UNIVERSITÀ DEGLI STUDI DI GENOVA

DIPARTIMENTO DI SCIENZE DELLA TERRA, DELL'AMBIENTE E DELLA VITA

Corso Di Laurea Magistrale in Biologia ed Ecologia Marina

Implementation of an acoustic protocol on whale watching vessels for size estimation of sperm whales off São Miguel, Azores

Relatore interno:
Prof. Giorgio Bavestrello
Relatori esterni:
Prof. Gianni Pavan
Dott. Laura González García
Correlatore:
Prof.ssa Marzia Bo

Candidato:
Cristina Marcolin

Anno accademico 2020/2021

Abstract

Sperm whales, the largest toothed whales in the world, are known to use the most powerful sound in nature to search for their preys, the echolocation clicks. These clicks have a particular multipulsed nature, with decaying and equally spaced subsequent pulses. The time distance between two pulses (inter-pulse interval, i.e., IPI) has been related to the size of their sound generator organ, the spermaceti. Different equations have been then developed to relate the IPI to the total length of the animal. A similar structure is found also on coda clicks, sounds used for communication purposes. The aim of this study was to implement an acoustic protocol for the size estimation and size structure (thus, age structure) of sperm whales' population off São Miguel Island, Azores. The approach was based on data collection from vessels of a whale watching company. The protocol was based on data collection during whale watching trips using a single hydrophone. Echolocation clicks and codas were recorded in the area behind the animals or at the point of fluke-up. Two years of recordings have been collected in the south of the island. An implementation of the protocol was done from December 2020, with higher sampling frequencies and the area behind the animal as the final point of recordings. Three different approaches were then tested for the IPI measurement: 1) manual method, using Raven for spectrogram and waveform display; 2) an automatic method with PAMGuard software IPI plugin using cepstral analysis; and 3) an automatic method for testing purposes with CABLE software. The manual method revealed to be highly precise, as found in previous studies, but highly time consuming. The IPI plugin was found reliable and required less effort from the operator. CABLE was found not adapt for the dataset used. The size and age structure found with the study were consistent with previous assessments made in São Miguel, but this is the first case with an acoustic approach from whale watching vessels. Whale watching is an increasing touristic activity in all the Azores islands. Thus, an implementation of a rigorous acoustic protocol can potentially lead to population size and age structure of the sperm whale's population found in the waters around the Azores.

Table of Contents

1 Introduction

1.1 The Azores and the history of whaling

The Azores is an archipelago of nine volcanic islands located in the mid North Atlantic Ocean, 1584 km from the nearest point on the mainland (Cabo da Roca, Portugal) (Figure 1.1). They are crossed by the Mid Atlantic Ridge (MAR) and present a typical bathymetry of volcanic islands. The islands are divided in three groups: Western group (Corvo and Flores), Central group (Graciosa, Faial, Pico, São Jorge and Terceira) and Eastern group (São Miguel and Santa Maria).
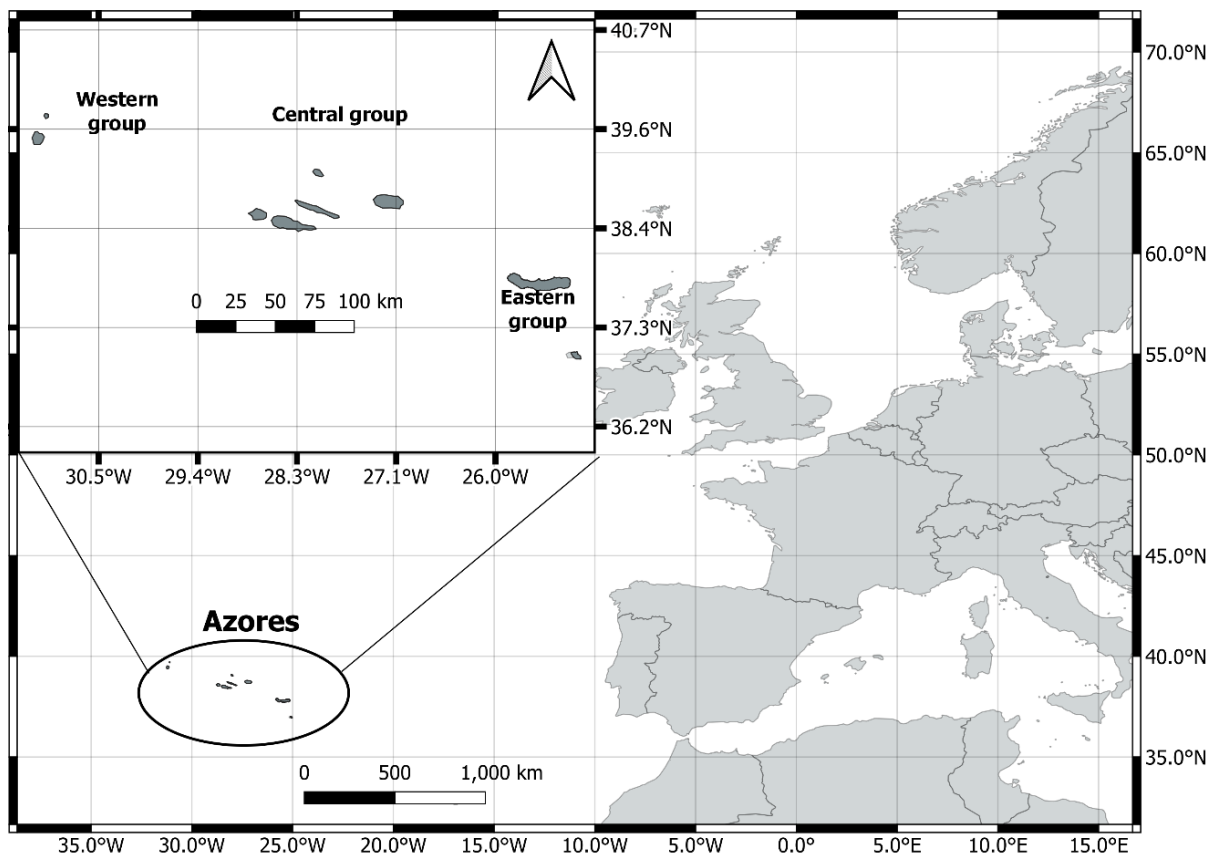
Figure 1.1: Position of the Azores in relation to western Europe. Archipelago enlarged on the top left panel. Data source: Eurostat/GISCO.

The Azores islands are mainly surrounded by deep waters (> 2000 m), with shallower waters (< 200 m) only near the coastline due to the absence of a real continental shelf (Gallagher et al, 2013, Silva et al., 2014; González García et al., 2018). The islands are continuously located on top of the triple junction of three tectonic plates (North America, Eurasian and African plates) and across the MAR (Santos et al., 1995). The volcanic and tectonically active nature of the seafloor creates a complex and heterogenic oceanic region, characterized by a large number of seamounts (Morato et al., 2008a). Four hundred seamounts and seamounts-like formation were discovered in the Economic Exclusive Zone of the Azores, covering around 37% of the area (Morato et al., 2008a). Seamounts are characterized by a wide range of depths, slopes and base area, leading to a large assortment of habitats suitable for many different species (Morato et al., 2008b).

The Azores are under the influence of the two main branches of the Gulf Stream: the North Atlantic Current (NAC), cold water mass passing north of the islands, and a warmer one, the Azores Front /

Azores Current system (AF/AC), flowing eastward south of the islands (Santos et al., 1995; Silva et al., 2014; Caldeira & Reis, 2017; Gonzalez García et al., 2018). The mixed layer is around 150 m deep during winter due to the density differences between the water masses and the stronger winds that characterize the islands in this season. Increasing temperatures within spring-summer seasons trigger the development of a thermocline between 40 and 100 m deep that isolate the surface layer, decreasing the availability of nutrients but increasing the light penetration (González García, 2019). This seasonality of temperature in the water column affects the concentration of nutrients in the upper layers and it is reflected by the highest abundances of phytoplankton found during springtime (the spring bloom) and lower during the summer (Rogerson et al., 2004; Silva et al., 2014)

A narrow continental shelf and deep waters near the shore characterise all the islands (Silva et al., 2014; González García, 2018). The interaction between the complex bathymetry and a dynamic oceanography creates a great variety of habitats. providing great opportunities to host ahigh biodiversity in the area (Silva et al., 2014; González García et al., 2018; Afonso et al., 2020). Twenty-eight cetacean species have been listed in the Azores, occurring all year round, migrating or sighted occasionally (Silva et al., 2014). Common dolphins, bottlenose dolphins, Risso's dolphins and sperm whales can be spotted year-round. All the other species can be spotted only in some periods of the year, with a certain pattern, such as Atlantic spotted dolphin (*Stenella frontalis)* and blue whale (*Balaenoptera musculus*), or just occasionally, without a defined scheme, such as killer whale (*Orcinus orca*) or several species of beaked whales (Hartman et al., 2008; Silva et al., 2014; González García, 2019; Afonso et al., 2020).

From the beginning of the XVIII century, the offshore whaling was already a well-established practice. American whalers targeting sperm whales started to expand their hunting grounds[1]. In 1765 the first attempt to catch sperm whales in the Azores was made. American whalers soon discovered that this area was profitable. By 1768, Horta, Faial, was the port of call in Azores for 200 whaling ships each year (Clarke, 1954; IWC, 2018). The islands of Faial, Flores, São Miguel and Terceira were used by whalers to recruit locals as crew members, to procure supplies for the long trips and to stock the spermaceti oil. In Horta, around 4000-5000 barrels of oil were discharged at the harbour every year (Clarke, 1954).

In the American whaling vessels, Azoreans learnt whaling skills and soon started their own shore-based whaling, the first large-scale commercial fishery in the area (Clarke, 1954; Prieto et al., 2013). It started from the village of Lajes, Pico Island, in 1867 (IWC, 2018a). An important role in this type of whaling was covered by the lookouts, called "Vigias" in Portuguese. They were experienced observers searching for sperm whales with powerful binoculars from towers settled on top of high cliffs around the islands (Prieto et al., 2013; IWC, 2018a). Hand harpoons and canoes were used during all the years of hunting. The efficiency increased during the 20th century to keep up with the greater demand of sperm whales' oil from the new industries (Clarke, 1954; Prieto et al., 2013; IWC, 2018a). The peak was reached in 1948, and in the following years the total number of captures declined because of lower demand of oil (Brito, 2008; Prieto et al., 2013).

The review of Brito (2008) found that 23,525 sperm whales were captured from 1896 to 1987 in the Azores ground. Prieto et al. (2013) showed a slightly different value due to discrepancies of the sources, with 23,557 individuals captured. These numbers considered only the shore-based whaling, because the number of captures from offshore hunting is unknown (Clarke, 1954; Brito, 2008; Prieto et al., 2013). The methodology remained artisanal and the efficiency of processing and exporting whale products was limited. For these reasons, the Azores ground was not considered overexploited (Neves-Graça, 2004; IWC, 2018a). Nevertheless, the oil demand decreased gradually after 1950s, and in 1982 Portugal decided to support the International Whaling Commission's moratorium on whaling activities, which

---

[1] Referring to sperm whales, grounds are area of higher density of individuals. Term was given by whalers (Whitehead, 2003).

came into force in 1986 (IWC, 1982). In 1987 the three last captures were made in the Azores (Vieira and Brito, 2009).

In 1930s, the first whaling restrictions for whale protection included size limitations and overall limit of captures (Gambell, 1993; Scheiber, 1998). In 1946, in Washington DC, the International Whaling Commission (IWC) was created. Certain species were declared protected, specific areas were designed as whale sanctuaries, a maximum number of individuals and a minimum size was set per season, and a closing period for whaling was also established. In addition, whalers were forced to write catch reports with statistical and biological data. Within the next year, 1947, the IWC created a committee of scientists to set regulations with a solid scientific base. In 1972 catch limits were set by species and observers were stationed to control the whaling activity of each country. In 1982, there was almost no agreement between the scientists on the catch limits and this helped to apply a moratorium to the commercial whaling, active from 1986 (Gambell, 1993; Clapham and Baker, 2002). Japan, Norway and Soviet Union rejected to support the decision and in 1987 they started the "scientific purpose" whaling, together with Iceland (Clapham and Baker, 2002).

In the Azores, the Regional Legislative Decree nr. 2/83/A of 2nd March 1983 was the first regional law applied for cetaceans' protection. Capture of dolphins in the area of the Azores was banned. A new Regional Legislative Decree, the nr. 15/2012/A, implemented the protection to all the species of cetacean in the Azorean's waters.

With the loss of the primary source of sustenance, the need of finding a new economical driver rise. In 1987 the International Fund for Animal Welfare (IFAW), financed by the European Community and the Regional Government, came to the Azores with the purpose to explore the possibility to settle a new touristic activity, the whale watching (Silva, 2015; IWC, 2018A). Since the Moratorium of the IWC, the whale watching tourism increased in the coastal areas around the world, as a new form of marine ecotourism. At the same time, in late 1980s, tourism in the Azores was increasing (IWC, 2018A). In 1989, the first whale watching company started in the village of Lajes (Neves-Graça, 2004; IWC, 2018a). Soon after, in 1993, whale watching was also established in São Miguel Island. Since then, the number of tourists increased exponentially in Lajes, from 50 in 1991, to 4000 in 1997. An estimation of 12000 whale-watchers was done for 2011 (Silva, 2015). In the decade 1998-2008 the mean annual growth rate was about 15.5% in all the Azores (O'Connor et al., 2009). From 468 whale-watchers registered in 1992 in one company, the whale-watchers became 48000 in 2011. In 2018, according to the Regional Board of Tourism, more than 100000 whale-watchers were estimated in the Azores (González García, 2019). An assessment done in 2007 by "Serviço Regional de Estatística dos Açores" (SREA) found that 12.5% of the archipelago tourism was primarily for whale watching trips (O'Connor et al., 2009).

With the rise of the number of boats and so the number of tourists, the need of a strict regulation became evident (Neves-Graça, 2004). In this perspective, the first Azorean Biannual Conference for Whales and Dolphins was held in Lajes in 1998. The rules agreed were the baseline for the first Regional Legislative Decree related to the observation of cetaceans, 9/99/A of 22nd March. An update was done with the Regional Legislative Decree n°. 10/2003/A. The main rules apply to the minimum distance from the animals, maximum time with the animals, the maximum speed of approach and the direction of approach, both for single boats and groups of boats.

In the international environment, the World Cetacean Alliance has compiled a manual for responsible whale watching, the "Global best practice guidance for responsible whale and dolphin watching" (WCA, 2018). This report is meant to be a guide for all the tour operators around the world, in order to maintain the balance between running a business and protecting cetaceans.

All around the world, whale watching activities are increasing. In 1998, 9 million of whale watchers have been estimated in 87 countries. In 2008, the number increased to 13 million of whale watchers in 119 countries (Vinding, 2015). The most updated data is related to the European countries: from 1.440 million in 2008, whale watchers became 1.800 million in 2018. It corresponds to a growth of about 20.1% in 10 years, while in the previous decade was about +14.6 % (Hoyt, 2021). These number can give an idea of the potentiality of the use of this activity for data collection.

1.2 Sperm whales in the Azores

1.2.1 Physical description

Sperm whales are part of the infraorder Cetacea. Cetaceans started their way back to the ocean around 54 million years ago (Ma) and around 25-35 Ma the Mysticeti (or baleen whales) suborder diverged from the Odontoceti (or toothed whales), the group to which sperm whales belong (Whitehead, 2003; Fahlke, 2015). Odontocetes are characterized by the presence of teeth (Fordyce, 2001), a cranial leftward asymmetry as adaptation for the nasal apparatus and a fat-filled melon for sound production; the ability of hearing the ultrasounds (Fahlke, 2015); and sexual dimorphism with males larger than females, with few exceptions (Fordyce, 2001). In contrast, mysticetes are filter-feeding animals (Fordyce, 2001) with no cranial asymmetry and melon; they instead use the larynx for sound production, and they can hear in low and infrasonic frequencies (Fahlke, 2015); great dimorphism is present too, but in favour of females (Fordyce, 2001).

Sperm whales are distinct from the other odontocetes, being part of the lineage Physeteroidea, that is considered the eldest of all the extant cetaceans. Apart from sperm whales, *Kogia breviceps* and *Kogia sima* are part of the group (Whitehead, 2003; Boersma and Pyenson, 2015). They have common macro-morphological characters of skull, as a severe left asymmetry in the nasal area and a large supracranial basin for the spermaceti organ. Sperm whales have a maximum length around 18 m, while *K. breviceps* and *K. sima* can reach 2.7 m and 3.5 m respectively (Boersma and Pyenson, 2015).

As a unique species among cetaceans, sperm whales have a distinct profile (Werth, 2004), interrupted around the head by a crease at the conjunction of the head with the trunk (Figure 1.2). This feature is prominent in the largest males, because of the bigger spermaceti organ. In cranio-caudal direction, the profile is quite straight till the dorsal fin, low and usually rounded; in some mature females it can have a whitish, yellowish callus, which is a secondary sexual character. After, the back features a series of wrinkles till the peduncle (Whitehead, 2003; Whitehead, 2017). The fluke is flat and triangular, but of relevance is the trailing edge, where a pattern of scars, bites and notches can be found: these features are a reliable tool for photo-identification when the whale raises the fluke before diving (Whitehead, 2003; Whitehead, 2017). Behind the eyes there are two small and paddle-shaped flippers (Whitehead, 2003). The lower jaw, rod-like and narrow, can reach 5 m (Werth, 2004) and present 20-26 pairs of conical teeth, which will erupt near puberty. As sperm whales are supposed to feed with the suction feeding technique, as other odontocetes (Werth, 2004; Fais, 2016), teeth are not useful and healthy individuals were found without teeth or even lower jaw (Whitehead, 2017).

The coloration is mostly dark grey, but white areas are present around the line of the lower jaw (and can be used for individual identification) and some white patches can be found on the belly. Especially mature males can have larger white patches around the head. After the eye, large corrugations cover most of the body (Whitehead, 2003; Whitehead, 2017). The front of the head is usually full of scars, signs left from the prey (Whitehead, 2003).

Sperm whales are, indeed, the animal of extremes (Whitehead, 2003; Whitehead, 2017). The size make sperm whales the largest toothed whale (Werth, 2004, Whitehead, 2003) and the cetaceans with the

highest sexual dimorphism, as mature males are about 1.5 times longer and 3 times heavier than females (Whitehead, 2003; Steiner et al., 2012; Whitehead, 2017). On average, males are about 16 m and 45 t, while mature females are around 11 m and 15 t (Whitehead, 2003; Whitehead, 2017). Males grow till they reach the age between 35-60 and they usually reach the maximum length (18 m) around their forties (Whitehead, 2003). However, the largest male ever recorded was 24 m, caught in the South Pacific Ocean in 1933, according to the IWC data. Moreover, the jaw stored in the Natural History Museum in London seems to belong to a 25.6 m long individual (McClain, 2015).
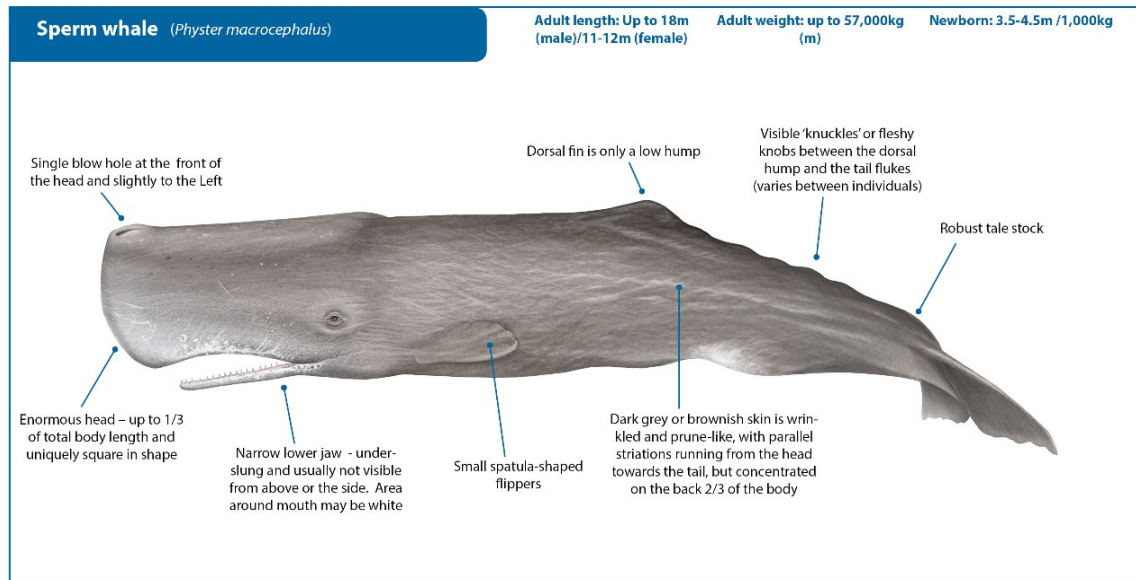

Figure 1.2: Sperm whale infographic about its morphological features. From IWC (2018), Whale watching handbook.

