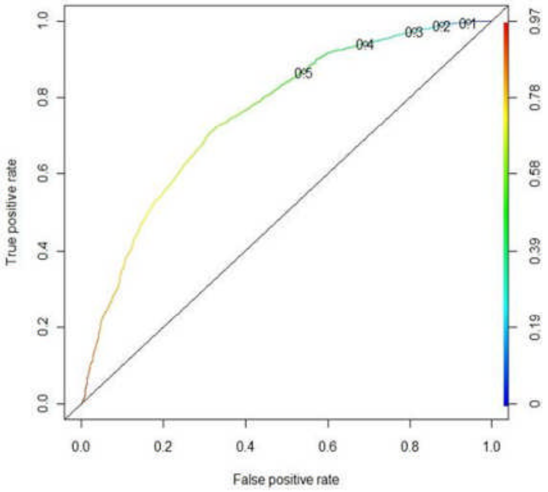


**Figure 12 Covariates of the final model for social unites' dataset a) Latitude b) Longitude c) Depth d) Slope e) SST, f) SLA and g) Distance from 1km contour.**

The cut-off probability derived from the ROC curve is 0.279 and the area under the curve (AUC) was 0.756. There were 6397 correct predictions and 3099 incorrect and the accuracy rate of the confusion matrix was 0.674 with a sd = 0.004 and error rate is 0.294.



| | | True Values | |
|---|---|---|---|
| | | Absence | Presence |
| Predicted Values | Absence | 0684 | 0.313 |
| | Presence | 0.014 | 0.065 |

**Figure 13 Confusion matrix % for female social unites' & ROC plot with area of 0.756**

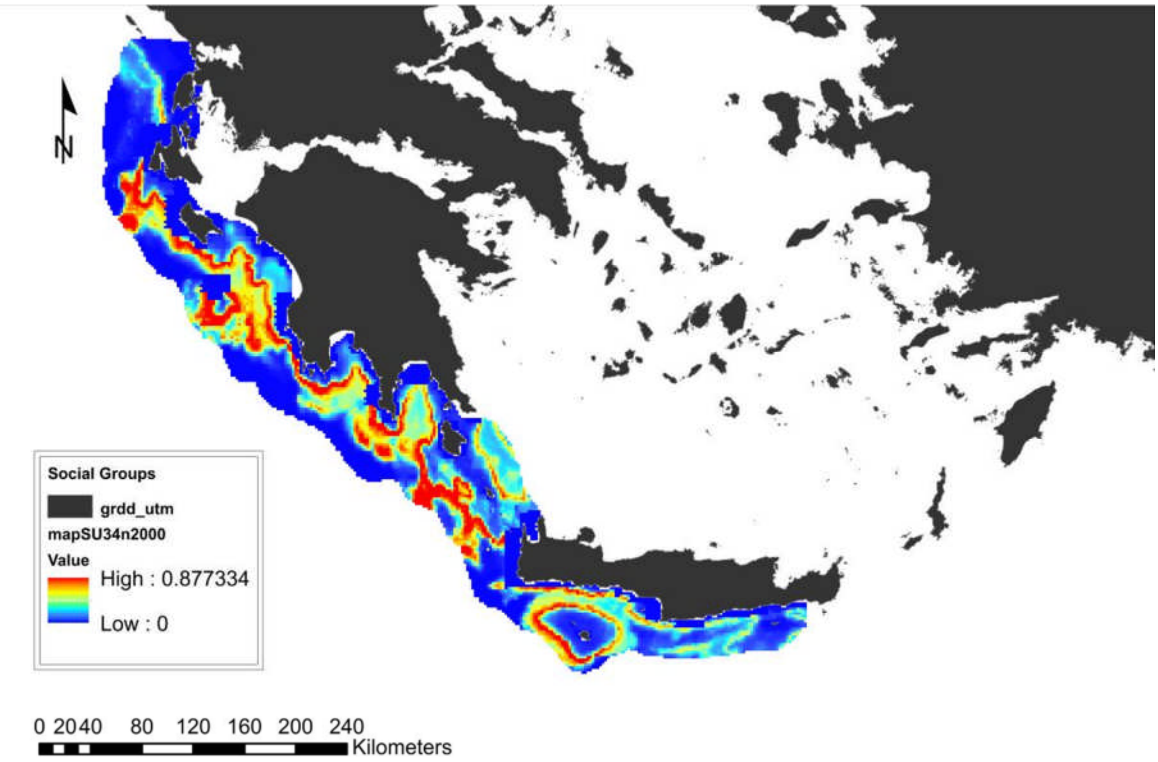


**Figure 14 Model predictions for social units along the Hellenic Trench. Colors from blue to red represent the increasing probability of social unites presence.**

## 4.4 Discussion

### 4.4.1 Sperm whale distribution

Sperm whales were seen every year throughout the 12 years of survey. The study area was characterized by the presence of summer hot spots of sperm whales and areas with low probability of encountering the species. In more detail, sperm whales were absent from the northern part of the study area near and above the island of Leukada and the south eastern part of Crete island, Ierapetra, where surveys have been conducted and no whales have been detected. Hot spots of sperm whales were found in the area on the south-west of Crete, further north and westerly of Kithira Island, south-west of Peloponnesos and north-westerly of Zakynthos island.

The modelling results demonstrate that sperm whales using the area of Hellenic Trench showed considerable preference for specific depths, slope, and certain distances from the 1km depth contour, as SST and SLA were the contribution of the temporal variables. Even though these variables were kept in the model, their ecological significance for the species must be further investigated. Bathymetry seems to be the driver of sperm whale presence in the area, while depth coefficients have the biggest range, showing bigger participation in the model. Sea Level Anomaly (SLA) has the second bigger coefficient range, showing a preference for areas that have lower or higher sea level anomaly values. The SLA is a measure for larger oceanographic processes such as gyres. Preference for slope values around 18° and a preference for larger SST values.

Aspect wasn’t included in any of the final models as was also found by Pirotta *et al.,* (2011). Sperm whales showed a preference for distances that are closer to the 1km depth contour, and the probability of their presence decreases with the distance from the contour and from the shore.

Sperm whales are found in waters above the continental slope and in open waters. Bathymetry seems to be the driving force for sperm whale distribution, as it has been previously suggested from other studies conducted in waters above the continental shelf such as the Gulf of Mexico and the Mediterranean. Slope seems to play an important role in most studies of sperm whale habitat (Jaquet, 1996). Preferable depth could differ between studies in different areas or when other environmental variable such SST change (Whitehead *et al.,* 1989), probably due to different prey spatial dispersion. Whitehead *et al.* (1989) found a negative correlation with faeces/fluke-up rates to SST and a higher rate in change of the direction in foraging whales suggested that the prey was more dispersed when SST was significantly higher, even though the distribution of whales in the area was the same in these two different years. Submarine canyons such as the Hellenic Trench or Blanes canyon (NW Mediterranean), Balearics in the Mediterranean, and gulfs such as the Gulf of Mexico, the Gulf of Lions in the Mediterranean where bathymetry is characterized by steep slopes, modify the local circulation (Flexas *et al.,* 2008) and become preferential recruitment habitats constituting aggregation areas for fish eggs and larvae (Flexas *et al.,* 2008). Guerra *et al.* (2011) found a positive correlation between submarine canyons close to the continental self with specimens of giant squid *Architeuthis*, which seem to provide a well-protected deep habitat near the productive areas of shallower feeding grounds. It is well known that sperm whales worldwide feed on mesopelagic squids (Whitehead, 2003), and in the Hellenic Trench *Histioteuthis bonnellii* represents their main prey species. Depth variability, slope and aspect in relation to the direction of currents may be important for the presence of squid and hence for the sperm whale presence. The results showing higher probability of sperm whale presence at negative values of Sea Level Anomaly and preference for smaller distances of the 1km isobath, agree with the results of Biggs *et al.* (2000) where sperm whales mostly encountered in regions with negative SSH (Sea Surface Height) along the 1km isobath in the Gulf of Mexico where cyclonic eddies were shaped.

### 4.4.2 Habitat of social groups vs. males

The model for single animals, specifically mature males, did not converge when the temporal variables were included and hence only the fixed variables were used. This may have happened because of the difference in frequency that males occurred throughout the years, so not enough data existed for the calculation of the model. As has been documented before (Whitehead, 2003) males are found closer to the shore waters, and our results confirm that the probability of males' presence decreases away from the shore. This may be due to the male preference for higher values of slope, which are found closer to the shore waters in our study area. It has been found that different sized sperm whales feed on different sized prey (Clarkes *et al.*, 1953), so this difference in slope preference from males and social groups may be explained in the difference of the prey habitat.

Social groups present a more complicated behaviour while males are more strictly driven by prey availability, social groups spend time in habitats where the food availability is low but they spend time for other purposes than feeding such as socializing, mating and giving birth. Explaining the habitat use of social animals brings about difficulties. The complexity of social behaviour is difficult to interpret because the reasons for their existence in that particular area is not always obvious, i.e. a male behaviour which continuously presents dives can be more easily interpreted as feeding behaviour in  areas with suitable feeding grounds. Solitary males and female social groups co-exist in the area presenting separate foraging behaviour but also share common ground for mating purposes. In areas such as the Hellenic Trench that seems to play an important role for the species for the Mediterranean subpopulation Social groups present an affinity for higher SST values, this preference combined with the preference for specific values for SLA could suggest that social groups may use areas of down welling where the productivity in lower levels of the water column would allow higher prey aggregation. Social groups with higher number of individuals would use the more productive areas in order to be able to obtain their energy requirements. The number of solitary male observations declined in more recent years of the survey.

### 4.4.3 Modelling work

The modelling procedure followed in this study and previously applied to sperm whales (Pirrotta *et al.*, 2011) and striped dolphins from (Panigada *et al.*, 2008) allows for differently derived data (regarding sampling procedures) to be pooled in the same analysis. Absences (acoustic follows) and presences (visual follows) had different spatial and temporal autocorrelation structures, however using the independence assumption (when the actual autocorrelation structure is unknown) allowed us to use the whole series of each data set. Using the actual follows and not just a point for each case of encounter or acoustic searching gives better information (bigger range of covariates values) regarding the environment for each data type. In addition, the use of smoothing splines for the environmental covariates allowed different relationships between our response variable (presence of sperm whales) and explanatory variables to be detected. The latter is important in trying to explain ecological processes which many times are more complicated than a linear relationship could account for. The use of latitude and longitude as a covariate in the model firstly takes into account the different amount of effort along the study area, this is important especially in long time series of surveys where the effort most probably differs. Furthermore, the two geographical variables could suggest areas that are important for the distribution of the species but the underlying processes couldn’t be explained from the available input variables. The predicted power of the model is higher when the model selection is based only through a likelihood based criterion such as AIC (or QIC in our case) but further reduction in the covariates is happening based in a significant statistical test. We chose to proceed in a further reduction as the most complicated model may have higher predicted power but there is also the risk of over fitting to the data. Over fitting can be found in more complex models vs. simpler models, so model comparison should be target specific.

## 4.5 Conclusions

Sperm whales inhabit the Hellenic Trench in the Eastern Mediterranean Basin and hot spots of probability presence exist along the study area. The Hellenic Trench is an important ground for sperm whales as multiple years of survey suggest permanent annual summer occurrence. Social groups and single or loose aggregations of males are using the area though the habitat of each seems to differ in the same range of environmental variables.