Another record is about their brain: weighing on average 7.8 kg, it is the largest brain of any animal. Recent findings (Ridgway and Hanson, 2014) showed that the record is shared with killer whales (*Orcinus orca*). Nonetheless, it is not surprisingly large compared to sperm whales body size, though, no direct information about their cognitive abilities is available (Whitehead, 2017).

Figure 1.3: a) Detail of the front-left blowhole of a sperm whale. b) The oblique blowhole of a sperm whale. Photographer: Cristina Marcolin

A unique feature of this species is the blowhole. Unlike mysticetes, odontocetes has only one nostril. Sperm whales are no exception and the blowhole, S-shaped, is at the top front-left of the head (Whitehead, 2003; IWC, 2018b). The unicity is even more marked by the blow itself, highly recognisable at a distance because of 45° inclination (respecting to the body axis) to the left (Figure 1.3).

1.2.2   Population structure

P. macrocephalus has the widest distribution around the world, second only to the one of *Orcinus orca*. *The sperm whale* can be found in all the ocean and sea basins, except for Black Sea and Read Sea. Sperm whales tend to live in oceanic areas (Antunes, 2009; IUCN, 2021).

Habitat preference analysis for sperm whales have found continental slopes and temperature as the main predictors of their distribution. In the break of continental platform, higher concentrations of their favourite preys (cephalopods) are found (Correia et al., 2021). In particular, in the Azores Clarke et al. (1993) found that genera Histioteuthis, Chiroteuthis and Ancistrocheirus, general small species, were the favourite prey (68% of the beaked found in stomachal content). Their data suggested also that they feed on larger squids (such as the giant squid, Architeuthis sp.), supported by previous observation of vomited preys and stomach contents (Kawakami, 1980; Clarke et al., 1993; Paxton et al., 2016). Although sperm whale's habitat is mainly oceanic, along oceanic islands they can be also found close to shore (Antunes, 2009; Correia et al., 2021; IUCN, 2021). Moreover, complex underwater structures such as canyons or trenches are highly recognised as enhancer of primary production and so of the food web till cetaceans (Morato et al., 2008b; Tepsich et al., 2014).

The life of a sperm whale begins after a gestation of around 14-16 months. Once born, they weigh around 1 t and measure around 4.0 m. Births are always single, with twins found rarely by whalers, but it is unknown if, eventually, they can both survive (Whitehead, 2003). Calves usually suckle milk till 2 years old, but this limit is flexible: solid food in the stomach can be found from 1 year of age, and individuals of 13 years old have been found with milk in the stomach (Whitehead, 2003; IWC, 2018b).

Juveniles of both sexes are active in the social life of the unit, but at a certain point their behaviour diverge. Females usually stay in the mother's social unit for the entire life (Whitehead, 2003). These groups of individuals of different ages and sex are nomadic and, inside a home range, they can travel up to 35.000 km per year (Whitehead, 2003, p. 12; Whitehead, 2017, p. 920; IWC, 2018b). They inhabit warm tropical and temperate waters, usually deeper than 1000 m and lower than 40° latitude, correspondent to a mean temperature of 15 °C (Whitehead, 2017; IUCN, 2021). Steiner et al. (2015) found that female sperm whales make limited latitudinal migrations of around 1500km in the North Atlantic. Photo-identification matches suggested that they may stay in the Azores during the summer and migrate to the Canaries for the winter months. Females became sexually mature at around 9 years old and they will give birth to one calf every 4-6 years (Whitehead, 2003). According to Evand and Hindell (2004), females of sperm whale reach their maximum length (around 11 m) at 19-20 years old. Females share the responsibility to raise the calves inside the social unit. Around 40 years old, they are less fertile, with a rate of one calf every 15 year. Elder females are usually of about 80 years of age, but they can reach a century (Whitehead, 2003). Antunes (2009) found that females and juveniles form persistent social unit of around 12 individuals in the Azores. Silva et al. (2014) stated that social units of females would stay in the Azores for 14.7 on average. van der Linde and Eriksson (2019) describe 12 social units, from 2 to 13 individuals, in the area of São Miguel from 2010 to 2017. Unlike other areas, as in the Pacific basin, different social units do not form temporary groups (Antunes, 2009). This was consistent with the findings of van der Linde and Eriksson (2019), in which association within units was statistically more significant than within units. Long term associations were found to be consistent, with the longest recorded of about 19 years, at least (Antunes, 2009). Moreover, Antunes suggested that different social units are rarely spotted in the same area together.

Males leave the mother's unit in between 4 to 21 years old. (Whitehead, 2017). In the northern hemisphere it is thought that they abandon the group when they are sexually mature (Steiner et al., 2012). Loose groups of males, called "bachelor groups", does not have the same complexity as the female's social unity. Moreover, as the individuals age, the groups become smaller, and they venture farther

(Whitehead, 2003; Antunes, 2009; IWC, 2018b). The largest males, in their forties, can be seen alone around the ice edges in both hemispheres (Whitehead, 2017). Males become sexually mature in their teens, though they seem to have an active reproductive role around their twenties, when they start to migrate back to the warm waters' ground of females (Whitehead, 2003). These long journeys allow the males to feed in productive areas, such as around Norway for the northern hemisphere.

The oldest data available regarding the presence of sperm whales in the Azores come from the whaling period. As stated by Clarke (1956), the captures recorded are mostly of mature males. In the period 1939-1954, with base at Horta, the mean fraction of mature males was of 64.6%; in the period 1948-1954, aggregated data for all the Azores showed a percentage of mature males of 68.82%. Clarke assumed that whalers usually would have preferred a mature male, for its bigger size per catch effort; moreover, in other part of the world, females and calves were already protected, so the sex ration in catches are not reliable for population structure. Clarke also found a seasonality: in the monthly catches it was evident that the percentage of males (1939-1954 from Horta) was always greater but was reduced in spring and summer time; in August sex ratio was quite equal and from September on it increases again. Therefore, Clarke supported the hypothesis of a migration from and to the islands during the year, both for females and males. The higher catches in summertime were supported by a greater hunting effort and the starting of the calving season. A catalogue of tails photographed between 1987 and 2008 in the Azores were analysed by Steiner et al. (2012). The study found that mature males represented only around 10% of the identified animals. Authors have taken in account also a Norwegian catalogue to study the pattern migration of mature males. In this area, only mature males where sighted. Three matches were found between Azores and Norway, with individuals sighted for the first time in the Azores and never the opposite. The authors hypothesize that individuals photographed in the Azores were juveniles at the time and that they after migrated north. Recently, more extensive photo-identification comparisons in the Atlantic, yielded two more matches, at least one of them of a mature male between the Azores and the Gulf of Mexico, and between the Azores and the Bahamas respectively (Mullin et al., 2022).

Even though seasonal peaks in reproduction occur, there is not a clear pattern of migrations, as it happens for many of the large baleen whales (Whitehead, 2003; Steiner et al., 2012). Monitoring of migration patterns from one ground to another are still rare, and so there is a lack of knowledge about their frequency, duration or extent (Whitehead, 2003; Steiner et al., 2012).

In Silva et al. (2014) there is a better insight on the social structure and occurrence of sperm whales in the Azores. Data from land-based and boat-based surveys carried out between 1999 and 2009 were analysed. Sperm whales were spotted all year long in land-based surveys. The boat surveys were carried on mainly from May to October, hence not covering all year round. The majority of the sightings from the boat (76%) were of females' social units (adult females, juveniles and calves) and 8% were of only adult males, individually or in loose groups. The remaining 16% included both females' groups and adult males. Calves were spotted all year-round but with a higher frequency in August and September. Newborns only in August and March, suggesting a longer reproduction season in winter-spring, in accordance with the findings of González García (2019). Groups of females were sighted while foraging every month.

A more recent study (van der Linde and Eriksson, 2020) had similar findings as the latest research in São Miguel Island, Azores. Photographs were collected on board of whale watching vessels in the period 2010-2017. Only 10.4% of the individuals identified were mature males. Moreover, the majority of the sightings were of females' social units. Sperm whales were encountered during all the year, with a summer peak for females and juveniles. No seasonality was found for mature males. These findings are in contrast with previous studies such as Clarke (1956), but agrees with more recent studies (Steiner et

al., 2012; Silva et al., 2014; González García, 2019). Newborns were sighted during summertime, supporting the idea of the Azores as a reproduction ground and that the mating season is mostly in springtime. A unique history is that of the so-called "Mr. Liable", a male sperm whale considered resident in the Azores. This individual has been resighted during all the years of the study, with 101 different days of resighting (van der Linde and Eriksson, 2020). A previous study (Gardoki et al., 2018) examined the possible residency of this male. The individual was resighted 81 time between 2011 and 2017 off the south of São Miguel.

Resighting of the same individuals was found to be low in many studies. In Matthews et al. (2001), 762 individuals were identified in the period 1987-1995. In the same season, the majority of the individuals were not resighted, and the ones sighted more than one time were females (31% probable females and 12% others). Between different years, the number of resighting was even lower and there were no mature males resighted in different years. Van der Linde and Eriksson (2019) identified 393 individuals and the 62.7% of them were sighted only once. 26.5% of all the individuals were encountered in more than one year and only 1.5% of the catalogued (6 individuals) have been sighted in all the 8 years of the study. Regarding the resighted individuals, the majority were females and juveniles: 39.8% of the females and juveniles identified were seen more than once; for the males, the percentage drops to 14.6%.

In this perspective, Matthews et al., (2001) have done two different estimations of the population size in the Azores. The estimation as a closed population was made referring to all the individuals that might occur in the islands during a season. For a wider range, an open population model was used to refer to individuals that might be around the Azores in different years. Results were of 300-800 females and immatures around the central group islands during summertime; of 400-800 in the years 1988-1990 and of 1600-2200 between 1991-1994. These data are limited because of the small study area taken in consideration, limited only to the central group islands (Matthews et al., 2001). The Azorean population is so suggested to be an open population. Boys et al (2019) estimated a population abundance of 1500 animals during the summer. The resident population is about 20% of the summer estimate. Moreover, they found a continuous flow of individuals, entering and leaving the area in different years.

It is still in debate if there is a unique stock of the North Atlantic Ocean or different stocks can be found (Matthews et al., 2001; NAMCO, 2021). Different attempts have been made to give an accurate estimation of the entire North Atlantic, with the last update in 2007 counting 30000 individuals. Great issues about population estimates of sperm whales are the lack of overlaps in the survey areas and the great chances of bias due to their diving behaviour. (NAMCO, 2021).

1.2.3   Sound production

The most distinctive characteristic of the sperm whale is the nasal complex (Figure 1.4), which extends over one third of the total body length (Whitehead, 2017). The massive spermaceti organ is predominant and can occupy around the 25-33% of the total body length (Møhl, 2001; Whitehead, 2003). It dominates the head and influences the morphology of the skull, jaws and the nasal passages. The organ contains a spongy tissue, immersed in the spermaceti oil. This weird name derives from the whaling era: whalers misunderstood the nature of this oil, believing it was the sperm of the animal (Whitehead, 2003). At its extremes, two air sacs are present: the frontal air sac is present between the spermaceti organ and the skull, while the distal air sac is in the opposite side. The two air sacs are connected by the right nasal passage that is not connected to the blowhole. The sole nostril present let the air inside through the left nasal passage. Beneath the spermaceti organ there is a large mass of connective tissue filled with oil, the junk organ. The air enters the distal air sac through the monkey lips (or 'museau du singe'), a sort of

valve system made of connective tissue that looks like lips. The monkey lips are also connected to the distal sac and the right nasal passage (Whitehead, 2003; Growcott et al., 2011; Caruso et al., 2015).
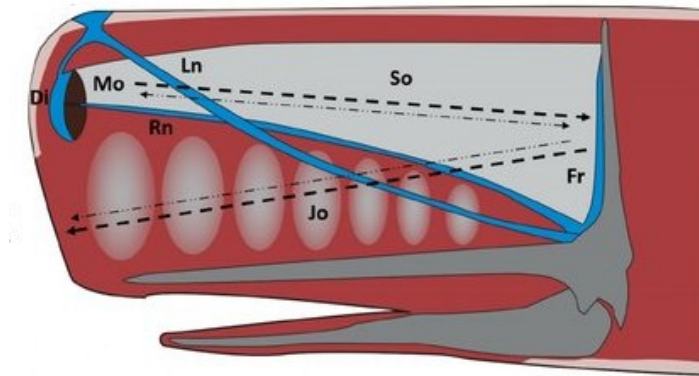

Figure 1.4: Scheme of a sperm whales' nasal complex: Di = Distal air sac; Fr = Frontal air sac; Jo = Junk organ; Ln = Left nasal passage; Mo = Monkey lips; Rn = Right nasal passage; So = Spermaceti organ. Arrows show the path followed by the sound produced trough the monkey lips. Adapted from Caruso et al., 2015.

Different hypotheses have been suggested for the function of the spermaceti organ. A buoyancy role during deep foraging dives was suggested by Clarke (1970; 1978), as well as a "battering ram" role in aggressive interaction with other males (Carrier et al., 2002). The most accredited hypothesis is the one proposed by Norris and Harvey (1972) that described the spermaceti as a sound generator (Whitehead, 2003; Antunes, 2009; Fais et al., 2016).

Indeed, sperm whales produces a variety of sounds, but their vocal repertoire is made of manly click sounds (Madsen et al., 2002a; Whitehead, 2003). However, also tonal sounds and trumpets are produced (Pace et al., 2021). Clicks are "sharp-onset, broadband, impulsive vocalizations with energy between 5 and 25 kHz" (Whitehead, 2003). Clicks are considered the highest biologically source level ever recorded (Madsen et al., 2002a), up to 223 dB (Møhl et al., 2000). They have been discovered as highly directional sounds (Møhl et al., 2000), although they have different patterns with unique characteristics and a variety of uses (Whitehead, 2003). The most recognised and studied are "usual clicks", "creaks", "codas" and "slow clicks", that have different characteristics depending on their main function (Table 1.1).

Usual clicks are the most common pattern, composed by a long repetition (called train) of regularly spaced clicks (Weilgart and Whitehead, 1988). They can last several minutes and have an Inter-Click Interval (ICI), or the time delay between two clicks, of about 0.5-2.0 s (Madsen et al., 2002a; Whitehead, 2003; Growcott et al., 2011; Shabangu and Andrew, 2020). These clicks are involved in echolocation, used during foraging dives to search for preys (Norris and Harvey, 1972; Gordon, 1987; Møhl et al., 2000; Madsen et al., 2002a). These dives are usually deep (400-1200 m) and long-lasting (up to 138 min). However, on average they spend 40-50 min underwater (Watwood et al., 2006; Aoki et al., 2012). Their high directionality was discovered recently (Møhl et al., 2000; Møhl et al., 2003; Zimmer et. al., 2005a).

Creaks, also called buzzes, are sequences of very rapid clicks, with an ICI that ranges between 5 to 100 ms and a short duration of single click (0.1-5 ms instead of 15-35 ms of the other types). Their rapid sequencing and short range make them suitable for short-range echolocation (Whitehead, 2003; Shabangu and Andrew, 2020). In more recent studies, buzzes have been associated with an active role in prey captures, as already seen in beaked whales and porpoises (Fais et al., 2016).

Codas are stereotyped sequences of 2 to 40 clicks, with an ICI of 0.1-0.5 s and a duration of about 0.2-5 s. Codas has a pre-defined pattern, so they can be distinguished in types depending on the number of clicks repeated and the ICI. The main function of this vocalisation is for communication. Indeed, they are mainly used by the members within a social unit, probably to maintain their complex social structure (Madsen et al., 2002a; Whitehead, 2003; Antunes, 2009; Shabangu and Andrew, 2020).

Slow clicks are mainly associated to large mature males for communication purpose. They are characterised by an ICI of 5-8 s and low centroid frequency. In this way, slow clicks are long range clicks, that can be heard by another sperm whale as far as 60 km away (Whitehead, 2003).

Table 1.1: Characteristics and main functions of the four main types of sperm whale clicks (Whitehead, 2003)

Click Type	Apparent Source Level (dB re 1μPa)	Directionality	Centroid Frequency (kHz)	Inter-Click-Interval (s)	Duration of Click (ms)	Duration of Pulse (ms)	Range Audible to Sperm Whale (km)	Primary function
Usual	230	High	15	0.5-1.0	15-30	0.1	16	Searching echolocation
Creak	205	High	15	0.005-0.1	0.1-5	0.1	6	Homing echolocation
Coda	180	Low	5	0.1-0.5	35	0.5	2	Social communication
Slow	190	Low	0.5	5-6	30	5	60	Communication by males

Norris and Harvey (1972) described for the first time the pulsed structure nature of the clicks and proposed a first version of the sound production. However, Madsen et al. (2002b) and Møhl et al. (2003) modified the initial model of the sound production, proposing the *bent horn* model. In this theory, one weak pulse (the P0), exit directly into the water, but contains a small portion of the energy. The majority of the energy is in the subsequent P1 pulse. This pulse is produced when the air is forced through the monkey lips. The pulse travels through the spermaceti organ till the frontal air sac, where it is reflected back. The reflection is partially directed trough the junk, constituted by a series of acoustic lenses that transmit the pulse into the water in front of the whale. Part of the initial pulse travels back and forth again (with the distal air sac as the second reflection surface), creating a second pulse that goes through the junk and out in the water. Clicks can contain three or more pulses regularly spaced (Whitehead, 2003; Antunes, 2009; Growcott et al., 2011; Caruso et al., 2015). Moreover, surface reflection (i.e., reflection of the direct click on the surface of the water) appears with a delayed time after the direct click and it might be or not overlapped with the direct click, mostly depending on the depth of the whale and of the receiver (Teloni et al., 2007; Hirotsu, 2010). Pulsed structure has been recognised in both usual clicks and codas (Pavan et al., 1997; Pavan et al., 2000; Madsen et al., 2002a). However, they differ quite clearly as seen by the decay rate (for peak amplitude) between the successive pulses. Codas tend to have a lower decay rate between pulses and display also a clearer multipulsed structure, with

more than two pulses (Figure 1.5). Codas' less directionality and reduced power allow this different multipulsed structure (Madsen et al., 2002a).
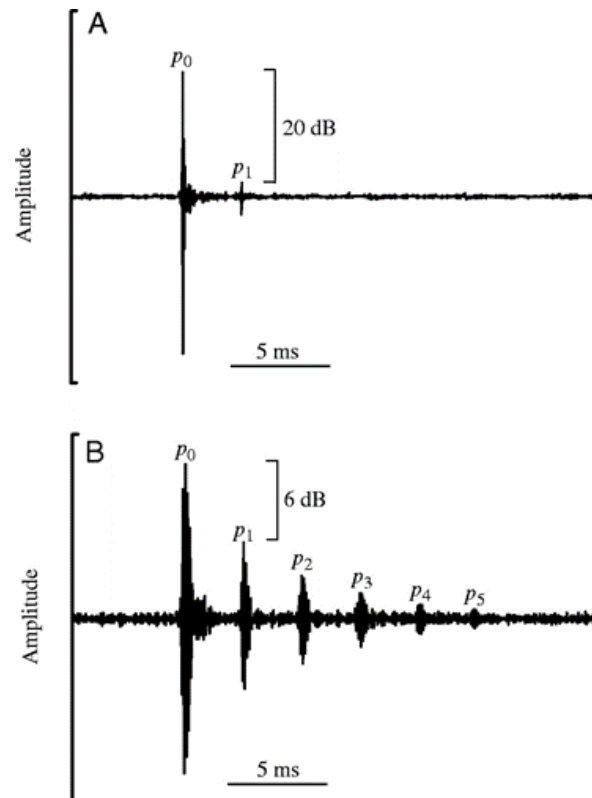

Figure 1.5: Waveforms display of usual clicks (A) and coda clicks (B). From Madsen et al., 2002a.