While the sperm whale subpopulation in the Mediterranean Sea has been characterized as endangered by the IUNC, it is important to find out the areas where the probability of presence is higher and to understand their habitat and the underlying ecological processes determining their occurrence in the area. The Mediterranean sperm whale subpopulation (Engelhaupt, 2009) is distinct from the Atlantic one and hence faces the well-known dangers of small, isolated populations such as changes in the environment, reduction of habitat, dependence in the prey availability and environmental changes. Anthropogenic activities such as shipping traffic and overfishing can negatively impact sperm whale population in the Mediterranean Sea, where the shipping traffic exists in high rates unlike other seas and open oceans. Long-term surveys are important for monitoring their population in order to identify shifts in space use, while knowing the underlying environmental forces that drive their distribution. New technological systems for ship traffic control like the Automated Identification Systems (AIS) could be used to assess the shipping density in the area of interest, identification of overlapping high-density areas could lead to decision making such as shifting the shipping lanes out of critical whale habitat in order to reduce risk of collision. Methods for comparative ship strike risk assessment and AIS data have been recently introduced by Leaper & Panigada (2012) in order to identify areas of high risk.

Further studies in the area are needed to collect longer follow data which would allow more complicated spatial models to be used for the analysis. Longer follow periods could allow us to combine existing models for foraging with the type of movement and behavioural data. Focus in different periods of the year in addition to summer surveys for gaining knowledge about their yearly distribution even though that would be costly. Thus recently automated passive acoustic monitoring equipment like the PAMBuoy could be placed along the Hellenic Trench and collect sperm whale density and distributional data all year round in order to support and strengthen the conclusions drawn from surveying the animals.

## 4.5 References

Aarts, G., MacKenzie, M., McConnell, B., Fedak, M. & Matthiopoulos, J. (2008) Estimating space-use and habitat preference from wildlife telemetry data. Ecography, 31, 140-160.

Arthur, S.M., Manly, B.F.J., McDonald, L.L. & Garner, G.W. (1996) Assessing habitat selection when availability changes. Ecology, 77, 215-227.

Azzellino, A., Gaspari, S., Airoldi, S. & 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 Part I-Oceanographic Research Papers, 55, 296-323.

Baumgartner, M.F., Mullin, K.D., May, L.N. & Leming, T.D. (2001) Cetacean habitats in the northern Gulf of Mexico. Fishery Bulletin, 99, 219-239.

Biggs, D.C., Leben, R.R. & and Ortega-Ortiz, J.G. (2000) Ship and satellite studies of mesoscale circulation and sperm whale habitats in the northeast gulf of mexico during GulfCet II. Gulf of Mexico Science, 18, 15-22.

Canadas, A., Sagarminaga, R., De Stephanis, R., Urquiola, E. & Hammond, P.S. (2005) Habitat preference modelling as a conservation took proposals for marine protected areas for cetaceans in southern Spanish waters. Aquatic Conservation-Marine and Freshwater Ecosystems, 15, 495-521.

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

Clarke, M.R., Martins, H.R. & Pascoe, P. (1993) The Diet of Sperm Whales (Physeter-Macrocephalus Linnaeus 1758) Off the Azores. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 339, 67-82.

Collum, L.A. & and Fritts, T.H. (1985) Sperm whales (Physeter catodon) in the Gulf of Mexico. Southwestern Association of Naturalists, 30, 101-104.

D'Amico, A., Bergamasco, A., Zanasca, P., Carniel, S., Nacini, E., Portunato, N., Teloni, V., Mori, C. & Barbanti, R. (2003) Qualitative correlation of marine mammals with physical and biological parameters in the Ligurian Sea. Ieee Journal of Oceanic Engineering, 28, 29-43.

Davis, R.W., Ortega-Ortiz, J.G., Ribic, C.A., Evans, W.E., Biggs, D.C., Ressler, P.H., Cady, R.B., Leben, R.R., Mullin, K.D. & Wursig, B. (2002) Cetacean habitat in the northern oceanic Gulf of Mexico. Deep-Sea Research Part I-Oceanographic Research Papers, 49, 121-142.

Drouot, V., Gannier, A. & Goold, J.C. (2004) Summer social distribution of sperm whales (Physeter macrocephalus) in the Mediterranean Sea. Journal of the Marine Biological Association of the United Kingdom, 84, 675-680.

Drouot, V., Goold, J.C. & Gannier, A. (2004) Regional diversity in the social vocalizations of sperm whale in the Mediterranean Sea. Revue D Ecologie-La Terre Et La Vie, 59, 545-558.

Drouot-Dulau, V. & Gannier, A. (2007) Movements of sperm whale in the western Mediterranean Sea: preliminary photo-identification results. Journal of the Marine Biological Association of the United Kingdom, 87, 195-200.

Embling, C.B., Gillibrand, P.A., Gordon, J., Shrimpton, J., Stevick, P.T. & Hammond, P.S. (2010) Using habitat models to identify suitable sites for marine protected areas for harbour porpoises (Phocoena phocoena). Biological Conservation, 143, 267-279.

Engelhaupt, D., Hoelzel, A.R., Nicholson, C., Frantzis, A., Mesnick, S., Gero, S., Whitehead, H., Rendell, L., Miller, P., De Stefanis, R., Canadas, A., Airoldi, S. & Mignucci-Giannoni, A.A. (2009) Female philopatry in coastal basins and male dispersion across the North Atlantic in a highly mobile marine species, the sperm whale (Physeter macrocephalus). Molecular Ecology, 18, 4193-4205.

Flexas, M.M., Boyer, D.L., Espino, M., Puigdefabregas, J., Rubio, A. & Company, J.B. (2008) Circulation over a submarine canyon in the NW Mediterranean. Journal of Geophysical Research-Oceans, 113.

Frantzis, A., Airoldi, S., Notarbartolo-di-Sciara, G., Johnson, C. & Mazzariol, S. (2011) Inter-basin movements of Mediterranean sperm whales provide insight into their population structure and conservation. Deep-Sea Research Part I-Oceanographic Research Papers, 58, 454-459.

Frantzis, A., Alexiadou, P. & Gkikopoulou, K. (Submitted) Sperm whale occurrence, site fidelity and population structure along the Hellenic Trench (Greece, Mediterranean Sea). Aquatic Conservation-Marine and Freshwater Ecosystems.

Frantzis, A., Swift, R., Gillespie, D., Mhhennett, C., Gordon, J. & Gialinakis, S. (1999) Sperm whale presence off South Crete, Eastern Mediterranean Sea

In: European Research on Cetaceans - 13. Proc. 13th Ann. Conf. ECS, Valencia, 214-217

Gannier, A., Drouot, V. & Goold, J.C. (2002) Distribution and relative abundance of sperm whales in the Mediterranean Sea. Marine Ecology-Progress Series, 243, 281-293.

Gannier, A., Drouot, V. & Goold, J.C. (2002) Distribution and relative abundance of sperm whales in the Mediterranean Sea. Marine Ecology-Progress Series, 243, 281-293.

Gannier, A. & Praca, E. (2007) SST fronts and the summer sperm whale distribution in the north-west Mediterranean Sea. Journal of the Marine Biological Association of the United Kingdom, 87, 187-193.

Garshelis, D.L. (2000) Delusions in Habitat Evaluation: Measuring Use, Selection, and Importance In: Research Techniques in Animal Ecology. Columbia University Press.

Gordon, J.C.D., Mathews, J.N., Panigada, S., Gannier, A., Borsani, J.F. & Notarbartolo di Sciara, G. (2000) Distribution and relative abundance of striped dolphins, and distribution of sperm whales in the Ligurian Sea cetacean sanctuary: results from a collaboration using acoustic monitoring techniques. J. CETACEAN RES. MANAGE., 2, 27-36.

Guerra, A., Gonzalez, A.F., Pascual, S. & Dawe, E.G. (2011) The giant squid Architeuthis: An emblematic invertebrate that can represent concern for the conservation of marine biodiversity. Biological Conservation, 144, 1989-1997.

Hardin, J., and Hilbe, J. 2003. Generalized Estimating Equations. London: CRC Press 240 pp.

Hastie, G.D., Swift, R.J., Slesser, G., Thompson, P.M. & Turrell, W.R. (2005) Environmental models for predicting oceanic dolphin habitat in the Northeast Atlantic. Ices Journal of Marine Science, 62, 760-770.

Hastie, T. & Tibshirani, R. (1986) Generalized Additive Models. Statistical Science, 1, 297-318.

Hutchinson, G.E. (1957) Concluding Remarks. Cold Spring Harb Symp Quant Biol, 22, 415-427.

Jaquet, N. & Gendron, D. (2002) Distribution and relative abundance of sperm whales in relation to key environmental features, squid landings and the distribution of other cetacean species in the Gulf of California, Mexico. Marine Biology, 141, 591-601.

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

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

Johnson, D.H. (1980) The Comparison of Usage and Availability Measurements for Evaluating Resource Preference. Ecology, 61, 65-71.

Kearney, M. (2006) Habitat, environment and niche: what are we modelling? Oikos, 115, 186-191.

Laran, S. & Drouot-Dulau, V. (2007) Seasonal variation of striped dolphins, fin- and sperm whales' abundance in the Ligurian Sea (Mediterranean Sea). Journal of the Marine Biological Association of the United Kingdom, 87, 345-352.

Laran, S. & Drouout-Dulau, V. (2002) Seasonal variation of striped dolphins, fin- and sperm whales’ abundance in the Ligurian Sea (Mediterranean Sea). J. Mar Biol Ass, 87, 345-352.

Larnicol, G., Ayoub, N. & Le Traon, P.Y. (2002) Major changes in Mediterranean Sea level variability from 7 years of TOPEX/Poseidon and ERS-1/2 data. Journal of Marine Systems, 33, 63-89.

Leaper R & Panigada, S. (2012) Some considerations on the use of AIS data to estimate shipping density for ship strike risk assessment. pp. 8. International Whaling Commission.

Leibold, M.A. (1995) The Niche Concept Revisited - Mechanistic Models and Community Context. Ecology, 76, 1371-1382.

Lewis, T., Gillespie, D., Matthews, L.J., Danbolt, M., Leaper, R., McLanaghan, R. & Moscrop, A. (2007) Sperm whale abundance estimates from acoustic surveys of the Ionian Sea and Straits of Sicily in 2003. Journal of the Marine Biological Association of the United Kingdom, 87, 353-357.

Liang, K.Y. & Zeger, S.L. (1986) Longitudinal Data-Analysis Using Generalized Linear-Models. Biometrika, 73, 13-22.

O'Hern, J.E. & 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.

Pan, W. (2001) Akaike's information criterion in generalized estimating equations. Biometrics, 57, 120-125.

Panigada, S., Notarbartolo di Sciara, G., Zanardelli Panigada, M., Airoldi, S., Borsani, J.F. & Jahoda, M. (2005) Fin whales (Balaenoptera physalus) summering in the Ligurian Sea: distribution, encounter rate, mean group size and relation to physiographic variables. J. CETACEAN RES. MANAGE., 7, 137-145.

Panigada, S., Pesante, G., Zanardelli, M., Capoulade, F., Gannier, A. & Weinrich, M.T. (2006) Mediterranean fin whales at risk from fatal ship strikes. Marine Pollution Bulletin, 52, 1287-1298.

Panigada, S., Zanardelli, M., MacKenzie, M., Donovan, C., Melin, F. & Hammond, P.S. (2008) Modelling habitat preferences for fin whales and striped dolphins in the Pelagos Sanctuary (Western Mediterranean Sea) with physiographic and remote sensing variables. Remote Sensing of Environment, 112, 3400-3412.

Pierce, G.J., Valavanis, V.D., Guerra, A., Jereb, P., Orsi-Relini, L., Bellido, J.M., Katara, I., Piatkowski, U., Pereira, J., Balguerias, E., Sobrino, I., Lefkaditou, E., Wang, J., Santurtun, M., Boyle, P.R., Hastie, L.C., MacLeod, C.D., Smith, J.M., Viana, M., Gonzalez, A.F. & Zuur, A.F. (2008) A review of cephalopod-environment interactions in European Seas. Hydrobiologia, 612, 49-70.

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

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

Pujol, M.I. & Larnicol, G. (2005) Mediterranean sea eddy kinetic energy variability from 11 years of altimetric data. Journal of Marine Systems, 58, 121-142.

Roberts, S.M. (2003) Examination of the stomach contents from a Mediterranean sperm whale found south of Crete, Greece. Journal of the Marine Biological Association of the United Kingdom, 83, 667-670.

Robinson, A.R., Golnaraghi, M., Leslie, W.G., Artegiani, A., Hecht, A., Lazzoni, E., Michelato, A., Sansone, E., Theocharis, A. & Unluata, U. (1991) The Eastern Mediterranean General-Circulation - Features, Structure and Variability. Dynamics of Atmospheres and Oceans, 15, 215-240.

Scott-Hayward, L. (2006) Modelling the spatial distribution of sperm whales, Physeter macrocephalus, in the Gulf of Mexico. M.Res. Environmental Biology, University of St Andrews.

Stephanis, R., Cornulier, T., Verborgh, P., Sierra, J.S., Gimeno, N.P. & Guinet, C. (2008) Summer spatial distribution of cetaceans in the strait of Gibraltar in relation to the oceanographic context. Mar. Ecol Prog Ser, 353.

Teloni, V., Mark, J.P., Patrick, M.J.O. & Peter, M.T. (2008) Shallow food for deep divers: Dynamic foraging behavior of male sperm whales in a high latitude habitat. Journal of Experimental Marine Biology and Ecology, 354, 119-131.

Tobler, W.R. (1970) Computer Movie Simulating Urban Growth in Detroit Region. Economic Geography, 46, 234-240.

Tudela, S., Kai, A.K., Maynou, F., El Andalossi, M. & 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.

Wakefield, E.D., Phillips, R.A. & Matthiopoulos, J. (2009) Quantifying habitat use and preferences of pelagic seabirds using individual movement data: a review. Marine Ecology-Progress Series, 391, 165-182.

Whitehead, H. (2003) Sperm whales: Social Evolution in the Ocean. The Univerisity of Chicago Press, Chicago and London.

Whitehead, H., Papastavrou, V. & Smith, S.C. (1989) Feeding Success of Sperm Whales and Sea-Surface Temperature Off the Galapagos-Islands. Marine Ecology-Progress Series, 53, 201-203.

Whittaker, R., Levin, S.A. & Root, R.B. (1973) Niche, Habitat, and Ecotope. American Naturalist, 107, 321-338.

# 5. ABUNDANCE ESTIMATION OF THE SPERM WHALES OF THE HELLENIC TRENCH

## 5.1 Introduction

The estimation of wildlife abundance is necessary for population monitoring and assessing the impact of human activities. Cetaceans live exclusively in the water and estimating their population size can be difficult due to their environment and cryptic way of living. The spatial distribution of cetacean species can be wide-ranging and usually unknown. In open oceans, seasonal variation in the geographical distribution patters of many cetaceans due to migration makes the estimation of abundance even more difficult. One of the most used methods for analysing data for abundance estimation for marine mammals are mark-recapture analyses of capture histories of individuals marked with tags or photographically (Hammond, 1986), and distance sampling based on sampling space (Buckland *et al.,* 2001).

Distance sampling uses transect methods and information from the distribution of the species in order to estimate their abundance. Methods exist for point or line transects (the latter being more commonly used in the marine environment). Distance sampling methods record distances of the observed (usually sighted) animals from the observer's position (a person or an automated detection system). Distance sampling uses cues for relating the presence/detection of a species. Cues for cetaceans are either visual or acoustic, depending on the type of survey. Visual surveys are not efficient for species that spend much time submerged such as sperm whales which dive for 45-50 minutes at a time and spend 86% of their time underwater (Gordon, 1987). Acoustic surveys hold a number of advantages in comparison to visual surveys, acoustic cues can be detected in much bigger distances than the visual cues (~ 2nm), they are received from all the directions (Leaper *et al.,* 1992), whereas visual coverage is only a small fraction of the area that has to be scanned acoustic surveys can be conducted for 24 hours a day regardless of light conditions and are less influenced by weather conditions. As a result, acoustic surveys can collect larger amounts of data. Acoustic surveys can be particularly effective especially with highly vocal species such as the Sperm whale. Most sperm whale surveys use acoustic surveys for initial detecting the animals because of their characteristic vocal foraging sound clicks and their long-lasting feeding dives.

Sperm whales are highly vocal animals producing a variety of sounds using their spermaceti organ, such as clicks, creaks (rapid click buzzes), codas (social sounds). Sperm whale clicks are sharp-onset, broadband, impulsive vocalizations with energy between 5 and 25 kHz. The clicks can be very powerful, up to 223 dB re 1 μ Pa @ 1 m, the highest biologically produced source levels that have ever been recorded and are highly directional (Zimmer *et al.,* 2005). Directionality of the clicks could influence the detectability of the animals in farther distances. Sperm whale clicks have been grouped in two broad functional groups; communication and echolocation. Echolocation, “usual clicks” are the most common type and are produced by the sperm whales during deep, feeding dives. Usual clicks have been used for cue distance sampling methods (Borchers *et al.,* 2007).

Sperm whales are social animals, forming matrilineal groups comprising mature females and immature animals of both sexes. When males reach their reproductive age they form bachelor groups and they disconnect from their natal groups (Best, 1979) and by the time they reach full sexual maturity they become mostly solitary animals that occasionally join other social groups for mating or interacting with other mature males (Whitehead & Weilgart, 2000).

Inter-click intervals within usual click trains average about 0.5 s for females and immature males and 1.0 s for mature males (reviewed from Whitehead, 2003). Sperm whale social groups spend more time on the surface socializing compared to solitary males that tend to conduct continuous feeding dives; however the overlapping and multidirectional vocalisations from the large number of individuals in social groups produce an almost continuous train of clicks (Whitehead, 2003), hence social groups have a bigger probability of detection.

For distance sampling methods, a detection function (probability of detecting an animal during the survey) must be calculated from the recorded distances from the animal of interest to the observer. Detection functions derived from only acoustic methods are hard to calculate due to the uncertainty in estimates of the position of the vocalising animals. In addition, difficulties arise in the calculation of the number of the individuals belonging to the group. A number of methods have been used for calculating the position of the vocal animals during acoustic cetacean surveys, such as Cartwheels (from Conservation Research Group, 1989) using triangulation from different listening stations or more improved methods of signal direction algorithms with the use of a towed hydrophone (Leaper *et al.,* 1992). Sophisticated software such as Rainbow Click (Gillespie, 1997), and more recent the Pamguard program (Gillespie, 2008) allows for more precise estimates of the vocal animal position.

To estimate abundance, the number of observed animals must be calculated, in the case of acoustic methods there are difficulties in counting the actual number of observed individuals because of difficulties in separating individuals through sound and also because of the superimposing of the individuals calls in case of group members.

For sperm whale abundance estimation photo-identification methods, visual surveys and acoustic surveys (Lewis *et al.,* 2007; Leaper *et al.,* 1992) have been conducted. Although mark-recapture analyses of photo identification data and acoustic censuses have considerable promise, they have been used only occasionally, and the great majority of sperm whale population estimates come from visual censuses (reviewed in Whitehead, 2002). An estimate up to 360 000 animals has been estimated for the world oceans (Whitehead, 2002). Abundance estimation for submerged animals can be difficult. Due to the small proportion of time that sperm whales spend on the surface (20% for the males and 25-30% for the female group members visual surveys alone would tend to underestimate the abundance because of the small number of detections that can be derived from them. On the contrary, passive acoustic data collected by hydrophones towed along line- or point-transects constituted a more appropriate method for the highly vocalizing sperm whales.

In the open ocean, it is difficult to estimate the population of an animal especially in the case of sperm whale because of its pelagic life style (Jaquet, 1996). Even though the Mediterranean is a semi-closed sea there are no population estimates for the sub-population of the basin (Whitehead, 2002), only relative abundance estimates (Gannier *et al.*, 2002; Laran, 2007).

The first acoustic sperm whale abundance estimate in the Mediterranean, obtained by IFAW (2007) in the Ionian Sea and the Straits of Sicily, was 62 animals. Due to the small sample size of detected sperm whales in the Ionian Sea, supplementary data were obtained from other parts of the western Mediterranean in order to estimate the detection function (Lewis *et al.*, 2007). Estimates of group size of the animals only by acoustic methods can be difficult because of their complicated social behaviour and their unsynchronized diving times (Barlow *et al.*, 2005). Silent periods during socializing at the surface or either sleeping periods may negatively bias these estimates. A better method would combine acoustic with visual surveys (Barlow *et al.*, 2005).

Between 2003 and 2007, IFAW carried out cetacean surveys in the Mediterranean that suggested that the greatest density of sperm whales occurs in north-western waters of the Mediterranean (Boisseau *et al.*, 2010). The total number of sperm whales in the Mediterranean is more likely to be in the hundreds rather than the thousands and it is suspected to be declining (Reeves and Notarbartolo di Sciara, 2006).

As acoustic and visual surveys are not efficient on their own, a combination of acoustic surveys with data from visual surveys could be used to maximizing the use of the data and the estimates of absolute abundance (Barlow *et al.*, 2005).

### 5.1.1 Objectives

The main objectives of this chapter are to estimate the acoustic detection function and the effective acoustic range for sperm whales along the Hellenic Trench, estimate the abundance of the sperm whale population in the study area and identify areas of high abundance along the Hellenic Trench.

## 5.2 Materials and Methods

### 5.2.1 Study area

The study area is located in the Eastern Mediterranean Basin along the Hellenic Trench (39°N, 20°E to 34°N, 26°E), which is the main topographic characteristic resulting in extreme changes in bathymetry. The area has a highly dynamic oceanographic environment which is influenced by the variable slopes in bathymetry and coastline.

### 5.2.2 Data collection

A combined acoustic - visual survey was conducted in the area of the Hellenic Trench for the summer periods 1998-2009. Pre-determined cruise tracks were conducted in the study area, while a towed hydrophone was used for the acoustic detection of the sperm whales.

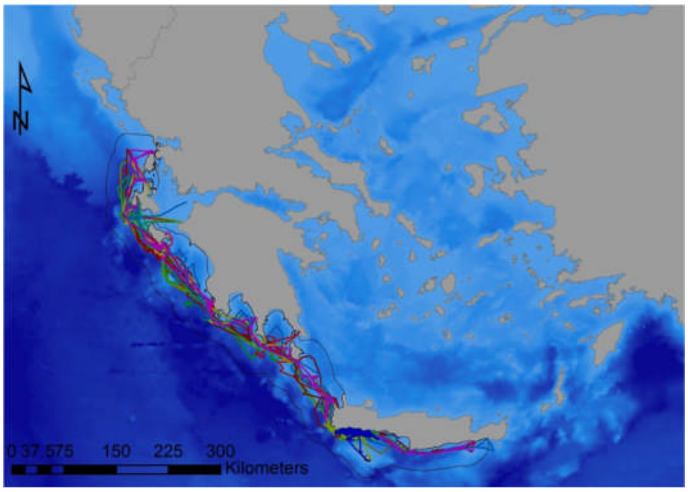


**Figure 15 Yearly survey tracks from 1998 to 2009; each year is represented by different colored line.**

The vessel had an approximate speed of 7-9 knots while in searching mode and 2 visual non-independent observers scanned the horizon covering an area of 180° in front of the boat. Following hydrophone deployment, an acoustic sampling of 1 minute duration (listening station) was conducted every 15 min. In each listening station, an experienced acoustic observer was analysing real time data and recording the binary response, presence/absence, of sperm whales or other cetaceans that were found in the area. In addition, a qualitative assessment of the acoustic information was recorded for the following variables: vessel noise; other vessels; water noise; number of sperm whales heard and strength of the sperm whale signal. All the previous qualitative measurements had values from 0 – absence - to 5 - very loud. The engine of the vessel was off in most of the acoustic sampling stations in order to minimize the noise, except in rough weather conditions. The software Rainbow Click (Gillespie, 1997) was used and was running during the acoustic survey, which identifies putative sperm whale clicks, calculates their bearings and attempts to distinguish sperm whale clicks from other transients based on their duration and spectral content. The software calculates bearings to each click from the relative time of arrival of the click at the two hydrophones in the array.