The bent horn model has led the scientist to develop a way to estimate the size of whale, based on the delay between the pulses. In fact, sound speed through the spermaceti organ and size of the spermaceti organ in relation to the total length of the animal are known. If the delay time between two subsequent pulses (Inter-Pulse Interval, or IPI) is correctly measured, the length of the animal can be estimated (Norris and Harvey, 1972; Gordon, 1991; Goold, 1996).

The first attempt was made by Norris and Harvey (1972) with an easy formula, as they thought the delay time was two times the length of the spermaceti organ. However, new findings changed the model (Madsen et al., 2002b; Møhl et al., 2003; Zimmer et al., 2005a) and the nature of the multipulsed structure of clicks. In fact, Goold and Jones (1995) proposed that the structure may change in different recordings and clicks. Madsen et al. (2002b) and Møhl et al. (2003) further explored this theory, adding that the multipulsed structure is aspect-dependant. Møhl et al. (2003) assumed that recordings can be recorded on-axis (monopulsed) or off-axis (multipulsed) in respect to the acoustic axis. Off-axis clicks make the majority of the clicks recorded. Zimmer et al. (2005b) better explained this property of the clicks. On-axis clicks shows a multipulsed structure that is more reliable for IPI measurements. On the other side, off-axis clicks shows a multipulsed pattern that is confused by interference from the off-axis angle of the whale. Moreover, they described a P1/2 ("pi-half") that may occur whenever in between the P1 and P2 pulse when recorded in front of the whale or between P0-P1 when behind the animal. Recording directly behind (backward clicks) or in front of the animal will increase the chance of on-axis click, with the P0 pulse for the former or the P1 for the latter, as the most intense pulse (Rhinelander and Dawson, 2004; Teloni et al., 2007). This was consistent with evidence of changing sound speed in the spermaceti organ with different pressures. Moreover, temperature has an effect, leading to a non-linear decrease of the sound speed. Goold (1996) discovered also that IPIs tends to decrease with depth.

The IPI estimation can be conducted in different ways: at the very beginning, only manual analysis from the waveform were used (Norris and Harvey, 1972; Gordon, 1991). A step forward was made with introduction of semi-automated method that used the cepstral and autocorrelation analysis (Goold, 1996). Cepstral analysis is a non-linear signal analysis described firstly by Bogert et al. (1963). The power cepstrum was defined as "the power spectrum of the logarithm of the power spectrum of a signal" (Antunes, 2009) and it was introduced for detecting delayed versions of a signal (Antunes, 2009). The autocorrelation analysis is based on the repetitive correlation of the click waveform with its replica progressively time delayed. Peaks in the function occurs when the shift in time between the replica and the original is equal to an integer multiple of the IPI (Goold, 1996). These methods were used subsequently in many different studies, together or separated (Pavan et al., 1997; Pavan et al., 1998; Pavan et al., 2000; Antunes, 2009; Caruso et al., 2015). The problem of these methods is that most of the clicks do not show a clear multipulsed structure, as found by Alder-Fenchel (1980) with only 11% of the clicks available for IPI estimation. Moreover, most of the clicks are recorded at an unknown aspect and clicks are aspect-dependant, so analyses of few clicks are not reliable (Zimmer et al., 2005). Following these findings, Teloni et al. (2007) developed a semi-automated method based on the cepstrum average of a large number of clicks. In the same direction, Miller (2010), created a tool for the software PAMguard, an opensource software developed for acoustic analyses of marine mammals. The plugin is a totally automated method to estimate the IPI based on the cepstrum analysis. Teloni et al. (2007) and Miller (2010) followed the hypothesis that averaging the cepstra of a large number of clicks will lower the effect of the aspect-dependant nature of the clicks, thus estimating the true IPI or stable IPI.

In parallel with these findings and modifications to the bent horn model, different formulas to estimate the total length have been suggested. However, two equations are widely recognised to be reliable.

The first equation was proposed by Gordon (1991), that compared size data from IPI estimations and from photogrammetry.

$$TL = 4.833 + 1.453 \cdot IPI - 0.001 \cdot IPI^2$$

where TL is the total length of the animal (m) and IPI is the Inter-Pulse Interval measured in (ms). This formula is recognised as being reliable for individuals with a total length ≤ 11 m (Madsen et al., 2002b; Growcott, 2010).

The second equation was obtained by Growcott et al. (2011) with another comparison between IPI size estimation and photogrammetry data.

$$TL = 1.258 \cdot IPI + 5.736$$

Where TL and IPI are the same from the previous equation. This was found to be reliable for individuals > 11 m.

1.2.4   Aim of the project

This project was developed to test sperm whale acoustic data collected from whale watching platforms as a useful source of scientific data for sperm whale size estimation. Platforms of opportunity (as in this case, whale watching vessels, have been used for a long time for opportunistic data collection. Nonetheless, working whit this kind of data can be challenging, as the primary goal of these activities is the satisfaction of the tourists, and rigorous application of scientific protocols onboard is not always possible. Acoustics recordings can be even more challenging, since they require a certain amount of time in which tourists are not directly observing the animals (because the animals are underwater) and tour guides play a key role onboard to educate and entertain clients while data is being recorded. However, due to its high potential value, the usage of these data should be tested, and recommendations

to improve the existing acoustic protocol, and make it widely implemented and reproducible on daily whale watching tours for recording of sperm whales would be an important tool for future research in the Azores (and potentially abroad).

Main objectives are:
- Estimate total body length of sperm whales acoustically recorded off south São Miguel Island during whale watching trips between 2019-2021.
- Create a baseline with size structure data for future assessment of the population of sperm whales around São Miguel Island. These data would increase the dispersed knowledge of the population structure in the Azores, mainly based on photo-identification techniques.
- Establish an acoustic protocol for estimation of sperm whale size, easily reproducible and subject of being implemented on daily whale watching trips on a regular basis.
- Provide suggestions and recommendations for the improvement of acoustic data collection from whale watching vessels in the Azores; and for data processing and analyses.

Population structure assessments are required as a complement of other information for a species conservation purposes. Knowing the health of a stock is fundamental to carry on conservation management plans, as errors can lead to failure in evaluating the vulnerability of a particular population. This evaluation, that usually is reported in the IUCN Red lists, is crucial to apply useful management rules and recognise areas that need to be protected. Population assessments are fundamental, especially for species that have been hunted and are now protected, as sperm whales, in order to determine if management regulations are effective (Whitehead et al., 1998; Alves et al., 2013; Evans and Teilman, 2019; NAMCO, 2020).

But why there is the need to protect whales? Different studies have given the answer. Cook et al. (2020) explored the ecosystems services that are provided by whales. Whales are important fertilizers in the ocean: as they migrate during long trips and they migrate vertically to feed, they carry nutrients from the depth to the surface (Roman et al., 2014; Cook et al. 2020). This allows a greater grow of phytoplankton, the first source of oxygen in the planet and a considerable sink of carbon dioxide. By increasing the primary production, whales contribute to the food chain supply. Moreover, they are a sink of carbon when they die and sink in the deep ocean, becoming the primary source of food for the deep species (Roman et al., 2014). Whales are also beautiful creatures to the public imaginary and therefore they are charismatic species, i.e., more appreciated by the public than other species. This aesthetic character is also at the base of the whale watching activity (Cook et al. 2020). Whale watching is an increasing activity all around the world and has a large economic impact (Silva, 2015; González García, 2019). Chami et al (2019) gave a monetary value to the importance of whales. With their supply to the food web chain, and so to the fish stocks, whales contribute to the fishing industry with more than $150 billion. Value of whale watching industry was then estimated as more than $2 billion, globally.

2 Methods

2.1 Study area

Data collection was conducted in the area off the south of São Miguel Island, the eastern-most island of the archipelago. As the other islands, it is characterized by a narrow continental shelf. Shallow waters can be found near-shore, while depth goes to more than 2000 m relatively close to shore (Figure 2.1). Both in the north-west and south-east of the island there are two oceanic trenches that reach more than 3200 m (González García et al., 2018). Surface temperature varies from around 15 °C in wintertime to 25°C in late summer (González García, 2019). NAC and AF are responsible for the oceanographic mesoscale features and have a strong effect on the area, as they transport and retain nutrients, favouring primary production. The island features also local upwelling spots, favoured by strong winds and the bathymetry (Lafon et al., 2004; González García, 2019). A biophysical coupling is usually observed around the islands, denoting upwelling spots by high chlorophyll concentrations and low SST (Lafon et al., 2004; González García, 2019). Mixed layer reaches a depth of around 10 m during summer months (June-Augusts), while from October to March is deeper than 30 m. In winter months (December-February) it can be found from 60 m to 90 m of depth (Coelho, 2021).
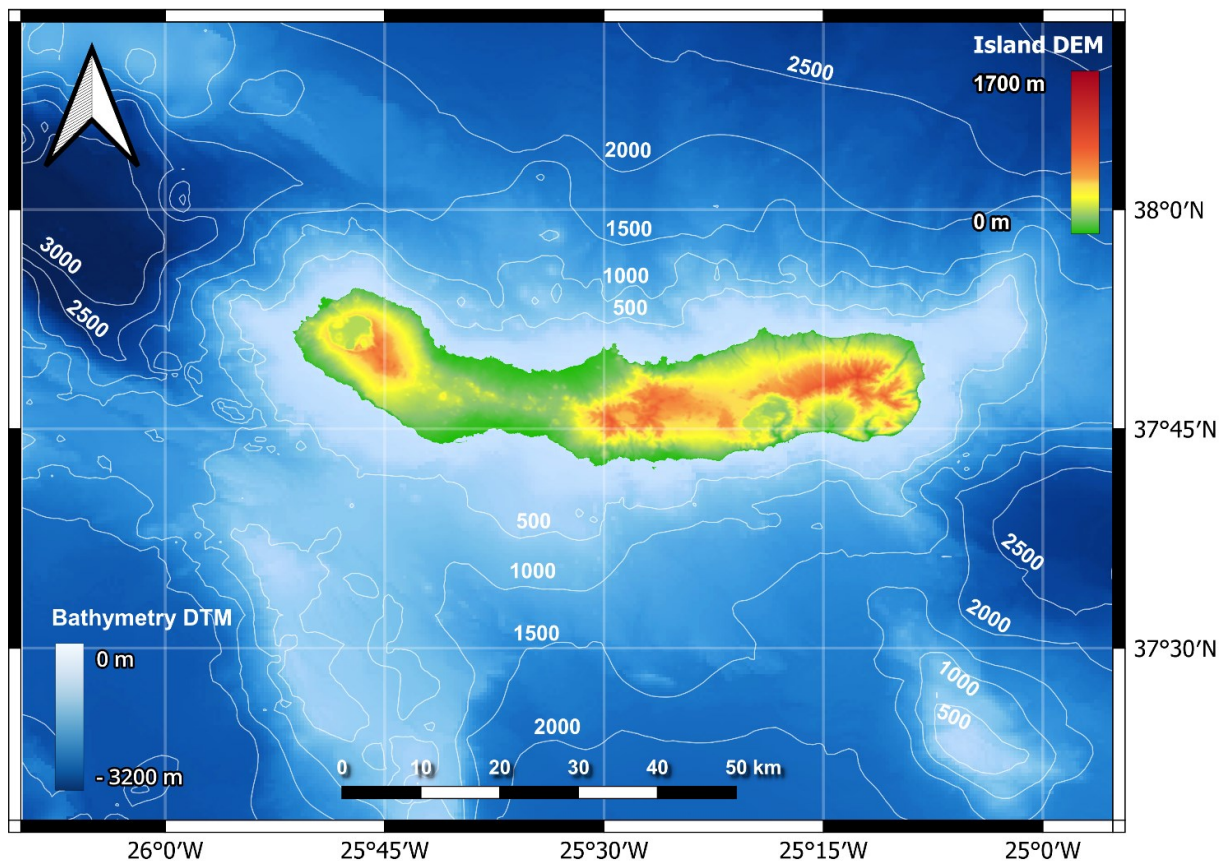

Figure 2.1: São Miguel Island map with bathymetry, isobaths every 500 m; colour legend on the bottom left. DEM of the island with colour legend on top right. Data source: Azores DEM from European GNSS/EU Copernicus programme 2018 Bathymetry DTM from GEBCO 2021 (1 km resolution).

2.2 Data collection

All the data were collected between June-2019 and October-2021 off the south coast of São Miguel Island from whale watching platforms of "Futurismo Azores Adventures". Sperm whales' data consist of: (1) audio recordings of sperm whales; (2) correspondent sighting records, which include species,

location, date and number of individuals present, among other info; and (3) when possible, photographs valid for identification purposes. Data were collected by qualified biologists during whale watching touristic trips (i.e., opportunistic data). Three dedicated trips were done in the period.

2.2.1 Sightings information

All the trips started from Ponta Delgada harbour and lasted on average 2.30 h to 3 h. The area surveyed covered mostly the entire south coast.

Whale watching companies from Azores rely on the so called "Vigias" (in Portuguese): these are experienced people equipped with military binoculars who, from early morning, observe the ocean from specifics spots on land. Vigias can see traces of animals till 45 km from land. Depending on the cetacean, they can directly recognise the species by their blow (the characteristic 45° tilted to the left of the sperm whales), by the visible colour or shapes, or by their behaviour on the surface. Most of the observations made during the tours are guided from land. Once the vigias spot something, they communicate with the boats and guide them to the location of the animals using the radio. To maximise chances to find sperm whales and to complement the land search for the animals (e.g., low visibility, great distance from the coast) and the boat search done onboard by the crew, a custom-built directional hydrophone is often used. This equipment is used from the RIBS, allowing the biologist and skipper on board to find the direction from which the sound of the eventually present animals comes, and therefore to increase chances to find it on the surface on that direction.

Cetacean observations, and therefore, correspondent data collection, were done always complying with the rules set by the Government of Azores. Approaching (Figure 2.2) should be done from behind, in order to leave a 180º sector free ahead of the animal. The maximum distance allowed is 50 m and the sighting can last for maximum 30 min (DLR 10/2003/A).

Figure 2.2: Approaching scheme of an animal during whale watching trips, as set according to the Azorean Government DLR 2003/A. Source: Whale and Dolphin Watching Guidelines | Futurismo (whalewatchingazores.com) (2018)

Data collection included GPS coordinates, date and time, wind and sea state (Beaufort and Douglas scale) and age structure of the group sighted, among other information.

2.2.2 Protocol for acoustic data

All acoustic recordings were obtained from three of the inflatable boats (or RIBS) of the company. The dataset of recordings is divided in two subsets: subset 1 composed by recordings taken by different biologists of Futurismo, from June 2019 to October 2021; subset 2 composed by recordings mainly

made by the PhD student Fadia Al Abbar, from May 2019 till August 2019, during whale watching trips with Futurismo. Protocols followed to record the data on each subset were slightly different and are explained in detail in the following lines.

For this study, both echolocation clicks and codas were collected, as both type of sound have been found reliable for IPI analysis (Pavan et al., 1997; Pavan et al., 2000; Madsen et al., 2002a). Codas were recorded only when at least two individuals were spotted together. Echolocation clicks were recorded right after the fluke-up (with 1 min of tolerance), with the hydrophone introduced in the water behind the animal (i.e., usually in the fluke print left by the whale when submerging). Following the whale watching rules and in order to take photo-identification photos, usually the boats are already behind the animals. Sperm whales in the Azores start to click on average in the first 2-3 min after showing the tail (Goold and Jones, 1995). The time delay between the first clicks emitted by the animal and the starting of the recording can be reduced by recording directly from the point of observation of the animal. Moreover, as the IPI is highly directional and aspect dependant, being behind the animal will give higher probability to record click on-axis and backwards, with the P0 pulse the most intense in respect to the subsequent pulses (Rhinelander and Dawson, 2004; Teloni et al., 2007).

Protocol subset 1

The equipment used for the subset 1 was composed by a set of a handy recorder, Zoom H4n PRO (bit depth: 16/24-bit; sampling frequency: 44.1/48/98 kHz; CH4, MTR, STEREO mode) financed by Idea Wild, and a single hydrophone, Aquarian H2a (-180dB re: 1V/µPa; ± 4dB 20Hz-4KHz; frequency range <10 Hz to >100KHz; sensitivity at 100 kHz: -220dB re: 1V/µPa) with 15 m cable, owned by Futurismo.

A temporary distinction has to be made for the protocol: (1) Before December 2020 (2) After December 2020. From June 2019 till December 2020 (period 1) different settings have been tested. For the first audios, MTR and primarily CH4 mode were used. MP3 format type was set for 2019 recordings. All the audio files were recorded at 44.1 kHz of sampling frequency and 16-bit depth. Recordings were obtained mainly from diving individuals, starting at different times after the fluke up. The "fluke-print" spot (the mark left by the fluke of the animal after submerging) was mainly used as the point to drop the hydrophone to record. No duration limits were set.

From December 2020 (period 2) the final protocol was adopted, with the main purpose of using the recordings for acoustic research on sperm whales. STEREO mode and WAV format were used. A sampling frequency of 96 kHz and bit depth of 24 bit were set, as they may provide better results for sperm whale acoustic analyses (Thode et al., 2010; Miller et al., 2013; Caruso et al., 2015). A higher sampling frequency was required to have better accuracy on the clicks. According to the sampling theorem, for digital audio the sampling frequency must be always at least two times higher than the maximum frequency to be recorded (Lai, 2003). Moreover, digital recorders may introduce aliasing artifacts at the upper limit of their frequency range, e.g., close at the Nyquist frequency that is the half of the sampling rate. A greater bit depth was necessary to have a better resolution (in terms of amplitude values) of the different sounds recorded and, important for this study, a better image of the pulsed structure of the clicks.

The hydrophone was dropped in the water as soon as the animal prepared to dive (typical behaviour can be observed and recognised). The hydrophone was used with the full-length cable, aiming to have it at a depth of 15 m (or as close as possible). However, sea state or wind can affect the straightness of the cable and result in a shallower position. For this reason, from September 2020 two weights of 200 g each were positioned in the cable, at 50 cm from the hydrophone. Moreover, the person recording held the cable more distant from the boat to avoid direct contact (and noise) between the cable and the rib.

The duration of the recording was set to be at least 2 min and a maximum of 10 min, considering what Rhinelander and Dawson (2004) established, based on the knowledge acquired by Gool, 1996 regarding the change on IPI values with higher pressure levels.

Protocol subset 2

The second subset is composed by recordings made between May 2019-August 2019. Data collection was done for the development of a pilot study on the acoustics of sperm whales related to the number of boats in the area. Two different hydrophones were tested, one Aquarian H2a with 25 m of cable (from Wageningen University) and one Magrec HP30 (old hydrophone. Frequency range: 200 Hz-20 kHz) from Futurismo. The digital recorder was a TASCAM HD-P2 (sampling rates: 44.1/48/88.2/96/176.4/192 kHz; bit depth:16/24-bit; STEREO mode). A simple voice recorder was used for Magrec HP30 hydrophone.