Upon acoustic detection of sperm whales the searching effort was stopped and tracking effort was followed in order to locate the animals. The necessary time was spent with the encounters for the photo-identification study, in the case of social groups, at least two dive cycles (time of fluke until the first blow) were spent with the animals.

Search effort is measured in a number of points, corresponding in the number of listening stations. A total of 4399 listening stations occurred between 1998-2009.

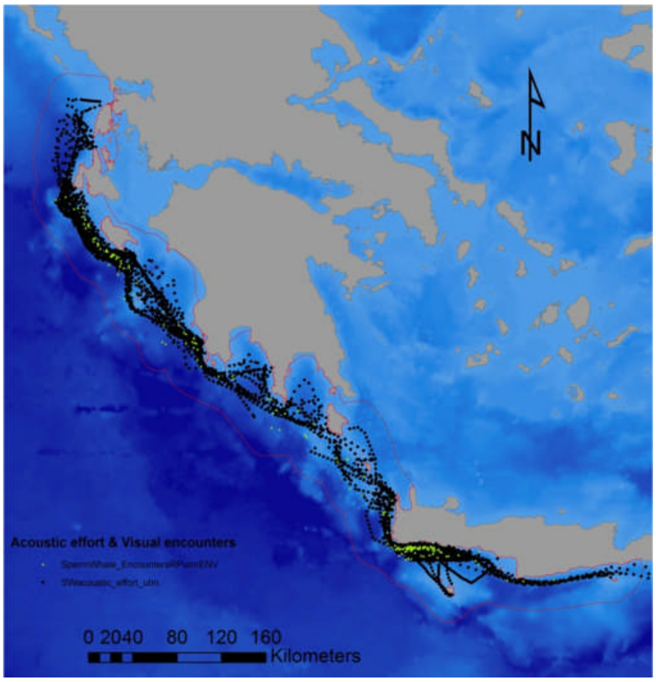


**Figure 16 Map with acoustic effort and visual encounters in the Hellenic Trench. Black dots represent the acoustic listening stations and green triangles the visual encounters.**

### 5.2.3 Data used for the analysis

Two datasets were used for the abundance estimation of sperm whales in the Hellenic Trench. An experimental data set was used for the calculation of the acoustic detection function g(x) (and information derived from it, such as the effective acoustic radius distance), and a dataset of all the detected sightings of the sperm whales during the 12 dedicated yearly surveys.

The experimental data set consisted of 62 *departure* experiments which collected data while leaving the whales. Each experiment resulted in 2-8 cases of recorded distances recorded as follows: While we were with the animals (social group or single male), we started moving away from them in an non pre-designed direction and, at random time intervals, a listening station was conducted in order to identify if sperm whales were still in our acoustic detection range. The binary information Detection (1) or Absence (0) was recorded each time, in addition to the noise level as described above.

The total number of the final experimental dataset from which the detection function was derived consisted of 202 occasions of acoustic stations, each with associated distances, binary detection information, number of animals sighted, type of encounter, position, and time of listening station that corresponded to the 62 departure experiments.

### 5.2.4 Data analysis

Point transect distance sampling was applied to analyse the absolute abundance of sperm whales in Hellenic Trench. Due to the social nature of the sperm whales and their occurrence in groups, a stratification method was used. The sample unit of the analysis was cluster. The experimental dataset was used to calculate the acoustic detection function which was further used for estimating the abundance.

### 5.2.5 Movement model

During our departure experiments we recorded the distance between the listening station and the point of where we left the whales. However, the animals were usually in motion and therefore we needed to correct the estimate of distance by using a simple movement model: Mean direction and speed of the encounters for each occasion was calculated from the follow dataset. For each of the departure experiments the corresponding sighting dataset were used to calculate mean direction and speed of the encounter. The new positions were derived by linear extrapolation from their last known location depending on their direction and speed.

### 5.2.6 Detection function g(x)

The detection function is the probability of detecting an animal within the covered area, given its characteristics and other associated environmental or survey-level variables. Generally, the detection function decreases with increasing distance, but always has values from 0 to 1 (Buckland *et al.,* 2001). The detection function comprises two parts: A key function is adequate as a model for the detection function and a ‘series expansion’ is used to adjust the key function and to improve the fit of the model to the distance data. (Buckland *et al.*, 2001). The general form for modelling the detection function is as follows:

$$g(y) \propto key(y)[1 + series(y)] \qquad (1.5)$$

Where $g(y)$ is the detection function and $y$ is the vector of recorded distances.

Mostly the uniform, half normal distribution and hazard rate models have been used as a key function of the detection probability $g(x)$. Three series expansions are considered by Buckland *et al.* (2001) 1) the cosine series, 2) simple polynomials and 3) Hermite polynomials, these three expansions are linear in their parameters.

Detection function is often not only dependent on distance. It may depend on the ability of the surveyor, the characteristics of the individual animals, environmental conditions and other factors (Buckland *et al.*, 2001). Sometimes it is useful to model the detection probability as a function of variables other than distance, a detection function $g(y, \underline{z})$ which depends on distance $y$ and some other appropriate variable $\underline{z}$. Modelling the detection probability as a function of other variables can be useful in the following cases: 1) density is correlated with detection probability, 2) a large component of the variance of the abundance estimate is due to estimation of the detection function, and this variance can be explained by variables other than distance and 3) detection probability changes across strata but there are inadequate detections in some strata to allow separate estimation of detection probability within each stratum.

Multi-covariate detection models have been used in previous literature. A Generalized Linear Model (GLMs) with a logit binomial function was used in our case to model the acoustic detection function of the sperm whales. The use of GLMs allowed us to include additional covariates with distance.

A generalized linear model (GLM) is an extension of the linear model that allows the distribution of the response to be from the exponential family such as Gaussian, Binomial, Poisson or Gamma. The detection probability was modelled as a binary response variable (with a Bernoulli distribution). A general formula for this GLM is as follows:

$$g(E[Y_i]) = g(\mu_i) = \log\left(\frac{p_i}{1-p_i}\right) = \eta_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_p x_{ip} \qquad (0.6)$$

Where *i* is the model specified, *p* is the number of covariates, $g(\mu_\iota)$ is the link function and $\eta_\iota$ is the linear predictor.

A binary dataset of 0 and 1 was used to fit the detection function, with 0 corresponding to an absence of acoustic detection of the animals and 1 signifying an acoustic detection. The covariates that were used in the model were: 1) distance, 2) number of animals, 3) type of encounter (solitary matured males, loose male aggregation and social groups) , 4) depth, 5) noise (was quantified in a qualitative way 0-5, which derived from information about vessel noise, other vessels, water noise), 6) year and two additional interactions between 7) distance and type of encounter and 8) distance and number of whales.

The full model comprised the six variables plus two interactions as described below:

$$E[p(\text{detection})] = \log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1(\text{distance}) + \beta_2(\text{number})$$
$$+\beta_3(\text{type.encounter}) + \beta_4(\text{depth}) + \beta_5(\text{noise}) + \beta_6(\text{distance} \times \text{type.encounter}) \quad (1.7)$$
$$+\beta_7(\text{distance} \times \text{number}) + \beta_8(\text{year})$$

Where $p$ is the probability of detecting a group of whales and 1- $p_i$ is the probability of no detection.

### 5.2.7 Model selection

The model selection was based on Akaike's Information Criterion (AIC; Akaike, 1974). AIC is based on the likelihood and asymptotic properties of the maximum likelihood estimator (MLE). The AIC is a measure of fit of the model which is increasingly penalized with additional parameters estimated from the model. A relatively lower AIC score indicate a better fit of the model. The model with the smallest AIC value was selected.

$$AIC = -2\log_e(L) + 2q \quad (1.8)$$

Where $\log_e(L)$ the maximized log-likelihood function and $q$ is the number of estimated parameters in the model.

A detection function should have three desired properties: robustness, a shape criterion and efficiency. A shape criterion- detection function should have a "shoulder" near the point, which means that detection remains nearly certain at small distances from the point (i.e. the derivative of the detection function should be zero, $g'(0) = 0$). In addition the function should be non-increasing and have a tail that goes asymptotically to zero. Efficiency is desirable in the sense that the selected model should provide estimates that are relatively precise, i.e. have small variance.

### 5.2.8 Abundance estimation

The estimator of abundance is very closely linked with the detection function, for that reason it is crucial to select the best model for the detection function.

Point transect distance sampling with clustered objects was used for the estimation of abundance. The estimator takes the form

$$\hat{D} = \frac{E(n) \cdot E(s)}{a \cdot P_a} \tag{1.9}$$

Where $a$ is the area covered, $P_a$ is the probability of detection for an object unconditional on its actual position, $E(n)$ is the estimator of the detected clusters and $\hat{E}(s)$ is the estimator of the mean group size for each cluster size.

In the estimation of cluster populations estimating the mean group size is $\hat{E}(s)$ not always straight-forward. Analyzing clustered population the abundance estimator takes the following form:

$$\hat{D} = \frac{\hat{E}(n) \cdot h(0) \cdot \hat{E}(s)}{2\pi k} \tag{1.10}$$

Where $E(n)$ is the estimator of the detected clusters, $h(0) = \lim_{r \to 0} f(r)/r$ is the probability density function of detected distances from the point, evaluated at zero distance and $E(s)$ is the estimator of the mean size of each cluster and $k$ is the number of points sampled.

In line transect sampling, the probability density function (pdf) of distances *f ( r )* has the same shape as the detection function *g(x)* but in point transect sampling the two are different. The pdf of distances is proportional to *r g(r)*, with constrain that *f(r)* integrates to unity. So the probability density function of detected distances takes the form of:

$$f(r) = \frac{r \cdot g(r)}{\int_0^w r \cdot g(r) dr} \tag{0.0}$$

Where *w* is the truncation distance used for the detection function, *r* is the recorded distance and *g(r)* is the detection function. Where distances beyond the distance w were ignored in the analysis, in order bias to be minimized from outlying detected distances.

When animals occur in clusters the estimation of mean group size can be biased, due to the difference in detectability of group size. There are different solutions to overcoming the possible bias that can occur in clustered populations, including: 1) truncation, 2) stratification by cluster size, 3) weighted average of cluster size, 4) regression estimators, 5) use of covariates and 6) replacing clusters by individual objects. We used stratification by cluster size and abundance estimation was calculated independently for each stratum. In cases where there are not enough data for estimating abundance independently for each of the observed cluster size the pooling method could be used and each stratum could correspond in clusters with similar size. The data were divided in three strata of similar cluster size as shown below:

| **Stratum** | **Cluster size** |
|---|---|
| I | $1-3$ |
| II | $4-7$ |
| III | $8-15$ |

**Table 12 Stratification of cluster size**

Cluster size stratification is generally used to account for differences in detectability between clusters of different sizes and the resolution of stratification is reflected in the bin size used. Although stratification into broad group size bins may cause some loss of precision (as information about within-bin variability is discarded), it may be a pragmatic approach in the case of sparse data sets such as ours, where binning behaves as a rudimentary smoother and acts to counteract spurious sampling variation. Furthermore, calculating the detection function in three different strata instead of 15 different cluster sizes, results in a smaller variance in the total abundance estimate. This is because the mean size of each stratum will always be larger than the smaller group size of the detected cluster and smaller group sizes will have a lower detection probability and an associated higher variance.

The stratum I contains the solitary males and the loose male aggregation.

In the case of stratification in point transect sampling the mean cluster size in the population is estimated by

$$\hat{E}(s) = \frac{\sum_u n_u s_u \hat{h}_u (0|s = s_u)}{\sum_u n_u \hat{h}_u (0|s = s_u)} \tag{1.12}$$

Where $n_u$ is the number of detections for each stratum, $s_u$ is the cluster size for the $n_u$ and $h_u(0|s=s_u)$ is the estimator of derivative of probability density function (pdf) of the detection function for each of the strata, as suggested by Quinn (1979).

The effective radius of detection *p* is a product of detection function for which as many objects beyond *p* are detected as are missed within *p*. Is estimated by

$$\hat{p} = \sqrt{2/\hat{h}(o)} \tag{1.13}$$

Where *h(o)* is the slope of the probability density function of detected distances from the point, evaluated at zero distance.

### 5.2.9 Effort points k

The total number of points used as sampling effort in the abundance estimate was 4399. Because the acoustic effective radius was bigger than the distance between two successive listening stations, the number of points was reduced depending on the distance between them. The reduction procedure was stopped when no overlap between two successive listening circular areas existed. This process was applied for time correlated acoustic effort, so acoustic listening stations for each day of the survey were independent.

For estimating the abundance of sperm whales along the Hellenic Trench we further divide the study area A in poor, medium and good quality habitat estimated from the distribution map created as described in the previous chapter. Each cell of the distribution map included the information of the probability of sperm whale presence with values ranging from 0 to 1. Poor habitat was defined as that with probability of presence in the range 0-0.33, medium habitat 0.33-0.66 and good quality habitat the remaining probability of 0.66-1.

Final abundance estimates derived from summation of the separate estimates for each stratum in each habitat type. Three stratums and three habitat types divided the whole dataset into nine portions, leading to 9 components for the abundance estimate.

### 5.2.10 Abundance variance estimation

Variance in the estimated abundance has three components. First, there is a variance associated with the encounter rate (i.e detected groups), second, variance associated with estimating the detection function in the experimental dataset and third, variance associated with the group size of the encounters. A nonparametric bootstrap approach was followed for deriving the uncertainty of the abundance estimate. The departure experiments and the while following the whales dataset were randomly resampled with replacement. The whole analysis was repeated 500 times and the same numbers of abundance estimates were obtained. The sample variance of the bootstrap estimates of abundance was taken as an estimate of the variance of the estimator $\hat{N}$ . In order to extract the abundance estimate CI, the bootstrap abundance estimates are order and the 95% confidence intervals is given by $\left[\hat{N}_{(j)}, \hat{N}_{(k)}\right]$ with $j = (B+1)a$ and $k = (B+1)(1-a)$ where B is the number of repeats and $a$ is 0.025. The advantage of bootstrap over analytic variance estimates is that there is no need for independence assumption for the estimates n, $\hat{h}(0)$ and $\hat{E}(s)$. By applying the full estimation procedure to each replicate, components of the variance for estimating the number of adjustment terms and for estimating $E(n)$ and $E(s)$, and any additional multipliers, are all automatically incorporated.

## 5.3 Results

### 5.3.1 Effort and sightings

The dataset used for deriving the detection function was made up of 202 records from 62 departure experiments. Each experiment contained 2-8 independent records of distance recordings. During the years 1998-2009 a combined acoustic/visual survey was conducted in the area of Hellenic Trench with a total 361 days of active survey. These yielded 178 occasions of visual sperm whales detection, the acoustic effort in total was 4399 listening stations. The histogram of the recorded distances from the departure experiments is shown in Fig17.

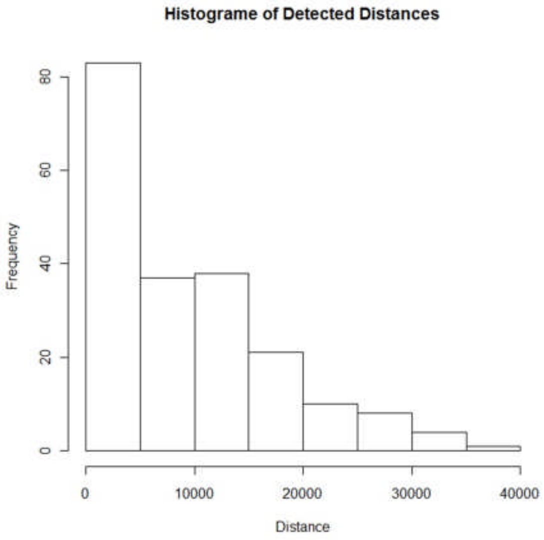


**Figure 4 Histogram of detected distances from the departure experiments.**

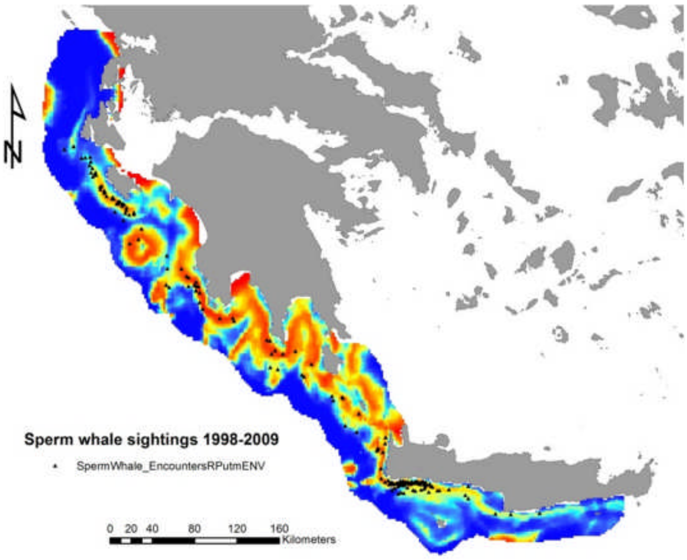


**Figure 5 Model prediction for the sperm whale presence in the area along the Hellenic Trench and sperm whale sightings during the period 1998 to 2009. Colors from blue to red represent the increasing probability of sperm whale presence and sightings denoted with black points.**

### 5.3.2 Detection function

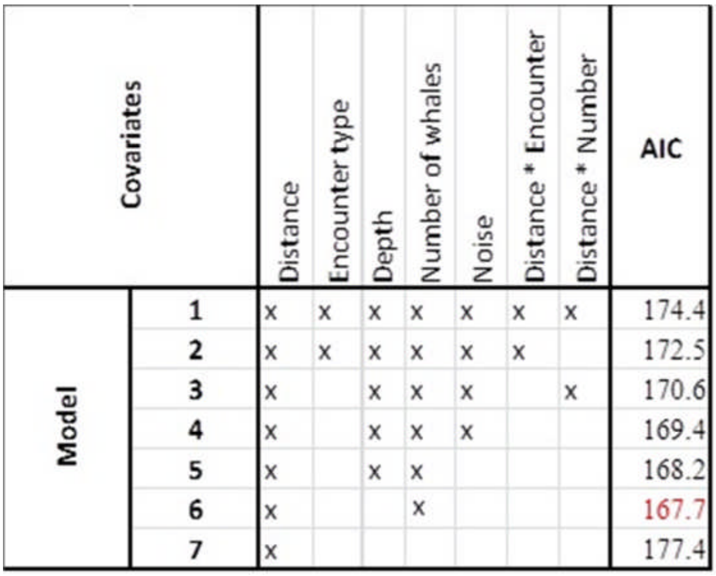
The data set of detection distances were truncated at 27.6 km (w) leaving outside 8 detections, 0.5% of the whole dataset. The dataset used for the derivation of the detection function also included the variables: 1) number of whales, 2) encounter type, 3) noise, 4) depth, 5) year.

The full model (GLMs) was composed by 8 variables as shown below:

$$E[p(\text{detection})] = \log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1(\text{distance}) + \beta_2(\text{number}) \qquad (1.14)$$

Where $p_i$ is the probability of detecting a group of whales and $1\text{-}p_i$ is the probability of no detection.

The model with the smallest AIC was kept for modelling the acoustic detection function. An AIC value of 177.68 suggested that the best model fit was the one where distance and number of whales was included in the model.



| **Covariates** | | Distance | Encounter type | Depth | Number of whales | Noise | Distance * Encounter | Distance * Number | **AIC** |
|---|---|---|---|---|---|---|---|---|---|
| **Model** | **1** | x | x | x | x | x | x | x | 174.4 |
| | **2** | x | x | x | x | x | x | | 172.5 |
| | **3** | x | | x | x | x | | x | 170.6 |
| | **4** | x | | x | x | x | | | 169.4 |
| | **5** | x | | x | x | | | | 168.2 |
| | **6** | x | | | x | | | | 167.7 |
| | **7** | x | | | | | | | 177.4 |

**Table 13 AIC values of the models taken under account in the model selection for the detection function. Covariates that participated in the model selection are: distance, encounter type (solitary males, loose male aggregation, and social group), depth, number of whales, noise and two interactions distance with encounter type and distance with number of whales.**

A $\Delta_{AIC}$ of value 9.7, when the covariate number of whales was omitted from the full model, indicates that the covariate number of whales had a significant effect.

on the detection function.