The protocol applied was slightly different from the one of the subset 1. Hydrophone was dropped in the water at a minimum depth of 1 m when the whale prepared to dive. Recording started as soon as the whale showed the tail. Then, a photo of the fluke, GPS coordinates and other related information were required. No duration limit was set. However, the first clicks were required and preference to record the entire dive was suggested. Recordings were made at 44.1 kHz and 16- or 24-bit depths.

2.3   Data classification

Recordings were classified in 6 categories (0 to 5) according to: (1) the presence of clicks; and (2) the number of pulses detected (Table 2.1). This classification provides information on the overall quality of the recordings.

Table 2.1: Classification categories of the recordings

0	REALLY BAD, NO CLICKS DETECTED
1	BAD, NO CLEAR CLICKS / CLICKS COVERED BY NOISE
2	UNUSABLE, NO CLEAR STRUCTURE OF CLICKS
3	USABLE, SOME CLICKS CAN BE ANALYSED
4	GOOD, CLICKS' STRUCTURE IT'S ENOUGH CLEAR, SOME CLIKS HAS MORE THAN ONE PULSE
5	REALLY GOOD, CLICKS PERFECTLY DISTINGUISHED, MORE THAN ONE PULSE

Raven Lite 2 was used to visualize the waveform and spectrogram of all the recordings. Visualization was improved in order to better detect the pulses present in the recordings. Contrast and brightness were changed to reduce the impact of the background noise on the spectrogram (Figure 2.3).

A "bandpass" filter between 1 kHz to 2 kHz was applied to filter out all the energy in a particular frequency band. Background noise covers usually the low-band frequency with biological and anthropogenic sounds. Abiotic noise (breaking waves, wind, bubbles) usually covers a mid-frequency range (Hildebrand, 2009; Peng et al., 2015). This range is usually overlapped to the frequency range of sperm whale's usual clicks, that has most of the energy between 2 kHz - 40 kHz (Thode et al., 2002). As a result, this bandpass filter lowers the wave amplitude. For better visualization, the recordings were amplified, both by the "Multiply by factor" (from 1 to 4) and the "Add dB" (+5 or + 10 dB) functions. This second step was fundamental when clicks had already a low amplitude before filtering.

The "spectrogram window size" modifies the resolution of the spectrogram, allowing the view of the multipulsed structure of a click at values between 37 to 42 units (n° of samples). The pre-set value is 512 (Figure 2.3). Decreasing the window size lowers the frequency precision and increase the time resolution, the parameter needed as pulses have a small duration (15-30 ms for usual clicks and 35 ms for coda) (Whitehead, 2003). Precision on frequency is not relevant for this study.
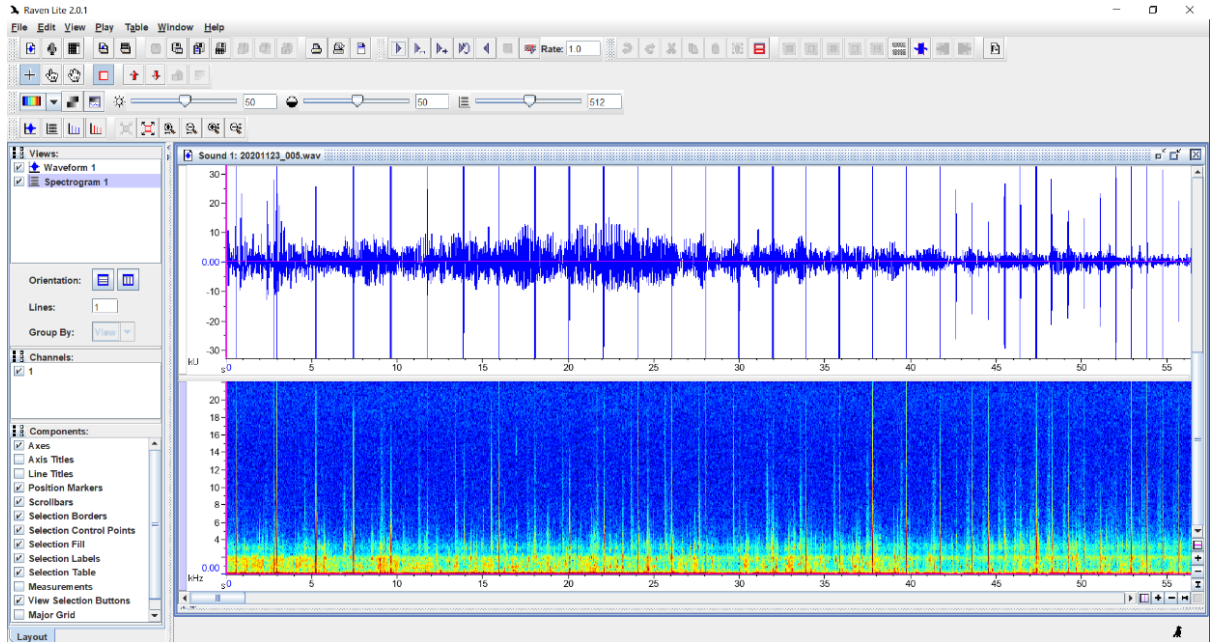

Figure 2.3: Screenshot of Raven Lite 2, showing the Spectogram and Waveform of a typical recording.

Different classification rates mean different reliability. Recordings classified as 3 has only two clear pulses, while recording classified as 5 has more than one pulse for most clicks. This means that the signal-to-noise ratio in "type 3" recordings is lower than in "type 5" recordings, as the higher decay rate for the subsequent pulses reduce their amplitude until they result covered by noise.

The pulsed structure of two different recordings, with different classification ranking, can be quite different (Figure 2.4a, 2.4b). In a), a recording classified as "3" and in b) a recording classified as "5". Spectrogram and waveform of the audio file in a) has been adapted to the same scales of the recording in b) to have a better comparison. Bandpass filter with lower limit at 2 kHz was applied. No amplification was applied. In the case of recording in a) an amplification is suggested for a better visualization of the amplitude. In both of the recording is evident the surface reflection of the single clicks, that appears after the original click. Surface reflection is evident from the phase reversal (i.e., same waveform but reversed) that is due to the nature of the reflection itself. Time delay between original click and surface reflection is variable. In these two recordings is evident that there is no overlap between the click and its reflection.
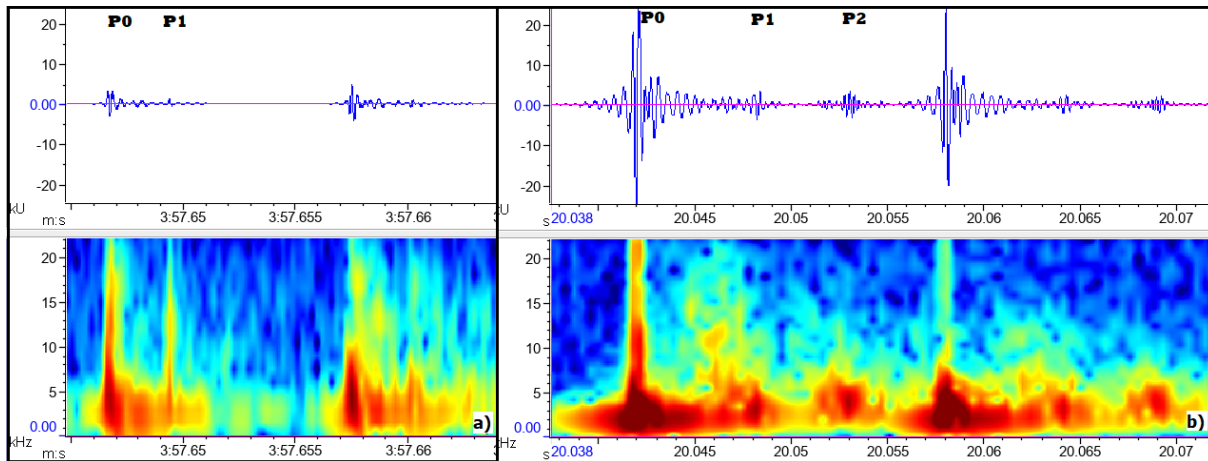
19

Figure 2.4: Waveform and spectrogram showing the pulsed structure of a) a "3" recording and b) a "5" recording

2.5 Acoustic analyses

The analysis of the recordings, including the IPI extraction, was done in three different ways:

1) A manual method with the software Raven Lite 2 (Cornell Lab of Ornithology)
2) Automatic method with PAMGuard IPI Plugin (Miller et al., 2013)
3) software CABLE version 0.3.

All files were used for analyses with method 2 and 3, but only the audio files with a grade ≥ 3 were used for the manual analysis (method 1). Recordings with less than 30 s (classified as 0) were excluded because too short to have enough clicks. All the methods used assumed that, for a single recording, only whales with a slightly different stable IPI (± 0.3 ms) were discernible for the analyses. Thus, individuals with a similar IPI were not distinguished and their IPI estimations were merged together in one unique value (Caruso et al., 2015; Beslin et al., 2018).

2.5.1 Raven analysis

Raven Lite requires the WAV format to work. Thus, audio files with a different extension were converted to WAV with the software Audacity® (v. 3.0.0).

In order to be coherent with the findings of Rhinelander and Dawson (2004) for IPI stability, only the first 8 min for all recordings were used. After opening the file, filtering and amplification were applied as discussed in paragraph 2.3.

Only clear clicks (classified as ≥ 3) with at least two different pulses, were considered for the IPI estimation. In addition, only clicks that had an evident and clearly separated surface reflection (i.e., not overlapped to the original click) were analysed. Due to the protocol adopted, only a "backward on-axis click" structure should have been present for the majority of the clicks, with the P0 pulse the most intense (Teloni et al., 2007). Thus, the IPI was considered between the P0 and P1 pulses. However, due to the uncertain nature of the direction of the animal recorded to the hydrophone, all the clicks usable were analysed. Averaging a large number of clicks should lower the aspect-dependent nature of clicks and estimate the "true IPI" (Zimmer et al., 2005; Teloni et al., 2007).

The "selection" function was used to measure the distance from two subsequent pulses. When more than two pulses were available, the distance between all the pulses was determined. For a better accuracy, selections were made from the spectrogram but compared also with the waveform. Selections in Raven

provide the exact time (s) corresponding to the beginning and the ending of the selected area. Times are given with four decimal digits.

To estimate the IPI, the beginning and ending time of all the selection were exported to an Excel file. In order to avoid errors with doubled data, spectrogram or waveform visualization was disabled before coping. The dataset was prepared with five columns: CODE (for individual recording), CLICK ID (a progressive number only for order purpose), START (s) (begin time), END (s) (end time) and IPI (ms). For the calculation of IPI, the difference between end and start time was multiplied by 1000 (for conversion from s to ms). In case of multiple whales present, click trains were distinguished, in order to assign the IPI found to a specific cluster. IPI clusters were considered different if a difference of $\pm$ 0.3 ms was found.

For each individual recording, there were calculated mean (the mean of IPI in ms), mode (the mode of IPI in ms), CV (coefficient of variation), SE (standard deviation from the mean in ms), min, max (minimum and maximum value of IPI estimated for a recording) and n° click (the number of click used to calculate the IPI). Length estimations (Total Length or TL) were then reported. TL was obtained with both formulas of Gordon (1991) (2.1) and Growcott et al. (2011) (2.2), using the mean IPI resulting from previous analysis.

$$TL = 4.833 + 1.453 \cdot IPI - 0.001 \cdot IPI^2 \qquad (2.1)$$

$$TL = 1.258 \cdot IPI + 5.736 \qquad (2.2)$$

2.5.2 PAMGuard analysis

PAMguard is an opensource software developed for passive acoustic monitoring of marine mammals. It is freely available, and many plugins were created by contributors (Gillespie et al., 2008). One of the plugin modules is the *Sperm Whale IPI* plugin (Miller et al., 2013), firstly presented by Miller (2010).

PAMguard requires the WAV format and works with 16-bit audio. Thus, recording with a different extension were converted to WAV and all the recordings at 24-bit modified to 16-bit with the software Audacity® (v. 3.0.0).

The *Sperm Whale IPI* plugin uses the cepstrum analysis as done by Teloni et al. (2007) to estimate the IPI. Cepstral analysis is a non-linear signal analysis (i.e., non-Gaussian model) described firstly by Bogert et al. (1963). The power cepstrum was defined as "the power spectrum of the logarithm of the power spectrum of a signal" (Antunes, 2009) and it was introduced for detecting delayed versions of a signal (echo arrivals times). In this definition, the cepstrum is a powerful tool for IPI estimation, as the subsequent pulses of a click are the echos (with a delayed time of arrival) of the first pulse (Antunes, 2009).

Miller (2010) computed the cepstrum for individual clicks as

$$C_t = |FFT^{-1}(\log|FFT(x_t)|)| \qquad (2.3)$$

x_t is the digital representation of the time domain waveform; FFT and FFT^{-1} are the fast Fourier transform and its inverse, as described by Teloni et al. (2007).

Time delay at peak cepstrum is saved as the IPI of the analysed click. A frequency histogram for all the IPI found is then computed. Mean and mode IPI of all clicks used are saved. Together with the histogram, the plugin was built to compute also the ensemble average of the cepstrum from all the clicks

analysed. As described by Teloni et a. (2007), the time delay of the peak value of the ensemble is stored as the ensemble IPI. Averaging a large number of clicks allows the analysis of recordings made at an unknow aspect to the whale recorded. The peak width of the ensemble (set at 75% of the maximum value, empirically found to be consistent) was used by Miller (2010) as a discriminatory factor for the analysis. A maximum peak width of 1 ms was set, in order to avoid estimations with high uncertainty. The number of clicks used for IPI estimation are also stored.

Parameters were set following Miller (2010). The software automatically detects the sampling rate and the number of channels of each recording. A preamplification of 10 dB was set to homogenise the amplification process for all the recordings. Before proceeding with the analysis, a high-pass filter (to filter out the low-frequency noise) with an eighth order Butterworth filter was applied (Digital pre-filter in the click detector core module). Corner frequency was set to 2 kHz to homogenise the filtering process for all acoustic files. Filtered data were then used as input into PAMguard click detector. This module compute automatically all the clicks that can detect in an audio file. More features, such click train identification and classification, can be used.

Settings for the "Sperm Whale IPI parameters" were set. Source for the IPI computation are the clicks coming from the click detector. The cepstrum duration was set to 40 ms and 2-9 ms was the interval in which to calculate the IPI. These limits were set according to Marcoux et al. (2006) and Beslin et al. (2018). The upper limit corresponds to a sperm whale body length of 17.8 m. The lower limit corresponds to a body length of 7.7 m (4 years old for male and 6 years old for females according to REF). This was set to avoid estimations of unreal IPI due to first pulses with an extended duration (> 2 ms). Limiting to a minimum of 2 ms may exclude young calves from further analysis (Marcoux et al., 2006; Beslin et al., 2018). The peak width was left as at 75%. An automatic filter for clipped clicks is run by the plugin, due to the distortion that the cepstrum of a clipped waveform can show (Miller, 2010).

Important settings of the click detector plugins are found in the "Click Detection Parameters" section, that regulate the number of clicks that will be recognised from the software. Click detector works comparing the acoustic energy in the recording and a set of control frequency bands. The energy in the recording band must exceed that in the control band by a threshold (in dB) to produce a detection (Jacobson et al., 2013). A minimum threshold of 11.0 dB was set to maintain the signal-to-noise ratio (SNR) > 10 dB (Rhinelander and Dawson, 2004). Increasing this value will exclude multiple whales from being detected and only the loudest individuals will be analysed. "Click length" was changed according to the IPI plugin manual. Default parameter "Max click Length" (called nSamples in the manual, the duration of the click) is not suitable for the sperm whale IPI plugin, so it has to be changed accordingly to the developer's hint. The parameter was calculated with the formula:

$$nSamples = \frac{Cepstrum\ Duration}{1000} * sample\ rate \qquad (2.2)$$

Sample rate is the sampling frequency of the recording, measured in Hz. The post sample parameter must be between 90-100% of the nSamples. The remaining 10-0% is the pre sample parameter. These settings are important so that every click detection will have at least nSamples.

An *Echo Detector* and *Noise Sampling* feature are set as default to clean the recording.

Recordings were used at their entire duration (with a maximum of 15 min). The ensemble technique of the IPI plugin allows to analyse recordings after the maximum time of 8 min set by Gordon (1991) and Rhinelander and Dawson (2004).

In this study, multiple whales IPI estimation was tested. Different values of threshold were used per recording. The right value was decided when one clear peak showed by the ensemble was present, and when the peak width was < 1 ms. In the case of multiple whales present, with more than one clear

ensemble peak, the pre-set IPI interval was changed. In order to isolate one peak at a time, different values for the interval were applied. Some recording can show two distinguished peaks both from the ensemble and the histogram (Figure 2.5). The peak at the left correspond to louder clicks. Without changing the IPI interval (in this case, 2-4 ms for the first and 4-9 ms for the second) the software would recognise only the first peak. Only the clearest peaks, with comparable height and histogram peaks present, were chosen. If the isolation process revealed a peak width > 1 ms, the IPI cluster was discarded.
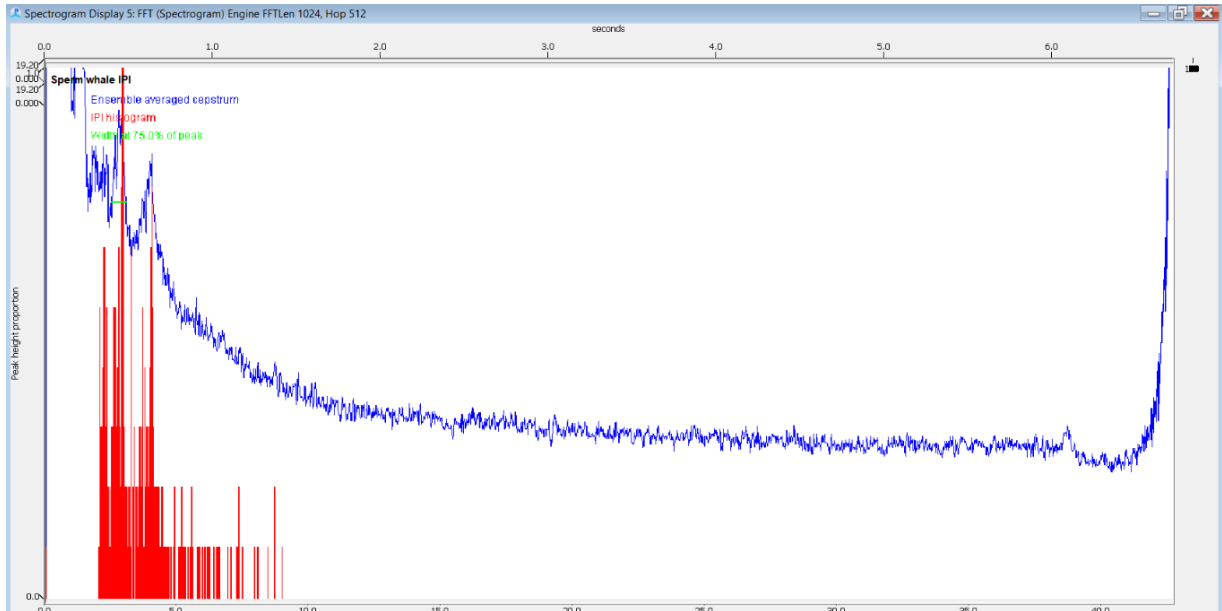
Figure 2.5: Screenshot of the PAMguard results from IPI plugin. In blue, the ensemble average cepstrum; in red, the histogram. A green line shows the peak width at 75% of peak.

Data from the IPI module are stored in the PAMguard database. In order to visualize the database, SQLiteStudio (v 3.3.3) library was used. Ensemble IPI (from the peak value), Ensemble IPI lower limit and upper limit (defined by the peak width), mode IPI and mean IPI (from the histogram) and number of clicks used for the estimation were noted. All the data were exported to an Excel file in the same order. Difference between upper and lower limit of the ensemble was used to calculate the peak width.