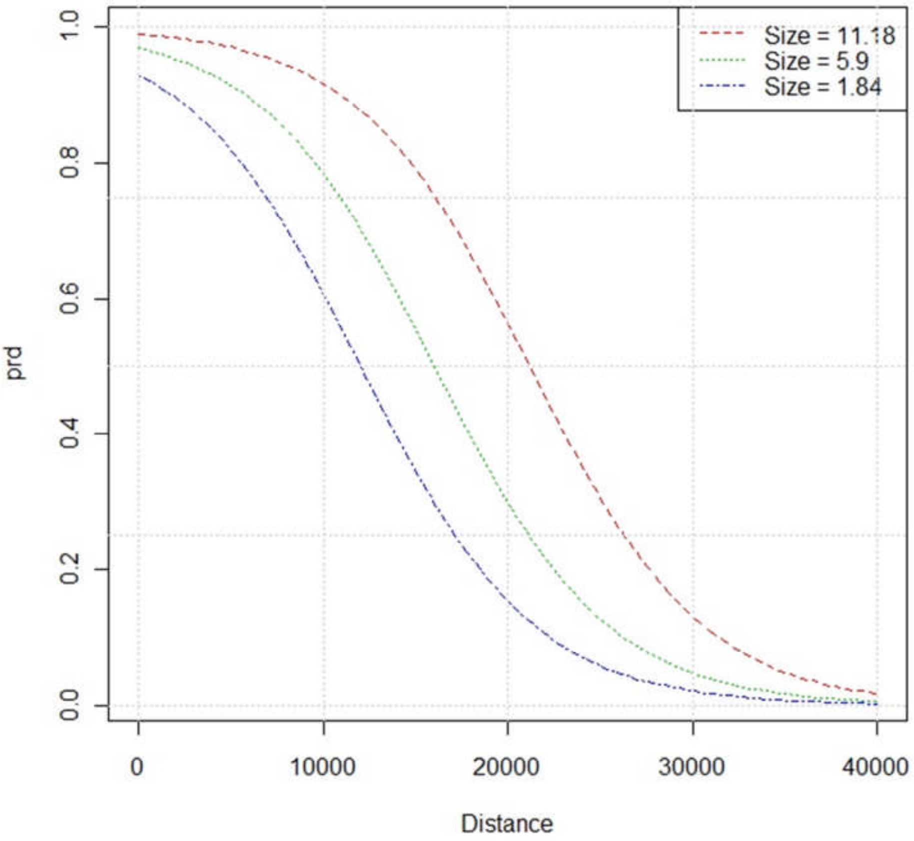


**Figure 19 Detection function for the three different strata with mean number of whales of 1.84, 5.9 and 11.18.**

The detection function for each stratum with 95% CI is shown below:

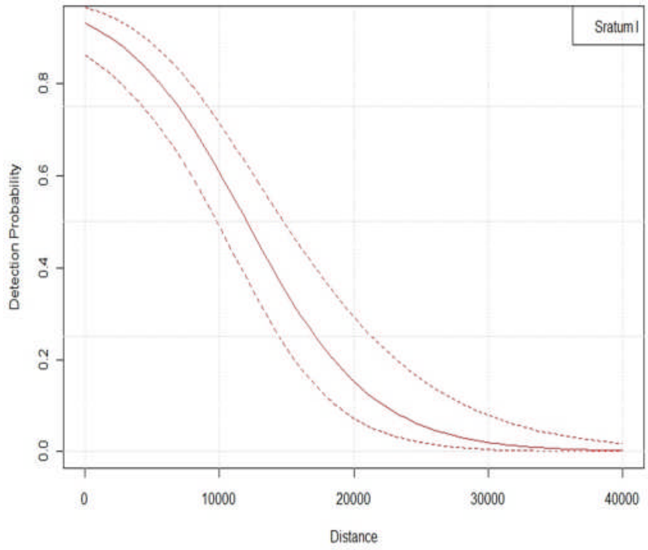


Figure 20 Detection function for Stratum I (1.84 animals) with 95% CI.

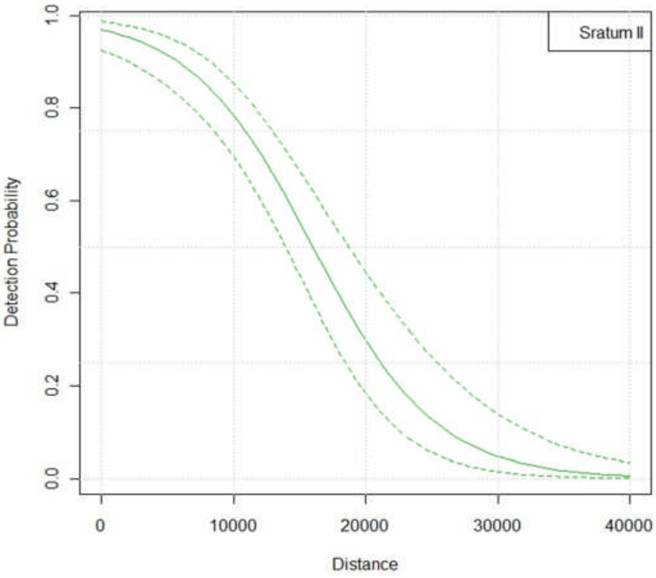


Figure 21 Detection function for Stratum II (5.9 animals) with 95% CI.

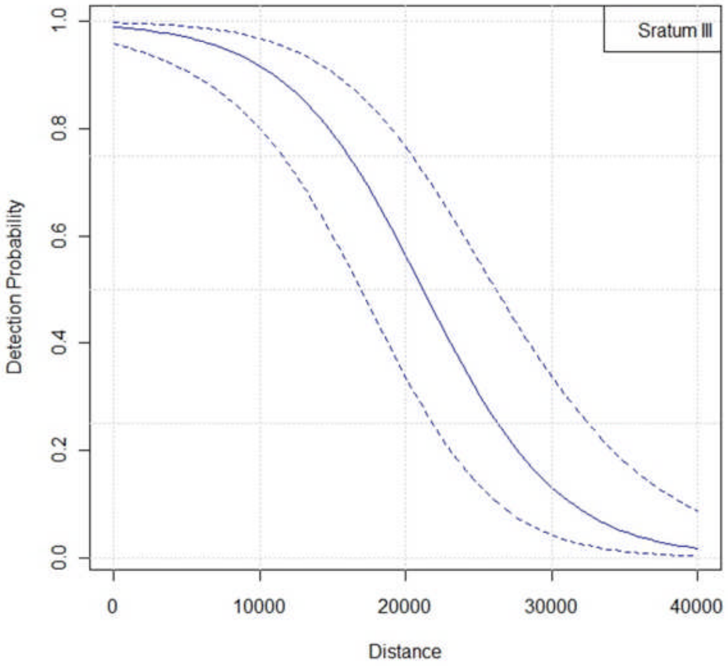


Figure 22 Detection function for Stratum III (11.18 animals) with 95% CI

| Detection function coefficients | | |
|---|---|---|
| | **Estimate** | **Std error** |
| (Intercept) | 2.20E+00 | 4.02E-01 |
| **Distance** | -2.15E-04 | 3.23E-05 |
| **Number of whales** | 2.10E-01 | 6.58E-02 |

Table 6 Summary of the final model used for the detection function.

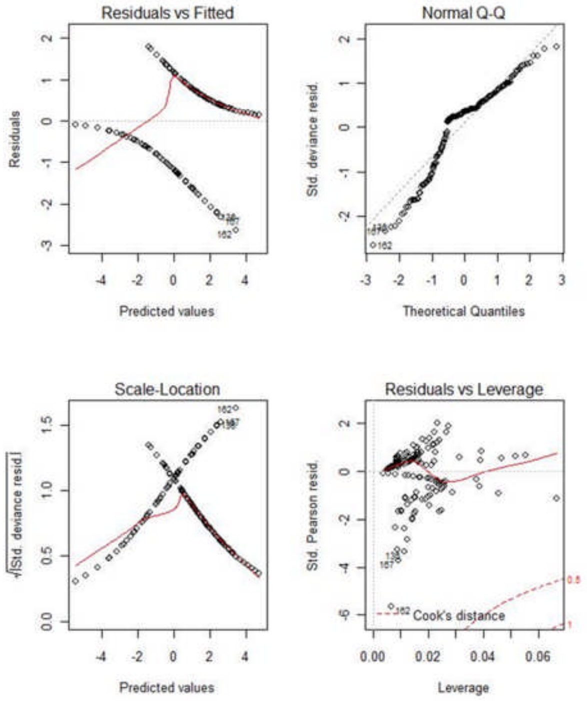


**Figure 23 Diagnostic plots for final GLM model fitting the detection function.**

Diagnostic plots for binary data is not very helpful as a residual can only take two possible values for a given predicted response. The most useful plot is the fourth (lower right) plot that shows that no outliers influenced the model.

The g(o) for the stratum I, II and III calculated to be 0.93 , 0.97 and 0.99 respectively.

| **Stratum** | **g(o)** |
|---|---|
| I | 0.93 |
| II | 0.96 |
| III | 0.99 |

**Table 15 Probability of acoustic detection at zero distance. The g(o) derived from the GLM fitted the detection function**

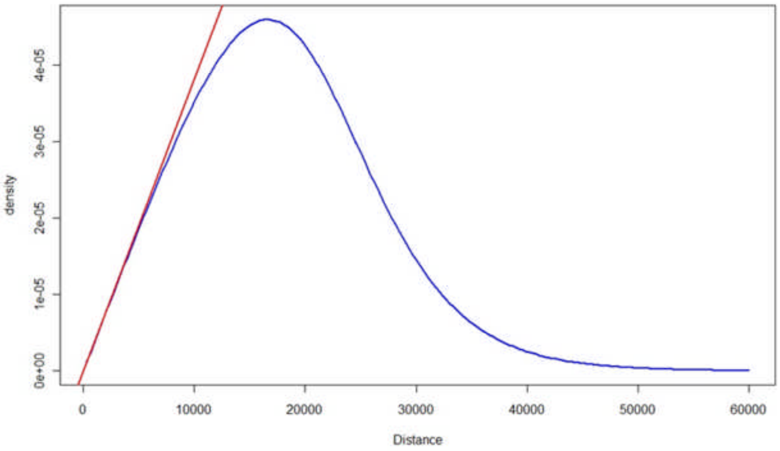


**Figure 24 Probability density function (pdf) of detected distances (blue line) and h(o) – the slope of the probability density function of detected distances from the point, evaluated at zero distance (red line)**

The *h(o)* for the stratum I, II and III was 1.041674e-08 , 6.796694e-09, 4.611756e-09 respectively.

The effective radius (the distance where the detection function is equal to 0.5) for the three different strata can be shown below.

| **Stratum** | **Effective radius** |
|---|---|
| I | 13856.36 |
| II | 17154.03 |
| III | 20824.85 |

**Table 16 Effective radius (in meters) for the three different strata**

### 5.3.3 Abundance estimate

The final abundance estimate derived from the 9 separate abundance estimates for each combination of stratum (I, II, III) and type of habitat (poor, medium, good). The detected objects (Table 16), effort points (Table 17) and the *h(o)* which differed along the strata used for the estimation of density. The density was then multiplied with the area corresponding to each type of habitat.

| | Habitat | | |
|---|---|---|---|
| | Poor | Medium | Good |
| Stratum I | 2 | 16 | 62 |
| Stratum II | 5 | 7 | 40 |
| Stratum III | 1 | 10 | 27 |

**Table 17 Number of detected clusters for the three different strata in each type of habitat (Poor, Medium, Good).**

| | Habitat | | |
|---|---|---|---|
| | Poor | Medium | Good |
| Stratum I | 268 | 731 | 929 |
| Stratum II | 249 | 677 | 839 |
| Stratum III | 229 | 629 | 757 |

**Table 18 Number of effort points for the three different strata in each type of habitat (Poor, Medium, Good).**

A total abundance estimate of 26.99 whales with 95% of [19.7, 32.08] derived from and analytical method and a non-parametric bootstrap of 500 resamples respectively.

The abundance estimates for the each combination of stratum and habitat can be seen below:

| | Habitat | | | |
|---|---|---|---|---|
| | Poor | Medium | Good | **Total** |
| Stratum I | 1.094332 | 1.115524 | 3.451148 | 5.661004 |
| Stratum II | 6.190337 | 1.104195 | 5.163007 | 12.45754 |
| Stratum III | 1.728831 | 2.183652 | 4.967021 | 8.879504 |
| Total | 9.0135 | 4.403371 | 13.58118 | **26.99805** |

**Table 19 Abundance estimates for the three different strata in each type of habitat (Poor, Medium, Good).**

The abundance estimates in each habitat type (Good, Medium, Poor) are as follows: 13.58, 4.40 and 9.01 animals.

## 5.4 Discussion

This study presents an analysis of combined acoustic/visual detection data for the abundance estimation of sperm whales along the Hellenic Trench. Point distance sampling application with stratification in the number of animals was used because of the clustered nature of occurrence of these social animals. An acoustic detection function was derived from experimental (departure) data, obtained from the study area, showing evidence of detection dependence on group size. The effective acoustic range for 3 different strata of group sizes was estimated at 13.85 km, 17.15 km and 20.82 km for the corresponding group size of 1-3, 4-7 and 8-15 animals respectively. The “good” quality habitat for sperm whales had the biggest proportion of estimated density 0.5 versus number to medium and poor habitat together.

A total abundance estimate of 27 ([19.7, 32.08] 95% CI) sperm whales was derived from our analysis. A Previous study in the Ionian Sea and Sicily, and the first of sperm whale abundance estimate in parts of the Mediterranean whales ,from acoustic surveys conducted in 2003, found 62 animals to exist in the Ionian Sea with 95% lognormal confidence limits of 24 to 165 (Lewis *et al.*, 2007). Our estimate of 24.99 whales falls at the lower limit of this estimate. This may be due to the facts that our estimated acoustic effective radius was larger than that study’s (10.0 km) and our study area was smaller by Due to the combined acoustic and visual method followed in our surveys, it was possible to ascertain the actual number of each group in contrast to acoustic-only survey methods which often arrive at underestimates of abundance. Simplifying the behavioural stage of the group formation when the acoustic detection function is dependent in the group size could potentially lead to less accurate estimates of abundance. Lewis et al. (2007) stratified the groups to either dispersed or clustered; with clustered groups to present smaller perpendicular distances at which they detected.

In a comparison study using acoustic and visual methods for different areas of the Mediterranean, the Ionian Sea (east Mediterranean) presented the second highest density of sperm whales after the Gulf of Lion (west Mediterranean) (Gannier *et al.*, 2002). In the Ligurian Sea during the summer, the relative density of sperm whales is lower compared to other areas of the western Mediterranean basin (Gannier *et al.,* 1999; Gannier *et al.*, 2002) and the annual density is increased during the period August to October (Laran and Drouot, 2007). Sperm whale annual density seems to show a geographical shift, so abundance estimates could differ if the survey data are collected in different times of the year. In addition, different areas comprise different grounds for the sperm whales and behavioural differences could lead to differences in detectability that would also influence the abundance estimates of the sperm whales.

Cetacean abundance estimates have been most commonly inferred by visual surveys. Nevertheless, because of the highly vocal nature of the sperm whale and its long dive-times, acoustic surveys are by far more efficient in detecting and estimating their abundance (Leaper *et al.*, 1992; Leaper *et al.*, 2000; Hastie *et al.*, 2003; Barlow *et al.*, 2005; Lewis *et al.*, 2007). From acoustic surveys greater estimates of detection range derive (Gannier *et al.*, 2002; Barlow *et al.*, 2005) i.e 8km in comparison to visual of 4km (Gannier *et al.*, 2002). Leaper *et al.* (2000) found maximum detection distances of 30km and a half width of detected perpendicular distance of 8km. In the Faroe Shetland Channel an effective range lying between 5km and 7km (Hastie *et al.*, 2003). In our analysis the effective radius varied between 13 km and 20 km, depending on group size.

Abundance estimates are highly dependent on the detection function, which, in turn, depends on the characteristics of the survey design, the characteristics and behaviour of the study animal, the characteristics of the mark/cue used as the response data in the detection function and also on the quality of the habitat and geographical area where the study is taking place. Bias could exist in all the previous steps of the analysis and the analysis should consider the potential factors influencing the detection function.

Acoustic surveys alone have a tendency to underestimate the number of the animals detected because of the discriminant ability of the software regarding the number of individuals heard. In addition the survey mode - passing or closing – could influence the number of detections (Dawson *et al.,* 2008) due to the dispersal of the animals that observed in groups such as the sperm whales. A closing combined acoustic with visual survey design corrects for the biases in detections, but there is a risk of observing only part of the group detected if it is scattered over a big surfacing area. In the Mediterranean, sperm whale groups are smaller than those of the open oceans such as the Atlantic and Pacific, so the risk of partial detection is small, thus favouring closing mode surveys. Another source of bias on the detectability cause by the survey design is the speed of the vessel that can affect the data in two main ways: 1) The noise of the vessel itself is the main noise source affecting whale detection (Leaper *et al.,* 2000) 2) vessel speed influences the depth at which the towed hydrophone is located in the water column, with higher speeds the hydrophone is closer to the surface and the ability to detect animals decreases (Hastie *et al.,* 2003).

Detectability additionally depends on the behaviour of the study animal. In the case of the sperm whale, where behaviour between the two sexes differs considerably, the g(0) in the detection function may vary tremendously. The detection function of the sperm whales in acoustic surveys uses the “usual clicks” as a cue, produced by the animals during their feeding dives. Solitary males, produce these about 81% (Watwood *et al.,*2006) and 86% of the time; Gordon & Steiner, 1992). Social groups of females emit a smaller proportion of foraging clicks because they spend more time socializing at the surface; during this time, communication sounds known as codas

(Watkins & Schevill, 1977) are produced by the animals. Codas are less directional compared to the "usual clicks", so they are not as useful for acoustic detection. The dive stage could influence the range of which animals are detected. Barlow *et al.* (2005) have shown that the slow clicks produced during the first and last minutes of a feeding dive are detectable over greater distances compared to the usual clicks produced mainly during the dive. In addition to the type of the clicks, dive-stage can influence detectability because sound propagation varies depending on the depth of the source in the water column.

The behavioural characteristics of a species can differ in different geographical regions. For example, the duration of the silent period from the fluke up time until the first click and the duration of the last click until surfacing varies between different sub-populations (Douglas *et al.,* 2005). Using cue methods, click rate has been found to vary with region, age and sex and thus it is inappropriate apply one standard click rate in acoustic surveys of sperm whales (Douglas *et al.,* 2005). Hence, abundance estimates corrected for g(0) and detectability factors should not use information produced for other geographical areas. In our study values for g(0) were derived from the GLM that fitted the detection function.

By incorporating covariates such as group size in the model for the detection function, biases can be minimised. Furthermore our analysis for the detection function could be improved by measuring the ambient noise from actual recordings and not just qualitative levels that we used in our model.

Estimating the abundance of clustered populations induces bias existing in the groups or clusters of individuals, failure to account for the relationship between detectability and cluster size tends a positive bias in estimates of abundance. Hierarchical models have been used for modelling the bias of cluster populations (Royle, 2008). Sperm whales present an interesting occasion for abundance estimation due to their highly vocal nature and social behaviour. The cue method (that is a transformation of point sampling method) for acoustic surveys could be more appropriate as it takes into account the proportion of the silent periods of a vocal animal.

## 5.5 References

Akaike, H. (1974) New Look at Statistical-Model Identification. Ieee Transactions on Automatic Control, Ac19, 716-723.

Barlow, J. & Taylor, B.L. (2005) Estimates of sperm whale abundance in the northeastern temperate Pacific from a combined acoustic and visual survey. Marine Mammal Science, 21, 429-445.

Best, P.B. (1979) Social organization in sperm whales, *Physeter macrocephalus*. In Behavior of marine animals, vol.3 Winn, H.E., Olla., B.L. (Eds.) Plenum, New York. 227-289

Boisseau, O., Lacey, C., Lewis, T., Moscrop, A., Danbolt, M. & 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, 1589-1599.

Borchers, D.L., Brewer, C. & Matthews, J. (2007) Methods for estimating sperm whale abundance from passive acoustic line transect surveys. pp. 16. Center for Research into Ecological and Environmental Modelling, St Andrews.

Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., D.L., B. & Thomas, L. (2001) Introduction to Distance Sampling: Estimating abundance of biological populations. Oxdord University Press, Oxford.

Dawson, S., Wade, P., Slooten, E. & Barlow, J. (2008) Design and field methods for sighting surveys of cetaceans in coastal and riverine habitats. Mammal Review, 38, 19-49.

Douglas, L.A., Dawson, S.M. & Jaquet, N. (2005) Click rates and silences of sperm whales at Kaikoura, New Zealand. Journal of the Acoustical Society of America, 118, 523-529.

Drouot, V., Goold, J.C. & Gannier, A. (2004) Regional diversity in the social vocalizations of sperm whale in the Mediterranean Sea. Revue D Ecologie-La Terre Et La Vie, 59, 545-558.

Gannier, A. & Drouot, V. (1999) Distribution and relative abundance of the sperm whale in the central and western Mediterranean. European Research on Cetacean.

Gannier, A., Drouot, V. & Goold, J.C. (2002) Distribution and relative abundance of sperm whales in the Mediterranean Sea. Marine Ecology-Progress Series, 243, 281-293.

Gillespie, D. (1997) An Acoustic Survey for Sperm Whales in the Southern Ocean Sanctuary Conducted from the RSV Aurora Australis. Report INT. WHAL. COMMN, 47.

Gillespie, D., Gordon, J., Mchugh, R., Mellinger, D.K., Redmond, P., Thode, A., Trinder, P. & Deng, X.Y. (2008) PAMGUARD: Semiautomated, open source software for real-time acoustic detection and localization of cetaceans. Proceedings of the Insitute of Acoustics, 30.

Gordon, J.C.D. (1987) Behaviour and ecology of sperm whales off Sri Lanka. . Ph.D. dissertation, University of Cambridge.

Gordon, J.C.D. & Steiner, L. (1992) Ventilation and dive patterns in sperm whales, Physeter macrocephalus, in the Azores. REP. INT. WHAL. COMMN, 42, 561-565.

Hammond, P.S. (1986) Estimating the size of naturally marked whale populations using capture-recapture techniques. Report INT. WHAL. COMMN, 253-282.

Hastie, G.D., Swift, R., Gordon, J.C.D., Slesser, G. & and Turrell, R.W. (2003) Sperm whale distribution and seasonal density in the Faroe shetland Channel J. CETACEAN RES. MANAGE., 5, 247-252.

Jaquet, N. & Gendron, D. (2002) Distribution and relative abundance of sperm whales in relation to key environmental features, squid landings and the distribution of other cetacean species in the Gulf of California, Mexico. Marine Biology, 141, 591-601.

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

Laran, S. & Drouot-Dulau, V. (2007) Seasonal variation of striped dolphins, fin- and sperm whales' abundance in the Ligurian Sea (Mediterranean Sea). Journal of the Marine Biological Association of the United Kingdom, 87, 345-352.

Leaper, R., Douglas, G. and Papastavrou V. (2000) Results of passive acoustic surveys for odontocetes in the Southern Ocean. J. CETACEAN RES. MANAGE., 2, 187-196.

Leaper, R., Chappell, O. & Gordon, J. (1992) The Development of Practical Techniques for Surveying Sperm Whale Populations Acoustically. REP. INT. WHAL. COMMN, 42, 549-560.

Lewis, T., Gillespie, D., Matthews, L.J., Danbolt, M., Leaper, R., McLanaghan, R. & Moscrop, A. (2007) Sperm whale abundance estimates from acoustic surveys of the Ionian Sea and Straits of Sicily in 2003. Journal of the Marine Biological Association of the United Kingdom, 87, 353-357.

Quinn, T.J. (1979) The effects of school structure on line transect estimators of abundance Contemporary Quantitative Ecology and Related Ecometrics (eds G.P. Patil & Rosenzweig), pp. 473-491. International Co-operative Publishing House, Fairland, MD, USA.

Reeves, R.R. & Sciara, a.G.N.d. (2006) The Status and Distribution of Cetaceans in the Black Sea and Mediterranean Sea. pp. 137. Monaco.

Royle, J.A. (2008) Hierarchical modeling of cluster size in wildlife surveys. Journal of Agricultural Biological and Environmental Statistics, 13, 23-36.

Watkins, W.A. & Schevill, W.E. (1977) Sperm Whale Codas. Journal of the Acoustical Society of America, 62, 1485-1490.