Length estimation were than reported in both the two formulas of Gordon (1991) (2.1) and Growcott et al. (2011) (2.2).

2.5.3   CABLE testing

CABLE (Cachalot Automatic Body Length Estimator) is a freely available software developed by Beslin et al. (2018) at the Hal Whitehead's Research Group. It is a fully automated software developed for IPI size estimation of sperm whales from only on-axis clicks. This software is presented as a powerful tool for automatic IPI estimation, but it was developed in the area of Dominica. Classification process (highly selective) could not be adapted for other geographical areas and populations of sperm whales. In order to see the performance with the dataset used in this study, all the recordings were analysed.

WAV extension is required for the software and analyses are conducted only in one specified channel. Original files were modified with Audacity accordingly. *CABLE* works with 48 kHz sampling rate, but it resamples automatically for different sampling frequency. The program is based on Matlab version R2015a which is required prior to installation (The MathWorks, Inc., Natick, Massachusetts). Default parameters have been established by the developers as highly reliable and they suggest to not modify them, in order to obtain reliable results (Beslin et al., 2018).

Different steps are executed by the software to complete a run. Before starting, it applies automatically a 2–12 kHz Butterworth bandpass filter for noise exclusion. After that, a custom-built click detector is run (similar to the one of PAMguard). Important for detection is the SNR of the recordings. The software automatic click detector has difficulties in detecting clicks with low SNR. Each click is analysed singularly to distinguish the on-axis and off-axis clicks and discard the latter. Then, all the clicks are divided in "Good" and "Bad". Coda clicks are always considered "Bad" because of their different multi-pulsed structure. After that, detection of different pulses is applied. Finally, the IPI estimation is done by the two methods proposed by Goold (1996): autocorrelation and cepstral analysis. IPI values are limited to the interval of 2-9 ms. The cepstrum is computed as:

$$C_q = |FFT(\log|FFT(x_t)^2|)| \qquad (2.3)$$

The cross-correlation (or autocorrelation) analysis is based on the repetitive correlation of the click waveform with a replica progressively time delayed. Peaks in the function occurs when the shift in time between the replica and the original is equal to an integer multiple of the IPI (Goold, 1996). Values from both the autocorrelation and cepstrum are then averaged to have a unique value of the IPI. All the clicks with the two IPI estimations that differs more than 0.05 ms are discarded.

For the final validation of the IPI, the software search for IPI repetitions. The same IPI, for the nature of clicks of sperm whales, is expected to be present more than one time in few seconds of a click train. The default value for "IPI repetition" parameter is 1, so that at least one similar IPI has to be present for this validation step. The IPI repetition parameter can be an integer value, with a minimum value of 0. Beslin et al. (2018) tested it till a value of 4. The overall goodness set the classification quality limit for "Bad" and "Good" clicks. It can be 0.1, 0.3, 0.5, 0.7, 0.9, but the default value is 0.7. Normal ICI intervals of sperm whales clicks (default, 0.25 – 1.5 s) are used to scan for different IPIs.

A cluster analysis is finally adopted, with Gaussian Mixture Models (GMMs) to fit the frequency distribution of the IPIs found. Each peak found in the distribution is related to a different individual, assuming that only whales with a different IPI are present. The mean of each cluster in a GMM are estimates of each whale's true IPI. The result is a frequency histogram with a probability density curve from the GMM. Different clusters in different colours and a summary value of the IPI (with its standard error) of the different whales is present too (Beslin et al., 2018).

Different responses of the software were tested with the dataset of the study. The default values were used for the first analysis. After that, all the levels of goodness (0.1, 0.3, 0.5, 0.7, 0.9) were tested together with IPI repetition of 0 and 1.

2.6 Statistical analysis

All the subsequent statistical analysis were carried on with RStudio software. Different additional packages were used: car, dplyr, ggplot2, ggpubr, grid, multimode and tidyverse (v 3.6.1. RStudio team).

The quality classification of the recordings was tested to see the eventual effect of different factors: wind speed (Beaufort scale), wave height (Douglas scale), frequency sampling (as a factor of digital recording method) and season were considered for the subset 1. Only Beaufort and Douglas scales were considered for the subset 2, as frequency sampling and bit depth did not change among recordings.

Since the dependent variable (classification level) is a categorical and ordinal variable, and independent variables are also categorical, an Ordinal Logistic Regression model was adopted. The dataset considered and the purpose of the analysis was found consistent with other studies that used the OLR quality (Abreu, 2008; Harrell, 2015). Thus, the sample was too small for some of the categories to conduct the statistical analysis and have reliable results.

2.6.1 Comparison of IPI from Ensemble and Histogram estimation in PAMguard

In order to obtain a reliable mean value of IPI, the difference between the measurements obtained with the Ensemble and Histogram methods were statistically compared.

The distribution model of the data was analysed separately. A first visualization of the dataset was done with a frequency histogram. Results from both histogram and Kolmogorov-Smirnov normality test (p-value < 0.01) showed a non-normal distribution. Subsequently, a unimodality test (Ameijeiras-Alonso et al., 2019) was run (p-value < 0.01). Since the distribution of both sets of data showed a bimodal distribution, a Paired Samples Wilcoxon Test was chosen. This test is a non-parametric alternative to the paired t-test that requires a normal distribution. The choice of a paired test was done because of the nature of dependency between the measurement from the Ensemble and the Histogram. In other studies, the statistical differences between the two methods for IPI estimate and between the formulas for the total length were tested with a t-test (Growcott et al., 2011; Miller et al., 2013; Caruso et al., 2015).

Wilcoxon test requires a symmetrical distribution of the differences between the two sets of data. For this reason, the differences were calculated and then plotted with both a scatterplot and a histogram.

A boxplot showing the variability in the two different datasets was also plotted. The resulting p-value from the Wilcoxon test was added as text.

2.6.2 Comparison of total body lengths obtained from Gordon (1991) and Growcott et al. (2011) formulas

Total length calculated with the formulas of Gordon (1991) (2.1) and Growcott et al. (2011) (2.2) were compared. A histogram of both formulas was plotted. Normality test and unimodal test were run, showing the same result (p-value <0.05). Symmetry of the differences' distribution was check before proceeding with the Wilcoxon Test.

A boxplot showing the variability in the two different datasets was also plotted. The resulting p-value from the Wilcoxon test was added as text.

2.6.3 Comparison between results from Raven and PAMguard

The comparison between IPI estimations obtained from Raven and PAMguard were equal to the precedent comparisons. A Shapiro-Wilk test for normality was applied as the number of samples was low (n <30).

In order to have a matrix-like dataset, only similar IPI measurements were compared. In case of multiple IPI estimations on a recording, only the ones of similar order (i.e, ± 1 ms) were retained. The others were discarded.

A boxplot showing the variability in the two different datasets was also plotted. The resulting p-value from the Wilcoxon test was added as text.

2.6.4 Population size structure

A first attempt of a population size structure distribution was done. A histogram showing body lengths with interval classes of 40 cm (value chosen for better visualization only) was plotted for the average total length estimates.

Following Caruso et al. (2015), the mean total length estimated from PAMGuard were divided in 3 age classes. The subsequent classes were obtain from Caruso et al. (2015), that were base on growth curves from Rice (1989).

- Immature Male or Female: $TL < 9$ m;

- Adult female or Juvenile Male: 9 m $< TL < 12$ m;

- Adult Male: $TL > 12$ m

Resulting counts per classes were plotted with a barplot.

To strengthen data from the population size structure obtained, a seasonality was also applied. Counts of estimations obtained from PAMGuard were organized in the four seasons (Winter, Spring, Summer and Autumn). A histogram was then plotted, with counts grouped by season and by age class.

2.6.5   Different size estimation of the same individuals

Presence of multiple measurements for the same individuals was inquired. No photo-identification was related to the recordings used for this study, but eventually identified animals by the expert biologists onboard were taken in consideration. A comparison between the measurement obtained from PAMGuard and Raven was eventually done for identified animals.

2.6.6   Comparison in the subset 1: first period vs. second period

Different quality classification scores and numbers of recordings usable on the total of recordings available were compared for the two periods of the subset 1. Data from both Raven and PAMGuard were used to assess an eventual statistical difference between the two periods.

3 Results

A total of 119 recordings (subset 1 = 61 and subset 2 = 58) were collected during the sampling season. The duration was different for every file, ranging from 0.04 min to 24.32 min. Total number of days of data collection was of 27 days for subset 1 and 23 days for subset 2.

3.1 Acoustic analyses

Most of the recordings were classified as 2 in subset 1 (55.7%); and as 1 in the subset 2 (75.9%) (Table 3.1). Only 11 recordings have been used for manual analysis as where ranked ≥ 3. 7 recordings were excluded from further analysis as the total duration was < 30 s. For subset 1, without the 3 recordings classified as 0, the rate of acceptance was of 19%. For subset 2, no recordings were suitable for further analysis

Table 3.1: Summary of the classification of the recordings per level and subset

Classification	Subset 1	Subset 2
0	3	5
1	13	44
2	34	9
3	6	0
4	4	0
5	1	0

3.1.1 Raven analysis

Manual analysis with Raven resulted in 14 IPI measurements from 11 recordings. Recording 20201123_005 and 20201208_006 where the only ones in which two different IPI were measurable. However, recording 20201123_005 was the only one with more than two subsequent clear pulses. Thus, one IPI estimates correspond to the P0-P1 difference in time arrival, while the second IPI estimate for this audio correspond to the IPI measured between P1 and P2 pulses. A summary table of the recordings and the statistical descriptors calculated is displayed in Table 4.

The number of click averaged for the resulting mean value ranged from 38 to 309. No clear pattern was observed between the duration of the recording and the number of clicks usable for the averaged IPI. Only one of the recordings used was made of coda clicks instead of usual clicks. Mean IPI ranged from 2.19 ms to 6.28 ms. No values smaller than 2 ms or greater than 9 ms were found. The mean difference between the minimum and the maximum value of a single IPI estimate was 0.39 ms. Different measurements of the same IPI with a difference of also 0.60 ms were found. CV ranged from 1.78% to 5.34% (mean 3.37%), while SE associated with the mean ranged from 0.7% to 1.8% (± 0.01 ms to ± 0.02 ms). Total length varied from 8.02 m to 13.91 m with the equation (2.1) and from 8.50 m to 13.63 m with the equation (2.2). Differences between the two formulas were consistent, with a maximum difference of 0.48 m using the mean IPI estimate. Considering the minimum and the maximum value of the IPI for TL estimates, the difference between the two formulas was greater. 0.87 m was the maximum difference between TL using (2.1) and (2.2) estimates for minimum values of IPI. 0.75 m was the maximum difference using the maximum value of the IPI.

Table 3.2: Summary of the recordings and the statistical descriptors calculated. CODE (univocal numeration for recordings, with date (AAAAMMGG) and progressive number by year) ; MEAN (the mean of IPI per recording in ms); MODE (the mode of IPI per recording in ms); CV (coefficient of variation per recording), SE (standard deviation from the mean in ms. Standard deviation / square root of total number of samples), MIN, MAX (minimum and maximum value of IPI estimated for a recording) and N° CLICK (the number of click used to calculate the IPI). TL (m) – Go and TL (m) – Gr: TL calculated with (2.1) and (2.2) respectively. Data ordered by date.

CODE	MEAN	MODE	CV	SE %	MIN	MAX	N° CLICK	TL (m) - Go	TL (m) - Gr
20200830_002	3.16	3.00	4.35%	1.1%	3.00	3.40	160	9.41	9.71
20201123_005	6.28	6.40	1.78%	0.7%	6.00	6.40	250	13.91	13.63
20201123_005	4.78	4.90	3.03%	1.8%	4.60	5.10	66	11.76	11.75
20201208_006	2.82	2.90	5.34%	0.7%	2.50	3.10	301	8.92	9.28
20201208_006	2.19	2.20	4.61%	1.1%	2.00	2.40	104	8.02	8.50
20201208_007	2.75	2.70	4.31%	0.7%	2.50	3.10	309	8.82	9.19
20210527_018	3.07	3.00	4.39%	1.2%	2.80	3.40	129	9.29	9.60
20210603_024	2.67	2.70	2.76%	1.1%	2.60	2.80	49	8.71	9.10
20210708_028	3.54	3.60	1.92%	0.7%	3.40	3.60	94	9.97	10.19
20210915_035	3.08	3.10	2.33%	1.2%	3.00	3.20	38	9.30	9.62
20210915_036	3.16	3.20	2.68%	1.1%	3.00	3.30	56	9.41	9.71
20210928_038	2.78	2.80	2.72%	0.9%	2.60	2.90	72	8.87	9.24
20211003_039	2.87	2.90	3.93%	1.1%	2.60	3.00	104	9.00	9.35
20211003_039	3.20	3.10	3.05%	1.2%	3.10	3.40	64	9.47	9.76

3.1.2   PAMGuard analysis

From a total of 111 recordings analysed with PAMGuard, 63 (56.8 % of the recordings) showed at least one clear peak with width < 1 ms. 47 recordings were from the subset 1 (77% of the subset 1), while 16 were from the subset 2 (28% of the subset 2). In total, 87 different IPI estimations were obtained. From 17 recordings it was possible to calculate 2 different IPI and from 4 recordings 3 measures were obtained.

From the ensemble and the histogram mode of 39 estimations the same IPI value was obtained. For the remaining, the differences among the two methods varied between 0.02 ms and 0.35 ms. The mean difference was of 0.05 ms. These differences in the measurements reflect a difference on the total length estimation. Using Gordon's (1991) formula, the TL between the two methods varied from 0.0 to 0.51 m, with a mean of 0.07 m. With Growcott's et al. (2011) formula, the TL between the two methods varied from 0.0 to 0.45 m, with a mean of 0.06 m. The IPI averaged from ensemble and histogram values was used for total length estimates with the two equations. The mean IPI ranged between 2.04 ms to 7.02 ms. Total length ranged between 7.79 m to 15.03 m for (2.1) and 8.30 m to 14.57 m for (2.2). Differences between the two formulas were larger, as found in Raven analysis. The range was 0.56-1.73 m with a mean of 1.03 m.

3.1.3   CABLE testing

The recordings were firstly analysed with an *IPI repetition* of 1 and 0, with the *overall goodness* fixed as 0.7 (default). From the former, only 2 recordings gave a result with one whale recognised per recording. From the latter analysis, no changes were seen and the same two recordings gave the same results (Figure 3.1a, 3.1b). Changing the *overall goodness* did not produce changes for values of 0.9

and 0.5. For *overall goodness* 0.3, one recording produced two different histograms-GMM and so possible individuals. Again, no other changes were seen.
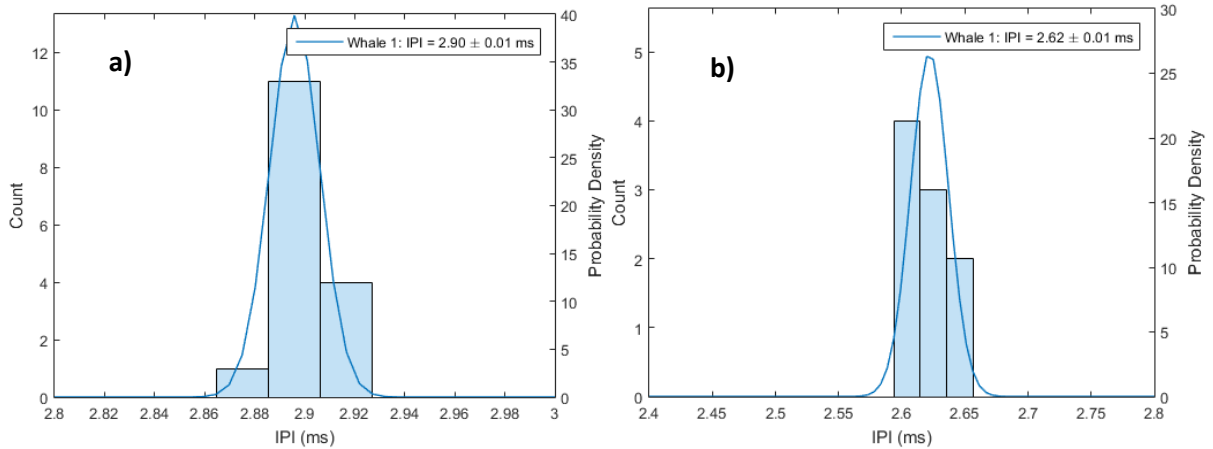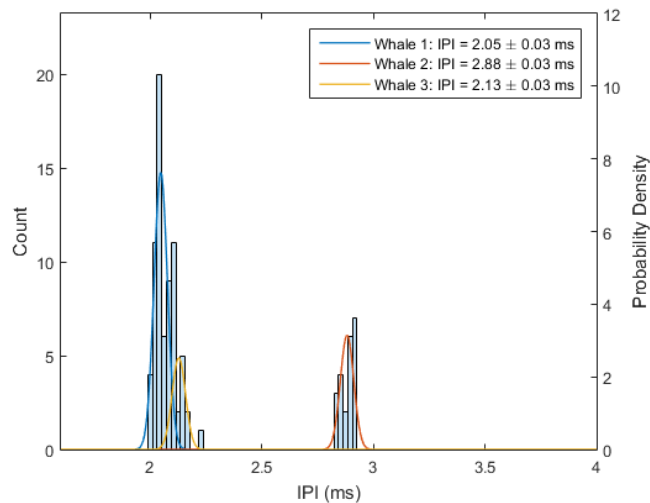

Figure 3.1: Histograms of the IPI estimate from two recordings with CABLE. The curve represents the GMM for the best fit based on the probability distribution. a) IPI repetition = 1. Overall goodness = 0.7. b) IPI repetition = 0. Overall goodness = 0.7.

For *overall goodness* = 0.1, 81.7% of the recordings produced results. No differences were found by changing *IPI repetitions* to 0 or 1. All the recordings showed more than one possible GMM fit (Figure 3.2). Decreasing the overall goodness, the number of clicks considered reliable from the software increased. It can be easily seen by comparing Figure 3.1a, 3.1b and Figure 3.2.

Figure 3.2: Histogram of the IPI estimate from one recording with CABLE. The curves represent the GMM for the best fit based on the probability distribution. IPI repetition = 1. Overall goodness = 0.1.

3.2 Statistical analysis

3.2.1 Comparison of IPI from Ensemble and Histogram estimation in PAMguard

The distribution of both the Ensemble and Histogram IPI measurements were bimodal (Kolmogorov-Smirnov normality test p-value < 0.01). The Wilcoxon t-test assume that there are no significant differences between two dependent samples. Thus, with a p-value > 0.05, the null hyphotesis was retained (Figure 3.3). The two estimations were averaged to obtain a unique value (mean IPI) to use in the formulas to estimate the total length (Growcott et al., 2011; Beslin et al., 2018).
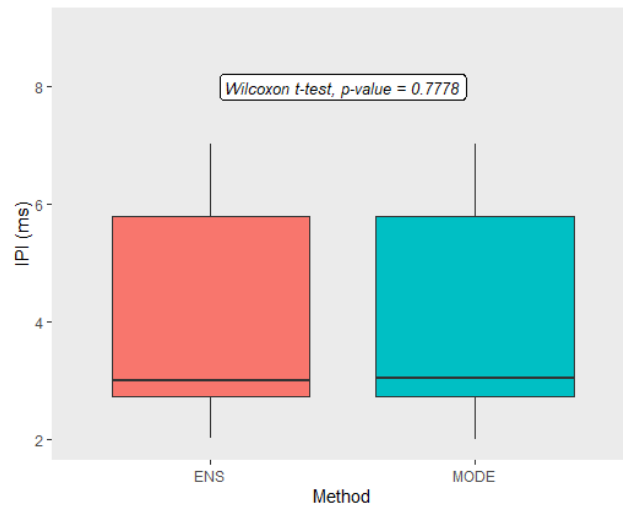
Figure 3.3: Boxplot of the two sets of IPIs measurements (ms) obtained from PAMGuard. ENS and MODE respectively = Ensemble estimates of IPIs and Histogram mode estimations of IPIs.