Watwood, S.L., Miller, P.J.O., Johnson, M., Madsen, P.T. & Tyack, P.L. (2006) Deep-diving foraging behaviour of sperm whales (Physeter macrocephalus). Journal of Animal Ecology, 75, 814-825.

Whitehead, H. (2002) Estimates of the current global population size and historical trajectory for sperm whales. Marine Ecology-Progress Series, 242, 295-304.

Whitehead, H. (2003) Sperm whales: Social Evolution in the Ocean. The Univerisity of Chicago Press, Chicago and London.

Whitehead, H. & Weilgart, L. (2000) The sperm whale: social females and roving males. In Cetacean societies: field studies of whales and dolphins. Mann, J., Connor, R.C., Tyack, P.L., Whitehead, H. (eds) University of Chicago Press, Chicago. 154-172.

Zimmer, W.M.X., Madsen, P.T., Teloni, V., Johnson, M.P. & Tyack, P.L. (2005) Three-dimensional beam pattern of regular sperm whale clicks confirms bent-horn hypothesis. Journal of the Acoustical Society of America, 117, 1473-1485.

# 6. GENERAL DISCUSSSION

The main objectives of this study were to quantify the habitat of the sperm whales along the Hellenic Trench, identification of the environmental covariates determine their distribution and an abundance estimate of the animals in the area to be retrieved.

The sperm whales that inhabit the area of the Hellenic Trench, and a constant summer distribution were observed throughout the 12 years of monitoring surveys which carried out by the PCRI. Both social groups and solitary males were found throughout the study area.

Long-term monitoring studies are crucial for understanding the distribution and dynamics of long-lived species such as sperm whales. Such time series are necessary to minimize the influence of unwanted randomness on inferences of spatial use and abundance. Such spatiotemporally high effort is important especially in the case of small populations under threat.

## 6.1 Distribution of sperm whales along the Hellenic Trench

Sperm whales are found in the deep waters relatively close to shore, where the seabed is steep and present a preference for higher SST and SLA values. Sperm whale social groups and solitary males co-exist in the area of the Hellenic Trench.

The distribution of the two different group types seems to differ slightly in the same narrow geographical area, with males sperm whales being closer to the shore where the slope is steeper. Social groups have a more restricted distribution than males.

Some parts of the Trench present higher sperm whale densities. These are the southwest part of the Crete Island, the west of Zakynthos Island and areas along the 1km isobath between these two areas. Sperm whales (both social groups and solitary males) seem to be absent from the north of Leukada and the southeast of Ierapetra (South Crete)

When trying to assess the habitat quality for a species, behavioural status should be incorporated in the model so that the habitat can be linked with specific behavioural state. It's worth noting that breeding and feeding grounds could be assessed differently, as animals may use differently distinct habitats.

## 6.2 Modelling long time data series & autocorrelation in habitat studies

Long time series of distributional data can differ in the amount of effort across the study area. In our study the first five years were restricted in the area SW of the Island of Crete and from 2003 onwards the study was extended until the Ionian Sea. Hence, by incorporating the latitude and longitude as variables in the model, this heterogeneity of effort is accounted for and predictions from such models are location-specific. Autocorrelation exists in longitudinal data (Liag & Zeger, 1986) in distributional and habitat studies data are spatial or/and time autocorrelated. Through the use of GEEs the autocorrelation in the data is taken into account by the correlation matrix and independence between the different autocorrelated dataset can be assumed.

## 6.3 Abundance estimation

An abundance estimate of 27 whales was derived for the study area of Hellenic Trench (95% confidence intervals) [19.7, 32.08] from a combined acoustic and visual survey. The Ionian Sea, northern part of the Hellenic Trench has been previously suggested to have high sperm whale densities (Lewis *et al.,* 2007; Gannier *et al.,* 2002). Acoustic surveys have been widely used and have been confirmed as the most appropriate way for studying sperm whales due to their generally high vocal nature and their long-lasting feeding dives. An acoustic effective radius of 14 to 21 km derived from this analysis, containing the size of the sperm whale group as a covariate of detection. These estimates of effective radius were larger than those of previous studies. A possible explanation is that the data and recorded distances derived from a combined acoustic – visual survey where the distances can be more accurate than only from acoustic surveys.

Distance sampling is the most commonly used method for cetaceans. The use of a regression method for modelling the detection function allows different covariates that could potentially influence the detectability of the animals to be incorporated through model selection. The point sampling method that has been widely used in vocal animals, especially for song-birds (Buckland, 2006), and was adopted as a method for the abundance estimation of sperm whales as well. A bootstrapping method for extracting the variance estimates has an advantage to point variance estimates, because the whole abundance procedure is repeated and the uncertainty contained in the estimated detection function, group size, number of detections and in all the products of those estimates is been considered simultaneously.

## 6.4 Conservation status

With the mitochondrial DNA evidence suggesting that the sperm whale sub-population is isolated from the Atlantic one (Engelhaupt *et al.,* 2009) and with the intensity of human activities in the semi-enclosed Mediterranean basin, it is crucial to protect this vulnerable and isolated sub-population. The total population of the sperm whale in the Mediterranean Sea is believed to be in the few thousand’s (Reeves & Notarbartolo, 2006) shows high levels of spatial patchiness because of their social grouping nature and their preferred habitats. Thus, there are priority areas with high population densities that should be part of a more wide conservation plan along the whole Mediterranean basin.

The southwest Crete and the area of Hellenic Trench (Greece) has been characterised as important for the sperm whales area by ACCOBAMS (Agreement on the Conservation of Cetaceans of the Black Sea, Mediterranean and contiguous Atlantic Area). This area has been suggested by ACCOBAMS for a designed MPAs area. Acoustic surveys in the whole Mediterranean (Gannier *et al.,* 2002) and in the eastern Mediterranean (Lewis *et al.,* 2007) have confirmed that the high sperm whale density in the Ionian Sea in addition to the long-term monitoring studies of the PCRI comprise a basis of knowledge for this sub-population. Our study found that at every point in time there are 27 sperm whales along the Hellenic Trench which is a big proportion of the 62 animals found in the Ionian Sea (Lewis et al. 2007).

Incidents like the mass sperm whale stranding that occurred in 2009 in the Adriatic Sea (Mazzariol *et al.,* 2011), resulted in the death of 9 mature males, one of which belonged to the eastern sub-population, could influence dramatically the state of the population. Furthermore, it has been suggested that the Mediterranean sperm whale sub-population is further divided in western and eastern components with little movement between them (Frantzis *et al.,* 2011). Taking that into consideration together with the high mortality of males, gives a clear idea of how vulnerable the population is. In addition, the high juvenile mortality in the Hellenic trench (Frantzis *et al.,* submitted), combined with the characteristic low reproductive rates of large mammals, mean that it is important to minimize any additional anthropogenic burden on this population.

Sperm whales in the Mediterranean are impacted by a number of other anthropogenic threats (Lewis *et al.,* 2007): Collisions, noise and water pollution, interaction with fisheries using gear such as driftnets (Tudela *et al.,* 2005) and plastic objects floating in the sea that end up in the digestive system of the animals and kill them.

The Mediterranean Sea is one of the busiest shipping areas in the world after the Panama Canal, connecting the Atlantic Ocean with the Indian Ocean through the Suez Canal and the Straits of Gibraltar. Monitoring whale presence, distribution and identifying high-density areas is important in order to apply suggestions such as decreasing ship speed while crossing through high whale presence areas (Panigada & Leaper, 2009) or moving ferry routes outside of proposed marine protected areas, like the area of Gibraltar (Canadas *et al.,* 2004). Noise pollution due to shipping traffic and oil exploitation remains an issue for the underwater environment for the marine species. Increased noise pollution could potentially cause a shift in habitat use towards areas that are less optimal for the energetic requirements and for specific behavioural use, such as breeding.

In order to sustain the population size, it is crucial to identify the high sperm whales density areas, the specific use for each of them and the additional areas that are important for the movement between those areas allowing genetic and information mixing between the different groups due to the highly cultural character of these species. The Mediterranean Sea sub-population being a closed one constitutes a unique occasion among other sperm whale populations. Future studies should consider investigating in the scale of the whole Mediterranean basin, since that would give us a more thorough view for the species.

## 6.5 References

Buckland, S.T. (2006) Point-transect surveys for songbirds: Robust methodologies. Auk, 123, 345-357.

Embling, C.B., Gillibrand, P.A., Gordon, J., Shrimpton, J., Stevick, P.T. & Hammond, P.S. (2010) Using habitat models to identify suitable sites for marine protected areas for harbour porpoises (Phocoena phocoena). Biological Conservation, 143, 267-279.

Engelhaupt, D., Hoelzel, A.R., Nicholson, C., Frantzis, A., Mesnick, S., Gero, S., Whitehead, H., Rendell, L., Miller, P., De Stefanis, R., Canadas, A., Airoldi, S. & Mignucci-Giannoni, A.A. (2009) Female philopatry in coastal basins and male dispersion across the North Atlantic in a highly mobile marine species, the sperm whale (Physeter macrocephalus). Molecular Ecology, 18, 4193-4205.

Frantzis, A., Airoldi, S., Notarbartolo-di-Sciara, G., Johnson, C. & Mazzariol, S. (2011) Inter-basin movements of Mediterranean sperm whales provide insigh into their population structure and conservation. Deep-Sea Research Part Ii-Topical Studies in Oceanography.

Frantzis, A., Alexiadou, P. & Gkikopoulou, K. (Submitted) Sperm whale occurrence, site fidelity and population structure along the Hellenic Trench (Greece, Mediterranean Sea). Aquatic Conservation-Marine and Freshwater Ecosystems.

Gannier, A., Drouot, V. & Goold, J.C. (2002) Distribution and relative abundance of sperm whales in the Mediterranean Sea. Marine Ecology-Progress Series, 243, 281-293.

Lewis, T., Gillespie, D., Matthews, L.J., Danbolt, M., Leaper, R., McLanaghan, R. & Moscrop, A. (2007) Sperm whale abundance estimates from acoustic surveys of the Ionian Sea and Straits of Sicily in 2003. Journal of the Marine Biological Association of the United Kingdom, 87, 353-357.

Liang, K.Y. & Zeger, S.L. (1986) Longitudinal Data-Analysis Using Generalized Linear-Models. Biometrika, 73, 13-22.

Mazzariol, S., Di Guardo, G., Petrella, A., Marsili, L., Fossi, C.M., Leonzio, C., Zizzo, N., Vizzini, S., Gaspari, S., Pavan, G., Podesta, M., Garibaldi, F., Ferrante, M., Copat, C., Traversa, D., Marcer, F., Airoldi, S., Frantzis, A., Quiros, Y.D., Cozzi, B. & Fernandez, A. (2011) Sometimes Sperm Whales (Physeter macrocephalus)

Cannot Find Their Way Back to the High Seas: A Multidisciplinary Study on a Mass Stranding. Plos One, 6.

Panigada, S., Pesante, G., Zanardelli, M., Capoulade, F., Gannier, A. & Weinrich, M.T. (2006) Mediterranean fin whales at risk from fatal ship strikes. Marine Pollution Bulletin, 52, 1287-1298.

Panigada, S. & R., L. (2009) Ship strikes in the Mediterranean Sea: assessment and identification of conservation and mitigation measures. J. CETACEAN RES. MANAGE.

Reeves, R.R. & Sciara, a.G.N.d. (2006) The Status and Distribution of Cetaceans in the Black Sea and Mediterranean Sea. pp. 137. Monaco.

Tudela, S., Kai, A.K., Maynou, F., El Andalossi, M. & 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.