3.2.2 Comparison of total body lengths obtained from Gordon (1991) and Growcott et al. (2011) formulas

Total body length obtained from Gordon's formula (1991) and from Growcott's et al. formula (2011), showed for both cases a bimodal distribution. Classes of 40 cm were chosen for a better visualization of the data (Figure 3.4a, 3.4b). Significant differences between IPI estimation obtained from the two equations were found (Wilcoxon t-test, $p < 0.05$) (Figure 3.5).
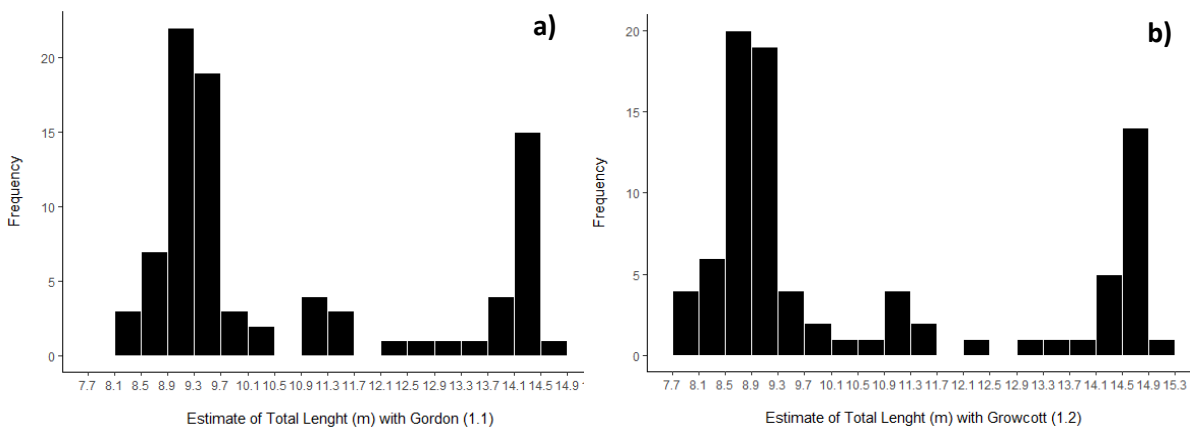
Figure 3.4: Histogram of the total length estimates in metres with Gordon's formula a) and Growcott's formula b)

Gordon's equation tends to be more conservative for TL < 11 m, while tend to estimate greater sizes for TL > 11 m (Table 3.2, Figure 3.4a). On the opposite, Growcott's equation, tend to estimate greater sizes for TL < 11 m and smaller size for TL > 11 m (Table 3.2, Figure 3.4b). For TL ~ 11 m, the two formulas tend to converge and has lower differences.

These findings are consonant with other studies (Gordon et al., 2011; Caruso et al., 2015). Thus, the two equations were used separately for a better reliability of the total length estimates. For TL ≤ 11 m, Gordon's equation was adapted. For TL > 11 m Growcott's equation was used (Caruso et al., 2015; Poupard et al., 2022).
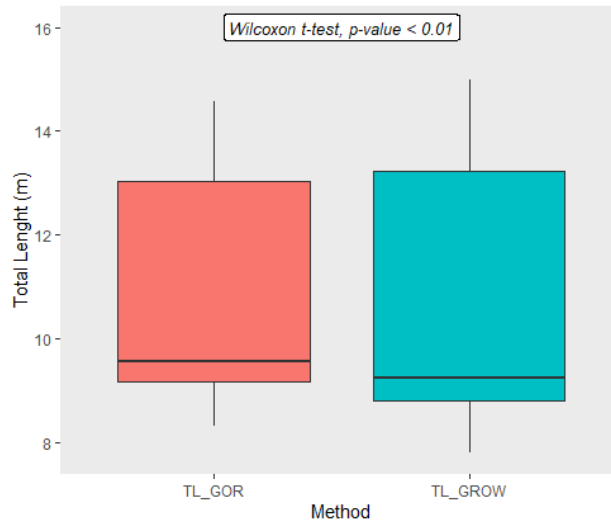
Figure 3.5: Boxplot of total length measurements (m) obtained from PAMGuard with Gordon's equation (orange) and Growcott's equation (green).

3.2.3   Comparison between results from Raven and PAMguard

For the comparison between the two methods applied (manual and automatic), the 13 IPI measures common to both methods were used. Both datasets showed a non-normal distribution. In the scatterplot of the differences, data were only partially distributed around the 0 line (i.e, line of differences = 0). Part of the estimations showed a difference $\neq 0$. This is reflected in the boxplot (Figure 3.6) where outliers are present. The two datasets do not correspond perfectly. The variability is low in both datasets, with only few measurements highly greater than the mean (the outliers). However, no statistically significant difference was found (Wilcoxon t-test, $p > 0.05$).
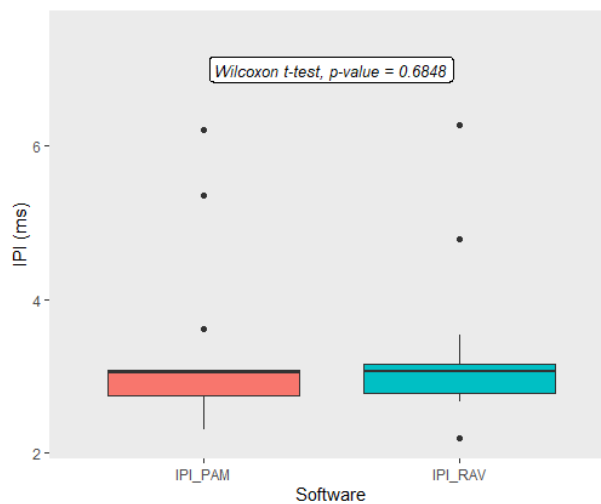
Figure 3.6: Boxplot of IPI measurements (ms) obtained from PAMGuard (orange) and Raven (green).

Distribution of the total sperm whale body length resulting from the two methods show similar distributions and differ mostly on two length classes 7.7-8.1 m and 11.7-12.1 m for Raven, and 8.1-8.5 m and 12.5-12.9 for PAMGuard (Figure 3.7a, 3.7b). Total length was computed as discussed in 3.2.2.
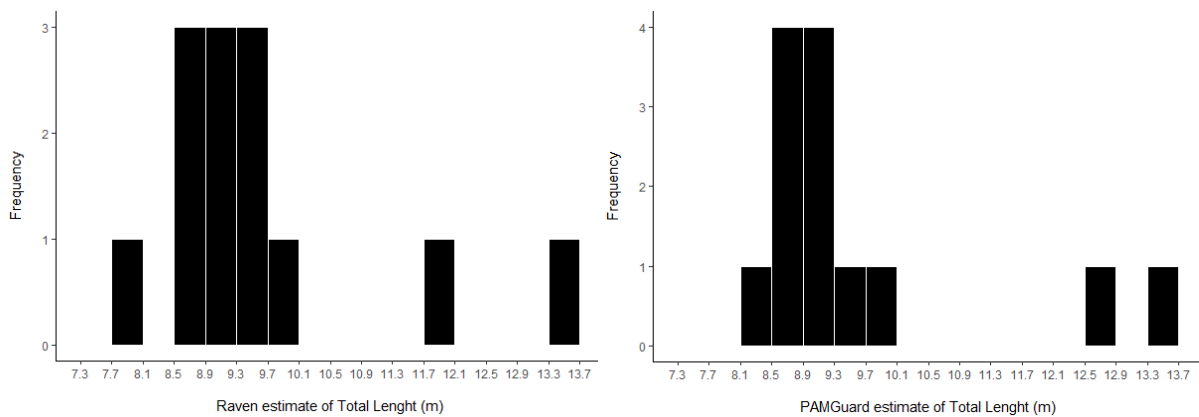
Figure 3.7: Histogram of the total length estimates in metres obtained with manual analyses in Raven a) and automatic analysis in PAMGuard b)

3.2.4 Population size structure

IPIs values resulted by PAMGuard analyses were then used for the population size structure. Gordon's (1991) and Growcott's et al. (2011) equations were integrated to obtain a mean total length, as described in paragraph 3.2.2. The resultant dataset was plotted with a histogram (Figure 3.8). Final estimations of sperm whale total body length ranged from 7.79 m to 14.56 m. The length classes between approximately 10 m and 13.7 m are less represented (< 5 measures) than other classes.
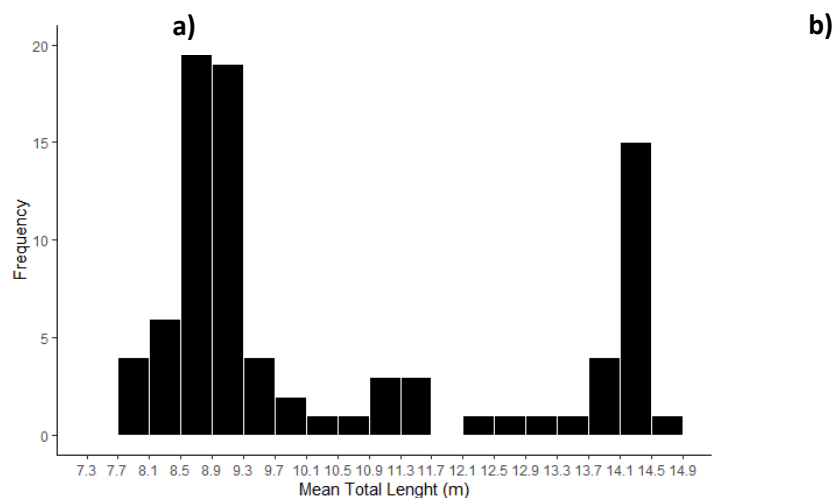
Figure 3.8: Histogram of the mean total length estimates in metres from PAMGuard IPI measurements

Total body lengths were then associated with general age classes and represented as a barplot (Figure 3.9a). No big differences were found between the three categories. However, the "immature" cluster (TL < 9 m) was the most represented with 34 estimations; 29 were grouped in "adult females and juveniles males" (9 m < TL < 12 m) and 24 in "Adult males" (*TL* > 12 m).

To integrate the age classes distribution, seasons correspondent to the measurements was taken in consideration. A total count per season found that 46 estimations came from summer trips, 20 from spring, 16 from autumn and only 5 from winter. Correspondent number of measurements per age classes were counted. These data were plotted with a histogram, showing together seasonality and age distribution (Figure 3.9b).

All the classes resulted to be present during the four seasons, with except of the "adult females and juveniles" in wintertime. However, only 5 measurements were available for winter months. Total counts per season (red line) shown to see the different availability of measurements during the different seasons.
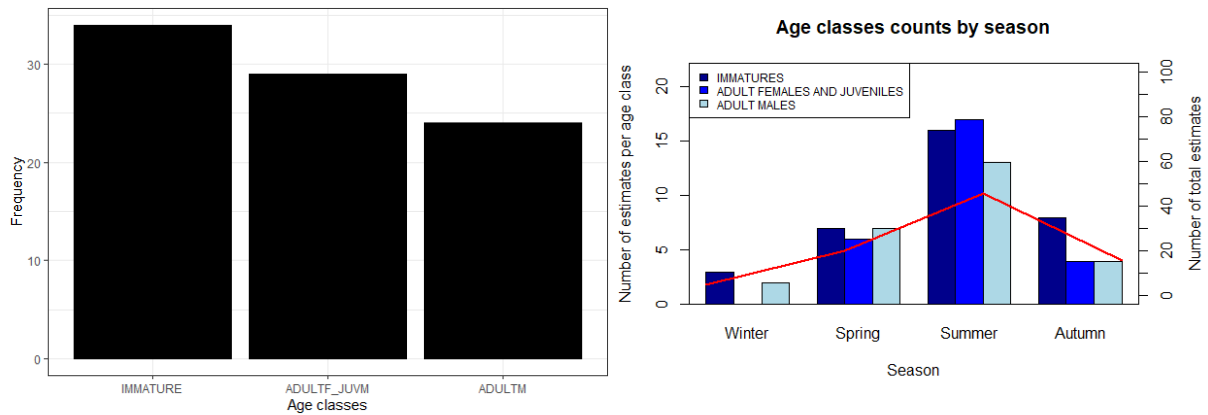
Figure 3.9: a) Barplot of the frequencies per age classes: Immature, adult females and Juvenile males, adult males. b) Histograms per age classes during the four seasons. Red line = number of total estimates per season.

3.2.5   Different size estimation of the same individuals

During the sampling period, 6 individuals were photo-identified during the trips from the expert biologists onboard. These individuals are already included in the catalogue of Futurismo. Of these 6 individuals, only 2 were recorded more than once. In addition, one individual was recorded twice in one trip (one recording for codas and one for usual clicks). These 3 individuals were named as *Orca (known adult female)*, *Penakabe (known adult female)* and *PM001 (adult female*, as recorded with its calf). For these individuals, two measurements each from PAMGuard were available. For *Orca*, two recordings from June 2019 and April 2021 were available. *Penakabe* was recorded once in June 2019 and once in May 2021. PM04 was recorded in September 2021 (Table 3.3).

Table 3.3: Table of the individuals measured more than once, with correspondent IPI (ms) and TL (m).

CODE	PM	IPI (ms)	TL (m)
20190630_038	*Orca*	2.74	8.81
20210415_006	*Orca*	2.88	9.00
20190606_007	*Penakabe*	3.04	9.24
20210527_018	*Penakabe*	3.07	9.29
20210928_037	*PM01*	2.79	8.88
20210928_038	*PM01*	2.75	8.82

According to these results, *Orca* should have been grown of about 0.19 m in around two years. *Penakabe* of about 0.05 m in one year. PM01 measurements showed a difference of 0.06 m between the estimate with its coda's clicks (recording 037) and its usual clicks (038).

For the other 4 individuals identified, one measurement was available (Table 3.4).

For PM01 and Mr November, the IPI estimation with Raven was done. PM01 was estimated as 8.87 m and Mr November as 13.63 m. Resulting mean length for PM01 was 8.86 m and for Mr November 13.59 m.

Table 3.4: Table of the individuals measured one time, with correspondent IPI (ms) and TL (m).

CODE	PM	IPI	TL
20190606_006	*Coconut*	2.88	9.00
20190712_046	*Mr Liable*	6.60	14.04
20201123_005	*Mr November*	6.21	13.55
20210415_007	*Dylis*	2.77	8.85

3.2.6 Comparison in the subset 1: first period vs. second period

A total of 15 recordings were collected within the first period, while 46 recordings were done in the second period. A general increase in the ranking can be observed from the first period to the second (Table 3.5). While only one "5" recording was present, from the first period, major part of the recordings with a rank ≥ 3 came from the second period protocol. To be considered the disproportion in the number of audios collected during the two periods. Rate of acceptance for the first period was of 14.3%, while for the second period 20.9% (0 recordings excluded, as the classification is based on their short duration and no other factors). An increase on the acceptance, and so on quality, can be already seen.

Table 3.5: Table of the ranking counts per different protocols of the subset 1

Ranking	First period	Second period
0	1	2
1	9	4
2	3	31
3	1	5
4	0	4
5	1	0

Regarding the IPI extraction, from Raven the recordings used from the first period are 2, while for the second period are 9. With PAMGuard, 10 measurements from 9 recordings were obtained from the former; 58 measurements from 38 recordings were obtained from the latter. The rate of acceptance for PAMGuard was 64.3% for the first period and 86.4% for the second period. These percentages show two main things: 1) the increasing quality of the recordings from first period to second period and 2) the higher rate of acceptance of the automatic software.

Due to the large disproportion between the two datasets, no statistical test was run to state the eventual difference between them.

4   Discussion

In this work we have analysed acoustic data collected from whale watching vessels off São Miguel, Azores. Despite the challenges, these opportunistic acoustic recordings became useful to estimate sperm whale size based on IPIs calculated manually, automatically with the software PAMguard or automatically with the software CABLE. Differences between the three methods were mainly on the number of recordings usable. For manual analyses, only a smaller part of the dataset was useful (9.1%), while the automatic methods used 56.8 % of the entire dataset. CABLE software gave results only for 2 recordings. Average IPI, divided for manual and PAMGuard datasets, were used to calculate estimations of total body length of the sperm whales recorded.

4.1   Dealing with opportunistic data

1.   Limited audio recordings (due to WW operation and limited dedicated surveys)

Collecting research data during touristic' trips is challenging and has a multitude of limitations, as research is secondary to touristic activity itself. During whale watching trips, time is limited per observation, both legally (maximum of 30 min according to DLR 10/2003/A), and in practice, to maximize chances of observing different species during the same trip, improve tourist comfort onboard avoiding long times stopped, and complying with established commercial timings of operation. This combination often leads to a reduced possibility to record and, in case, lower duration of the recordings. These circumstances lead to a limited number of recordings available, which are nevertheless of interest to test its usability for sperm whale size estimation, and finally provided a first approach to the size structure of the individuals recorded. In fact, some other studies also used a limited number of recordings and they found consistent results for the size distribution of the Mediterranean sperm whale population (Drouot et al., 2001).

Recordings used in this study applied two different protocols with different goals, therefore the two obtained datasets present slightly different characteristics. Subset 2 was collected in 2019 with the goal of development of a pilot study on the acoustics of sperm whales related to the number of boats in the area (not aiming for the present study). Four important problems were found: the Magrec HP30 hydrophone, the depth of the hydrophone not standardized (i.e., we do not know the depth of the hydrophone at each recording), the CH4 mode used for recordings and the file extension. The use of the Magrec HP30 hydrophone lowered the overall quality of the recordings. As an old hydrophone (with an old-fashioned pin), it was not possible to connect to a proper recorder and extract directly the data. The method for recordings extraction worked by recording again on the computer. This lowered the quality of the recorded file, together with the low quality of the voice recorded used. The non-standardized depth affected the overlap of the surface reflection to the original clicks, with a variable level of overalp. The CH4 mode led to a fixed cut frequency at 22.05 kHz. Indeed, discrimination of pulses for manual analysis with the spectrogram was impossible for recordings after May 2019. Another problem was The WAV extension is important, as all the acoustic software require it. Moreover, MP3 is a compressed format, that exclude part of the information recorded to decrease the size of the final audio. Converting files from MP3 to WAV will not allow to recover the information lost.

Subset 1 recordings were only collected aiming for this objective after November 2020, therefore, the first part of the dataset presents some limitations. Firstly, the hydrophone location with respect to the whale was not standardized. Secondly, data from 2019 until November 2020, follow the same CH4 settings as subset 1, so same problem of a cut frequency at 22.05 kHz occurred. After November 2020, data collection was made with the specific purpose of this thesis, therefore following a more standardize way. Important changes were the sampling frequency, bit depth, recording mode and the file extension. While bit depth did not result useful for automated analysis with PAMGuard, the other changes were

fundamental. Higher sampling frequency increase the number of samples that the software uses to detect clicks and compute the IPI, enhancing the reliability of the results. The recording mode set to STEREO allowed a 96 kHz sampling frequency and reduced the noise level in the recordings, as it is implemented also for field recordings. Same problem of the subset 2 with the extension was found.

An important point to be implemented in the future should be the time of recordings. As described by Miller (2010) and Growcott (2010), with PAMGuard software the recordings till 15 min after fluke-up can be used. A greater duration means a greater sample of clicks from which extract the IPI. Recordings shorter than 5 min should be avoided. Manual analysis should be restricted to 8 min after fluke-up (Rhinelander and Dawson, 2004). However, with less than 5 min audio, only a small number of analysable clicks can be found. This is consistent with Alder-Fenchel (1980) that found that, on average, only 11% of IPI is feasible for IPI extraction. From CABLE testing made by its developer, 4 min recordings were found to be the minimum time for a reliable result, but a minimum of 10 min is suggested because of the high rate of rejection of clicks (Beslin et al., 2018). However, different studies used 5 min recordings for their studies (Douglas et al., 2005; Caruso et al., 2015), so it is suggested as a minimum for standardization with other methods. The time spent to record varies mostly from the conditions of the single whale watching trip. Groups of tourists interested in the research part of the trips were found to be more feasible for longer recordings. In order to merge the whale watching activity with the research interest, each situation should be assessed independently before proceeding with the recording. It should be preferred to make only recordings that can be long at least 5 min.

2. Training

On another hand, the hydrophones and recorders, although not very difficult to use, require some training to use them properly and in time during the trips. Onboard there is usually just one guide/biologist guiding the trip and collecting the data, therefore, sperm whale diving time, engaging conversation with the clients and occurrence, photo-ID and acoustic data collection should be combined in a perfectly time management. To do so, a clear and step-by-step protocol should be available, and regularly explained and trained with staff who will be responsible for its use.

3. Data quality

Acoustic recordings quality is greatly hindered by external factors such as the environmental conditions, which are indeed one of the major limitations. The other big portion of quality is usually related with hydrophone and recorder quality and settings established.

Audio recordings with good quality require low (or none) levels of noise. This is often difficult to manage, both due to the weather conditions (rough seas -wind and waves- are noisier) or boats at sea when recording. Ambient noise was found to increase with sea state, in relation to local wind speed and wave height (Hildebrand, 2009). With rough sea conditions, recordings' quality lower because wind and waves noise is predominant over the clicks of the animals and the hydrophone stays nearer the surface. The former can be minimized by collecting data only with Beaufort ≤4 (Rhinelander and Dawson, 2004). The latter can be adjusted by testing the number of sinkers that can be added to the cable avoiding an increase on the noise produced by the sinkers. The two sinkers added after September 2021 improved the cable straightness, however it was not always sufficient. Finally, it should be preferred to record when no other boats are in the area. We must say that only with trial and error *in situ* in different situations, the best settings and equipment adjustments are met.

The hydrophones used for the collection of the data used in this study were mainly basic. This was mainly seen in two characteristics: sensitivity and frequency response. Sensitivity is an important characteristic of a hydrophone, as it is the expression of how close it can convert the acoustical wave into an electrical signal. Sensitivity range of the hydrophones was medium (usual range, -220/-150 dB). More negative ranges mean less sensitivity. For distant source of sound, as it might be a sperm whale, higher sensitivity levels are required. The reason is that, while propagating away from the primary

source, the sound spread around and attenuates with a variable rate, depending on the specific conditions of the sea. At higher distance, environmental factors assume an important role. The temperature, for example, change the sound speed; and the water column thermal structure affects the propagation path of the sound itself. This will result in different interactions with the ocean bottom and the surface and so the distance at which the sound can be recorded (Hildebrand, 2009). For instance, this value can be somehow reflected on the mixed layer depth, which here in the Azores reaches a depth of around 10 m only during summer months (June-August) – seasonal thermocline-, while from October to March is deeper than 30 m (even close to 100 m in winter) (Coelho, 2021). This means that, with the 15 m cable, the hydrophone is almost always shallower than the mixed layer, which affects the sound propagation from below. The 25 m was used fewer times, as it was found to be unfeasible to recollect it during commercial whale-watching trips, due to the extra time needed to take it out of the water.

Moreover, a different interaction with the surface will change the possible overlap between original clicks and reflected clicks. A higher sensitivity is also reflected in less preamplification gain necessary, in order to avoid clipping of loud clicks. Clipping leads to a lower signal-to-noise ratio (SNR), so that it is harder to distinguish the sound from the background noise (Miller, 2010). Clipping is a distortion of the audio signal when the level is too high for the recording system (Hahn, 2021). With sperm whale clicks as the highest biological sound ever recorded (Møhl et al., 2003), this can occur easily. In this study, the level of the recorders was not controlled through the entire duration of the recordings. Thus, it is a parameter to be implemented for further data collection.

The frequency response is another important characteristic of a hydrophone. It is the range of frequencies for which the hydrophone keeps a regular sensitivity within a predefined range. Variations from this flat sensitivity are usually given in dB. Smaller values of variations means that the sensitivity (and so the output of the hydrophone) vary less from the flat value. When out of this frequency range, in which the sensitivity is constant, the sensitivity tends to decrease, till the lowest value at a certain frequency. The Aquarian H2a has a greater sensitivity in the range 20 Hz – 4 kHz and then it decreases till the lowest sensitivity at 100 kHz. It means that the hydrophone has good sensitivity for low frequencies. As sperm whales clicks contains most of the energy in the band of 2 kHz – 40 kHz (with peaks of energy below 4 kHz) (Thode et al., 2002), it means that for higher frequencies and lower energy source, the hydrophone would have

As a summary, all the previous consideration are important to realize that, probably, one of the main limitations of this dataset was the low signal-to-noise ratio for many of the recordings.

Based on all this information and experience, we will recommend for future sperm whale recordings, a hydrophone with a sensitivity of at least -180 dB, and a frequency response of at least 20 Hz-40 kHz ± 4dB. Data collection should be done with the calmest seas possible to avoid external sources of noise; and using a long cable will help to reduce noise from the surface, and hopefully minimize thermocline effect on sound propagation.

4.2   Acoustic analyses

The quality classification of the recordings from subset 1 showed a consistent result with findings from previous studies (Alder-Fenchel, 1980; Gordon, 1991; Rhinelander and Dawson, 2004). Only a small proportion of recordings contains clicks of enough good quality for manual extraction of IPI. For subset 2, the greatest problem was the cut frequency at 22.05 kHz that has made impossible to distinguish the multipulsed structure withing the spectrogram. Other factors that may have an impact on the quality of the recordings (such as the n° of boats or other species presents) could not be taken in consideration because data were not available for all the recordings.

Møhl et al. (2003) found that a clear multipulsed structure is represented only when recordings are taken on-axis, i.e., directly in front or behind the animal. The direction in respect of a single hydrophone

cannot be stated, but, in relation to the diving profile of a sperm whale, it can be assumed that the animal would be in front of the hydrophone in the descent phase. For this reason, establishing in the protocol that recordings should be made always from behind the animal is of great importance to increase chances to get on-axis recordings. Occurrence of creaks in the bottom phase (i.e., time between the end of descent and beginning of ascent) (Watwood et al., 2006) are often associated with changes in the direction (Mathias and Thode, 2012).

The IPI estimated for the recordings had a SE range ±0.01 to ±0.02 ms with a mean of ±0.01 ms, smaller but consistent with that found by Caruso et al. (2015) (SE ±0.03). The CV showed a wide range (1.92-4.35%) as reported by other studies (Rhinelander and Dawson, 2004; Antunes, 2009). The accuracy of the manual method can be high; however, the precision may lack due to the method applied. In Raven, the IPI resultant was always precise to one decimal, so that a large approximation between different measurements was applied. To avoid errors, a large number of clicks is averaged to get the true IPI estimate. No clear relation was found between the number of clicks used for the IPI estimation and the duration of the recordings. However, it is just a matter of probability. For longer recordings, a larger number of clicks would be available, as sperm whales click with a click rate of 1.2 clicks/s (Beslin et al., 2018). Thus, a greater number of usable clicks would be present.

An important point of the manual analysis was the procedure to recognise the click trains and so associate them with the IPI range of the single individual. A precise discrimination was not possible for individuals with a similar IPI. Thus, different individuals' IPI might have been merged in the same measure. A strange result was the one from the recording 20201123_005, relative to the whale "Mr November", known as a mature male. In this case, measures from P0-P1 and P1-P2 were discordant (6.28 and 4.78 ms, leading to 13.63 m and 11.75 m, respectively) and have been retained both in the dataset as a comparison with PAMGuard efficiency. For the few clicks that showed also a P3 pulse, the P2-P3 time delay was measured and found to be consistent with the P0-P1 delay. No overlap between the original clicks and the surface reflection was seen in the clicks analysed. This is due to the methodology. In case of overlap between the original click and its surface reflection, the clicks have been discarded from the analyses. A possible explanation is the presence of a P1/2 pulse, described by Zimmer et al. (2005). This pulse is visible in off-axis recordings and it may appear overlapped with P0 or P1 pulses or it can be anytime in between the P0 and P1 pulses. In this perspective, the supposed P2 should be a P1/2 pulse and so this may explain the difference in the IPI. The hypothetical P1/2-P2 difference was found to be quite stable, however it showed the greatest SE (1.8%).

From the PAMGuard analysis, only 63 recordings were appropriate for analyses (55.8% of the dataset). This low rate of acceptance is explained by the same reasons above (i.e., quality of the recordings based on the equipment and the protocols applied). In total, 52 recordings more were accepted than in the manual analysis. This was the main reason to further investigate the usage of a totally automated software. As already described, PAMGuard works with a cepstral analysis and it is able to compute many more clicks than the ones that would be feasible with manual analysis. Moreover, the average ensemble technique used by the plugin allows to lower the off-axis effect and let more clicks suitable for the IPI estimation, as found by Teloni et al. (2007). In addition, longer recordings can be used as input in PAMGuard, as the change in IPI with the depth can be minimized by averaging a bigger number of clicks. Finally, larger datasets can be applied, as PAMGuard is totally automated and requires only a little effort from the operator. PAMGuard yielded 87 IPI estimations, which lead to 87 different total lengths estimated. No attempts were made to infer the abundance of individuals recorded and so the number of individuals estimated. This means that, possibly, the same individual might have been recorded (and therefore, measured) more than once. Moreover, for the same recordings, it was impossible to distinguish individuals with a similar IPI. Only clear peaks were considered for the IPI estimation, so animals with a similar stable IPI (< 0.3 ms) were not discerned.

Analyses run with CABLE were used to test the possible usage of the software with this kind of recordings. Results showed that, on a total of 113 recordings, only 2 produced reliable results with the default settings. Only with default settings the results were found to be reliable from the developer. With these settings, they found that the GMMs showed good fit and the accepted IPI measurements were coherent with the ones obtained from manual analysis (Beslin et al., 2018). These results are not surprising. CABLE requires three criteria to be observed to accept the clicks: 1) should be on-axis clicks 2) should be clearly distinguished among other clicks or echoes; and 3) SNR of the single click must be high. Only a few clicks will respect these criteria, especially considering the quality of the recordings discussed above. However, Beslin et al. (2018) found that the acceptance rate is lower than a manual analysis on the same dataset. This is because of strict filters to the classification and measurements that the software applies (i.e., "Overall goodness" and "IPI repetition" parameters). Decision to maintain this rate was done to avoid false positives, even if a great number of false negatives have been reported. In this way, even with as few as 5 samples per peak on IPI distribution allows to estimate the true IPI. With a rate of acceptance of 0.7% of the clicks and an estimated click rate of 1.2 clicks/s, a minimum recording of 10 min is required (Beslin et al., 2018). Testing the software behaviour by changing *Overall goodness* and *IPI repetition* parameter allowed to identify the main problem with these recordings, i.e., too low SNR. Lowering the *IPI repetition* did not product changes in the acceptance rate. However, lowering the *overall goodness* produced a larger number of clicks accepted, from a value of 0.3, with both 0 and 1 IPI repetition. The *overall goodness* controls the strictness of the classification, so that a certain misclassification from the software can be supposed. This misclassification can be explained with two main reasons: 1) a low SNR of the recordings 2) a difference in the type of recordings and region (as CABLE was developed with the sperm whales' population in Dominica). The second reason can be true mostly for recordings that were seen good in the previous analyses. In this perspective, results coming from lowering the criteria should not be seen as reliable, unless a good comparison with other methods can be done.

Finally, an important suggestion comes from the developer (Beslin et al., 2018). The software may perform better and may be more useful with big datasets and from long-term autonomous recordings projects. Availability of longer recordings in duration and from sperm whales at an unknown aspect may be more suitable for CABLE, allowing researchers to easily obtain a great amount of data. Therefore, with a bigger dataset and better-quality recordings, it would be of interest to test again this software.

4.3   Statistical analysis

The limited sample size for most of the classification levels did not allow to perform a statistical analysis to understand which factors affect more the quality of the recordings. In this sense, sea state (Beaufort and Douglas), sampling frequency, number of boats in the area and season (related to the depth of the mixed layer) should be taken into consideration for further studies. Due to the nature of categorical and ordinal dependent variable (i.e., ordered classification) and categorical factors, an Ordinal Logistic Regression model should fit best (Abreu, 2008; Harrell, 2015).

Two methods were used to obtain IPI with PAMGuard: average ensemble method and histogram. Both methods obtain a bimodal distribution of IPI (as found by Poupard et al., 2022). No significant difference was found between IPI values of both methods, so they were considered comparable and IPI were averaged to obtain a unique value to compute the total length equations (2.1.) and (2.2). These findings are consistent with those of Growcott (2010) and Miller (2010). However, it has to be highlighted that from the findings of Growcott (2010) and Miller (2010), the ensemble seemed to give more robust estimations when compared with the manual method used by Rhinelander and Dawson (2004).

The comparison of the two formulas of Gordon (1991) and Growcott et al. (2011) showed significantly different results, as tested with the Wilcoxon test. Differences in the values and distribution of the total body lengths obtained were found. Significant differences between both formula results are consonant to the studies published after Growcott et al. (2011) which suggest a better fit for the Gordon's formula with TL ≤ 11 m and for the Growcott's formula with TL > 11 m (Caruso et al., 2015; Giorli and Goetz, 2020; Poupard et al., 2022).

Manual (Raven) and automatic (PAMGuard) methods for IPI estimations did not show significant differences. In our case study, due to the reduced number of recordings used for manual analyses, majority of IPI's were in the interval 2-4 ms (similar to the other methods), but some outliers were present and not a normal nor bimodal distribution pattern was found. The manual analysis is deemed to be highly reliable (Growcott, 2010; Caruso et al., 2015) and that can be seen from the low value of SE. However, it is time consuming, and it can be quite subjective in the choice of the IPI extremes, leading to problems of accuracy. In addition, manual method is not always feasible due to data quality issues. Waveforms and spectrograms of the clicks in field recordings are often not as clear as supposed by the model of Norris and Harvey (1972), but contain anomalous clicks (Antunes, 2009) and noise. Automatic IPI detection with softwares like PAMGuard can overcome this limitation and easily work with larger datasets.

The comparison between the two different periods (before and after December 2020) for subset 1 showed some significant results. A general increase in quality was found from period 1 to period 2. The great abundance of recordings ranked as "2" for the period 2 is mostly due to not clear multipulsed structure of the clicks or just present for too few clicks. This aspect cannot be really controlled, as it is inherent to the animal sound production, and the aspect-dependant nature of clicks is the major reason of rejection of clicks for manual analysis. On the other hand, most of the recordings from period 1 were classified as 1, due to high presence of noise and no clear presence of clicks. Additionally, also the rate of acceptance by PAMGuard was greater for recordings from the second period. We confirm that, generally, recordings from the second period were more suitable for the analyses, both for manual and automatic analyses. These differences may be explained by the improvements applied to the protocol for the period 2. To be taken in consideration is also the different sample size of the two datasets that may have affected the probability of good recordings.

4.4   Sperm whales size structure

Sperm whale size distribution was found to be bimodal, with high representation of the 8-9 m class and the 14 m class. The classes between around 10 m and 13 m were underrepresented in respect to the others. These size classes correspond to the older females and younger mature males. The small number of recordings (or individuals recorded) in the dataset can be one of the main reasons for these findings. However, another hypothesis could be that of individuals of these size being less present during the study period. Old females can reach according to literature a maximum size of 11 m (Evand and Hindell, 2004); however, mean sizes of mature females can be smaller (Whitehead, 2003; Whitehead, 2017) and therefore, adult females being perfectly represented with TL<10 m. On the other side, young mature males may not be sighted very frequently in the area. The size classes of 10-13 m correspond to young males that already left the mother's social unit and travelled north. But according to Whitehead (2003), mature males do not start to migrate back to females' grounds until they are in their twenties. Seasonality should not be a reason for this lack of size classes for males, as indicated by van der Linde and Eriksson (2019) who found no significant differences on the seasonal presence of mature males. However, they found a peak in abundances of females and juveniles during the summer. This can be explained by the fact that females tend to move in their entire home range during the year. Resightings are uncommon but tend to be mostly of females and during the same season, rather than between different years

(Matthews et al., 2001; van der Linde and Eriksson, 2020). A similar bimodal distribution was found by Poupard et al. (2022) in the Nort-West of the Pelagos Sanctuary, Mediterranean Sea, where most of the data correspond to females and juveniles. On the other hand, Caruso et al., 2015 in the Ionian Sea (Mediterranean Sea), obtained a normal distribution with peak in the 10-13 m size classes using different acoustic methods. However, in the Mediterranean Sea there is a closed subpopulation of sperm whales, with individuals moving through the entire basin (Carpinelli et al., 2014). Two different areas of the Mediterranean Sea were analysed in Caruso et al. (2015) and Poupard et al. (2022) and it could show that sperm whales of different ages tend to be found in different areas of the Mediterranean Sea. This may also occur in the area of São Miguel Island, as the size structure found refers only to a small study area. Finally, bigger size of males can be less frequent for the previous whaling activity, that focused mainly larger males (Clarke, 1956). Male sperm whales reach their maximum length of around 18 m only in their late forties. Whaling activity in the North Atlantic stopped around 30 years ago, so it is possible that more time is required to restore the social structure present in the pre-whaling era. Clements et al. (2017) found a progressive decrease in the mean size of sperm whales caught during the whaling period. However, Clapham and Ivashchenko (2018) stated that reliability of whaling data is questionable. Size and sex information from whalers were largely modified (or even invented) to cover the massive harvest of individuals smaller than the minimum size. Thus, to support such hypothesis, further studies should be conducted.

All the factors proposed (small sample, individuals of those sizes do not present during the study period, males underrepresented as migrated northern, small study area and whaling pressure) contribute probably to the best explanation for the size distribution found.

The more frequently obtained sperm whale body length were those smaller than 9 m (34 measures), followed by those larger than 12 m, i.e., "adult males" (29), and "Adult females and juveniles males" (9-12 m) (24 measures). In the immatures class, also mature females can be present. In fact, sexual maturity in females is reached around 8 to 9 m, when they are about 9 years old (Whitehead, 2003; American Cetacean Society, 2004). Of 34 measures, only 2 were < 8 m. We should note that IPI < 2 ms were excluded from the analyses to avoid bias in the automatic estimation with PAMGuard. Calves and newborns were excluded from the beginning. Thus, the class of immatures was mainly represented by immature males or larger females and even adult females, as they can be already mature from 8 m. "Juvenile males and adult females" class includes older females. As found by Evans and Hindell (2004), females reach their maximum length of around 11 m at 19-20 years. Males, on the other hand, reach sexual maturity at around 10-12 m, about the same age of the females (Gaskin, 1970; American Cetacean Society, 2004). So, this class may include also already mature males. According to these two groups, it is possible to say that the composition of the population by size/age in the dataset is coherent with other studies in the Azores (Steiner et al., 2012; Silva et al., 2014; van der Linde and Eriksson, 2019). The majority of the records belong to females and immatures. However, in this study mature males (i.e., TL > 12 m) represented 27.6% of all the measures.

Our results support the previous knowledge about the sperm whale size structure in the Azores and São Miguel. Acoustic data can thus provide a complementary view to the traditional studies including occurrence and photo-identification data, further investigating in a non-invasive way the size, and therefore, age estimation of the individuals. This becomes essential when addressing the health status of a population, particularly in important areas for the species, such as breeding grounds.

Due to the limited recordings along the year (64.9% was done in summer), no seasonality can be inferred by these data. However, it can be seen that all the classes are present in any season of the year. Adult males, as expected (Steiner et al., 2012; Silva et al., 2014; van der Linde and Eriksson, 2020) were less frequently measured than females and juveniles. To keep in mind looking at this data is that the Azores population, and São Miguel population, is an open population. This means that different individuals

enter and exit the area during the year and within different years. While mature males generally make great migrations to the north Atlantic feeding grounds, females units move within their home range at lower latitudes (REF!). A minor presence of both males and females during winter was seen also by other studies, but it is usually related to a reduced effort during winter months (Silva et al., 2014; van der Linde and Eriksson, 2020).

4.5   Measuring known individuals

Finally, seven of the individuals measured were already known in the area thanks to photo-identification. Only three of them had two different size estimates in the dataset, one of the individuals, both estimates during the same trip. The three individuals were recognised as adult females (including the ones known as *Orca* and *Penakabe* recorded with their calves). The comparison of the measures from PM01 enabled to see how accurate different estimations of the same individual are:only a difference of 0.04 ms (0.06 m) was seen between the two recordings. Even if it is not a great difference compared to the size of the animal, the error is greater than the range found in the manual analyses (± 0.01 - ± 0.02 ms). For the other two individuals, one measure was from 2019 and one from 2021. Difference in size was found between the two years for both whales (*Orca* body length estimate increased by 0.19 m, while *Penakabe* by 0.05 m). Growth in sperm whales is age-dependent (older individuals has lower growth rate), as in other mammals (Evans and Hindell, 2004). It can be studied by fitting multiple size estimations with von Bertalanffy and Gompertz growth equations (Evans and Hindell, 2004). No attempts to study growth rate were made in this study due to the lower number of estimates available per recognised individual. However, it could be of interest for future studies.

Size estimations were available for other 4 individuals already known in the catalogue. This information will be of interest for educational purposes and awareness raising among tourists. For two individuals, Mr November and PM01, the comparison between Raven and PAMGuard estimations was available. Mr November measures were 13.63 m with the former and 13.55 with the latter. For PM01, 8.87 m from Raven and a mean of 8.85 m from PAMGuard. The biggest difference is found for Mr November (0.08 m). In the manual analysis, subjectivity canaffect accuracy. Measurement for Mr. Liable and Mr November could be further discussed, as they are well known mature males (particularly the former) by the biologists of Futurismo. Mr Liable was estimated as 14.04 m, while Mr. November as 13.55 m with PAMGuard. For Mr. Liable it is known that in 2004 it was already a mature male. If at that time, it was around 10-12 years, its total length estimation could be plausible. However, this workaround is just instructive, as more data is needed to provide reliable results on growth and, so, on the reliability of these size estimations of these known individuals.

As a general recommendation, correlated information is highly useful for validation of the data. Sea state, number of boats and recordings settings should be always written to correlate the quality of the recordings to the environmental conditions. The age structure inferred during the sighting can be a validation for the size classes found with the acoustic estimations; moreover, the number of individuals recorded can be a guide to know the number of individuals present in every recording and the number of estimates obtained. Time of the recordings, date and GPS coordinates should be always present for identify the single recording and to associate with photo-ID.

4.6   Potential improvements

Although the usage of opportunistic data poses different limitations, it has a huge potential to acquire information otherwise unavailable. In recent years, the "citizen science" increased greatly all around the world (Giraud et al., 2014). In the case study, this reflects into the increase of the whale watching activity in the Azores (Silva, 2015; González García, 2019), and an increase on the information available coming

from this source (as the data considered here, or other datasets of the region already published – Azevedo et al., 2021). With a standardized protocol, the data collection could be improved, resulting in a vast dataset able to provide more reliable results. Following the findings of this study, we encourage a long-term project with sperm whale acoustics to be started. Different approaches to complete the acoustic results should be of interest, including of course, occurrence data and photo-identification. The latter will allow the study of growth rates of individuals over time. A size population structure combining the different data approaches will be more robust than just with acoustic estimations. To validate the data and the accuracy of the estimates of São Miguel Island population, a complementary photogrammetric method (i.e., photographic method to estimate the length of a whale with a single photo with known size objects for scale; or a single photo with an accurate range measurement; or stereo images) could be implemented from boat surveys or even drone. As done by Gordon (1991) and Growcott et al. (2011) the comparison between acoustic estimates and photogrammetric measures allows to have a better final estimation. If needed, a new regression for the total length estimate can be also calculated, as done by Growcott et al. (2011). Finally, as the population around the Azores is an open population, with individuals that comes from different parts of the North Atlantic, the same methodology proposed here could be test for the entire Azores.

5 Conclusion

Estimations of the total length of the sperm whales recorded off the south of São Miguel during whale watching activities, were possible despite the quality issues and the reduced sample size available. Different methods were tested to obtain the best protocol feasible for the Azorean whale watching to conduct a long-term study on sperm whale acoustics. Cable software was also tested, but for the kind of data collected in the study seemed to be not useful. Manual method revealed to be precise, in accordance with previous studies. However, it was time consuming, allowing only a small sample size, and possibly biased by subjectivity of the operator. On the other hand, automatic analyses with the software PAMGuard revealed to be affordable and reliable. On the entire dataset, 56.8% of the recordings yielded sperm whale body length estimations, in contrast with the only 19.9% for manual method. If only the subset 1 is taken in consideration, the percentage increases to 77.0% and if only the period 2 from subset 1 is considered, 82.6% of the recordings gave results. Therefore, we can say that the improvements implemented when methodology was applied following a dedicated objective (i.e., data collection for this dissertation), they did make a difference and provided better results.

The objectives of the study were thus explored:
- Findings from this study showed that it is possible to combine the acoustic research with the whale watching activity. Even though some compromises need to be accepted due to the nature of the main activity, the potential for further studies is vast.
- With 2 years of occasional data collection, a first baseline of the size structure of sperm whales occurring in the south of São Miguel Island was created. Only small vessels were required and basic equipment was used for the data collection.
- Since whale watching is an activity increasing all around the world (Silva, 2015; González García, 2019), future studies like this one could be replicated in a larger temporal (and even spatial) scale.

This study was only a preliminary insight in the possibility of using acoustic data coming from whale watching vessels as a tool for assessing sperm whale's population structure. Bigger datasets, with a long-term goal, should be used for a more complete analysis. Moreover, better-quality equipment is suggested to avoid loss of time and effort. The project here presented used the infrastructures and equipment already available, which show, despite of the limitations, to be good enough to obtain useful data, particularly applying some improvements to the existing acoustic data collection protocol. At the same time, chances to complement acoustic information with occurrence registers and photo-identification are quite high during whale watching trips in the Azores. Additionally, whale watching business provide an excellent opportunity to set up long-term projects, as far as they are not at a high cost, particularly regarding tourist satisfaction. It has been shown that the latter can be easily increased with good information provided by qualified guides or biologists onboard, which at the same time will effectively raise awareness among tourists.

List of figures and tables

Bibliography

Abreu, M. N. S., Siqueira, A. L., Cardoso, C. S., & Caiaffa, W. T. (2008). Ordinal logistic regression models: application in quality of life studies. *Cadernos de Saúde Pública*, *24*, s581-s591. Ordinal logistic regression models (scielo.br)

Afonso, P., Fontes, J., Giacomello, E., Magalhães, M. C., Martins, H. R., Morato, T., ... & Vandeperre, F. (2020). The Azores: a mid-Atlantic hotspot for marine megafauna research and conservation. Frontiers in Marine Science, 6, 826. https://doi.org/10.3389/fmars.2019.00826

Alder-Fenchel, H. S. (1980) Acoustically derived estimates of the size distribution for a sample of sperm whales ~Physeter catodon, in the Western North Atlantic. Canadian Journal of Fish and Aquatic Science. 37: 2358–61. https://doi.org/10.1139/f80-283

Alves, F., Quérouil, S., Dinis, A., Nicolau, C., Ribeiro, C., Freitas, L., ... & Fortuna, C. (2013). Population structure of short-finned pilot whales in the oceanic archipelago of Madeira based on photo-identification and genetic analyses: implications for conservation. *Aquatic Conservation: Marine and freshwater ecosystems*, *23*(5), 758-776. https://doi.org/10.1002/aqc.2332

American Cetacean Society (2004). *Sperm whale. Physeter macrocephalus.* file:///D|/WebSites/ACS/fact%20sheet%20PDFs/fackpack/spermwhl-pdf.htm (memberclicks.net)

Antunes, R. (2009). *Variation in sperm whale (Physeter macrocephalus) coda vocalizations and social structure in the North Atlantic Ocean* (Doctoral dissertation, University of St Andrews). Ricardo Antunes PhD thesis (st-andrews.ac.uk)

Aoki, K., Amano, M., Mori, K., Kourogi, A., Kubodera, T., & Miyazaki, N. (2012). Active hunting by deep-diving sperm whales: 3D dive profiles and maneuvers during bursts of speed. *Marine Ecology Progress Series*, *444*, 289-301. https://doi.org/10.3354/meps09371

Appendices | CITES 22 June 2021

Assembleia Legislativa Regional (2003). Decreto Legislativo Regional n°. 10/2003/A. Altera o Decreto Legislativo Regional n.o 9/99/A, de 22 de Março (*observação de cetáceos*). https://dre.pt/dre/detalhe/decreto-legislativo-regional/10-2003-217369

Audacity® software is copyright © 1999-2021 Audacity Team. Audacity ® | Free, open source, cross-platform audio software for multi-track recording and editing. (audacityteam.org)

Beslin, W. A., Whitehead, H., & Gero, S. (2018). Automatic acoustic estimation of sperm whale size distributions achieved through machine recognition of on-axis clicks. *The Journal of the Acoustical Society of America*, *144*(6), 3485-3495. https://doi.org/10.1121/1.5082291;

Boersma, A. T., & Pyenson, N. D. (2015). Albicetus oxymycterus, a new generic name and redescription of a basal physeteroid (Mammalia, Cetacea) from the Miocene of California, and the evolution of body size in sperm whales. *PloS one*, *10*(12), e0135551. https://doi.org/10.1371/journal.pone.0135551

Bogert, B.P., Healy, M.J. and Tukey, J.W., 1963, June. The quefrency alanysis of time series for echoes: Cepstrum, pseudo-autocovariance, cross-cepstrum and saphe cracking. In Proceedings of the symposium on time series analysis (Vol. 15, pp. 209-243).

Borges, P. A., Bried, J., Costa, A. C., Cunha, R. T. D., Gabriel, R., Gonçalves, V., ... & Boieiro, M. (2010). Description of the terrestrial and marine biodiversity of the Azores. *A list of the terrestrial and marine biota from the Azores*, 9-33. http://hdl.handle.net/10400.3/2095

Borges, P.A.V., Costa, A., Cunha, R., Gabriel, R., Gonçalves, V., Martins, A.F., *et al.* (eds) (2010). *A List of the Terrestrial and Marine Biota From the Azores*. Princípia, Cascais. ISBN: 978-989-8131-75-1 p. 344-345

Brito, C. (2008). Assessment of catch statistics during the land-based whaling in Portugal. *Marine Biodiversity Records*, *1*.

Caddell, J. R. (2009). *International law and the regulation of cetaceans: An analysis of the role of global and regional instruments and institutions in the conservation of marine living resources*. Cardiff University (United Kingdom). 00001t.tif (cardiff.ac.uk)

Caldeira, R., & Reis, J. C. (2017). The Azores confluence zone. *Frontiers in Marine Science*, *4*, 37. https://doi.org/10.3389/fmars.2017.00037

Carpinelli, E., Gauffier, P., Verborgh, P., Airoldi, S., David, L., Di-Méglio, N., ... & De Stephanis, R. (2014). Assessing sperm whale (Physeter macrocephalus) movements within the western Mediterranean Sea through photo-identification. *Aquatic Conservation: Marine and Freshwater Ecosystems*, *24*(S1), 23-30. https://doi.org/10.1002/aqc.2446

Carrier, D.R., Deban, S.M., Otterstrom, J. (2002). The face that sunk the Essex: Potential function of the spermaceti organ in aggression. Journal of Experimental Biology. 205: 1755-1763. https://doi.org/10.1242/jeb.205.12.1755

Caruso F, Sciacca V, Bellia G, De Domenico E, Larosa G, Papale E, et al. (2015) Size Distribution of Sperm Whales Acoustically Identified during Long Term Deep-Sea Monitoring in the Ionian Sea. PLoS ONE 10(12): e0144503. https://doi.org/10.1371/journal.pone.0144503

Chami, R., Cosimano, T. F., Fullenkamp, C., & Oztosun, S. (2019). Nature's Solution to Climate Change: A strategy to protect whales can limit greenhouse gases and global warming. *Finance & Development*, *56*(004). https://doi.org/10.5089/9781498316880.022

Childers, D. G., Skinner, D. P., & Kemerait, R. C. (1977). The cepstrum: A guide to processing. *Proceedings of the IEEE*, *65*(10), 1428-1443. https://doi.org/10.1109/PROC.1977.10747

Clapham, P.J. and B aker, C.S. 2 002. M odern wh aling. In: Perrin, W.F., Würsig, B. & Thewissen, J.G .M. (eds.), Encyclopedia of Marine Mammals, pp. 1328-1332. Academic Press, New York. 01-whalingemm.pdf (whalenet.org)

Clarke, M.R. (1970). Function of the spermaceti organ of the sperm whale. Nature 238: 405-406. https://doi.org/10.1038/228873a0

Clarke, M.R. (1978). Buoyancy control as a function of the spermaceti organ in the sperm whale. Journal of the Marine Biological Association of the United Kingdom. 58: 27-71. https://doi.org/10.1017/S0025315400024395

Clarke, R. (1954). *Open boat whaling in the Azores: the history and present methods of a relic industry* (Vol. 26). University Press. https://indopacificimages.com/wp-content/uploads/2017/08/Open-Boat-Whaling-in-the-Azores_Robert-Clarke.pdf

Clarke, R. 1956. Sperm whales of the Azores. Discovery Rep. 28:237-98. pp. 282 Discovery reports (acorestraditions.com)

Coelho, Inês. 2021. Habitat preference of sperm whales (Physeter macrocephalus) off São Miguel island, Azores. MSc. dissertation. University of Algarve, Portugal.

Cornell Lab's K. Lisa Yang Center for Conservation Bioacoustics: Cornell Lab of Ornithology (2017). Raven Light 2. *Cornell University*. http://ravensoundsoftware.com/software/raven-lite/;

Cranford, T. W. (1999). The sperm whale's nose: sexual selection on a grand scale? 1. *Marine mammal science*, *15*(4), 1133-1157. https://doi.org/10.1111/j.1748-7692.1999.tb00882.x

Culloch, R.M, Davies-Marshall, P., Gallego, P., Krzyszczyk, E., Lamb, J., Lott, R., McCully, S., McGee, J., Nelson, M., Nuuttila, H., Scullion, A. J., Tetley, M. J., Vanman, C., Weller, M., Reid, R. J., and Goold, J. C. (2004) Sound velocity in spermaceti oil from different locations in the head of a sperm whale (*Physeter macrocephalus*). In *Proceedings of the eighteenth annual conference of the European Cetacean Society*. https://www.europeancetaceansociety.eu/sites/default/files/ECS%20Conference%20Proceedings%2018%20200 4.pdf

Douglas, L. A., Dawson, S. M., & Jaquet, N. (2005). Click rates and silences of sperm whales at Kaikoura, New Zealand. *The Journal of the Acoustical Society of America*, *118*(1), 523-529. https://doi.org/10.1121/1.1937283

Drouot, V., Goold, J. C., & Gannier, A. (2001). Acoustically derived size distribution of sperm whales in the Mediterranean sea. *European Research on Cetacean*. ACOUSTICALLY DERIVED SIZE DISTRIBUTION OF SPERM WHALE (cetaces.org)

Evans, K., & Hindell, M. A. (2004). The age structure and growth of female sperm whales (Physeter macrocephalus) in southern Australian waters. *Journal of Zoology*, *263*(3), 237-250. https://doi.org/10.1017/S0952836904005096

Evans, P. G., & Teilmann, J. (2009). ASCOBANS/HELCOM Small Cetacean Population Structure Workshop. *Bonn, Germany: ASCOBANS*. Methods for Discriminating Populations (ascobans.org)

Fahlke, J. M., & Hampe, O. (2015). Cranial symmetry in baleen whales (Cetacea, Mysticeti) and the occurrence of cranial asymmetry throughout cetacean evolution. *The Science of Nature*, *102*(9), 1-16. https://doi.org/10.1007/s00114-015-1309-0

Fais, A., Johnson, M., Wilson, M., Aguilar Soto, N., & Madsen, P. T. (2016). Sperm whale predator-prey interactions involve chasing and buzzing, but no acoustic stunning. Scientific Reports, 6(1), 1-13. https://doi.org/10.1038/srep28562

Fordyce, R. E., & de MUIZON, C. (2001). Evolutionary history of cetaceans: a review. *Secondary adaptation of tetrapods to life in water*, 169-233. Evolutionary-history-of-cetaceans-A-review.pdf (researchgate.net)

Gallagher, L., Porteiro, F., Santos, R. S. Observer's Guide to the Oceanic Marine Life of the Azores (2013), p. 2, Publisher: FISHPICS® MARINE LIFE DESIGNS Editor: FISHPICS® ISBN: 978-989-98324-0-4

Gambell, R. (1993). International management of whales and whaling: an historical review of the regulation of commercial and aboriginal subsistence whaling. *Arctic*, 97-107. https://www.jstor.org/stable/40511500

Gardoki, M., González, L., van der Linde, M., Fernández, M., & Azevedo, J. M. N. Mr. Liable: A resident male sperm whale in the Azores? In *Poster presente at the 32ⁿᵈ European Cetacean Society Conference, La Spezia, Italy*. (20) (PDF) Mr. Liable: A resident male sperm whale in the Azores? (researchgate.net)

Gaskin, D. E. (1970). Composition of schools of sperm whales Physeter catodon Linn. east of New Zealand. *New Zealand Journal of Marine and Freshwater Research*, *4*(4), 456-471. Composition of schools of sperm whales Physeter catodon Linn. east of New Zealand (tandfonline.com)

Gillespie, D. Gordon, J. Mchugh, R. Mclaren, D. Mellinger, D. Redmond, P. Thode, A. Trinder, P. and Deng, X. (2008) Pamguard: Semiautomated, open source software for realtime acoustic detection and localisation of cetaceans. Proceedings of the Institute of Acoustics, 30: Pt 5. http://toc.proceedings.com/04143webtoc.pdf

Giorli, G., & Goetz, K. T. (2020). Acoustically estimated size distribution of sperm whales (Physeter macrocephalus) off the east coast of New Zealand. *New Zealand Journal of Marine and Freshwater Research*, *54*(2), 177-188. https://doi.org/10.1080/00288330.2019.1679843

Giraud, C., Calenge, C., Coron, C., & Julliard, R. (2014). Capitalising on opportunistic data for monitoring species relative abundances. *arXiv preprint arXiv:1407.2432*. https://arxiv.org/abs/1407.2432v2

González García, L. (2019). Cetacean distribution in São Miguel (Azores): influence of environmental variables at different spatial and temporal scales (Doctoral dissertation, Física aplicada). http://hdl.handle.net/11093/1218

Goold, J. C. (1996). Signal processing techniques for acoustic measurement of sperm whale body lengths. The Journal of the Acoustical Society of America, 100(5), 3431-3441. https://doi.org/10.1121/1.416984

Goold, J. C., & Jones, S. E. (1995). Time and frequency domain characteristics of sperm whale clicks. *The Journal of the Acoustical Society of America*, *98*(3), 1279-1291. https://doi.org/10.1121/1.413465

Gordon, J. C. (1987). The behaviour and ecology of sperm whales off Sri Lanka. PhD thesis, Darwin College,Cambridge, UK.

Gordon, J. C. (1991). Evaluation of a method for determining the length of sperm whales (Physeter catodon) from their vocalizations. *Journal of Zoology*, *224*(2), 301-314. https://doi.org/10.1111/j.1469-7998.1991.tb04807.x

Growcott, A. (2010). Measuring sperm whales via acoustics and photogrammetry (Doctoral dissertation, University of Otago). GrowcottAbraham2010MSc.pdf (otago.ac.nz)

Growcott, A., Miller, B., Sirguey, P., Slooten, E., & Dawson, S. (2011). Measuring body length of male sperm whales from their clicks: the relationship between inter-pulse intervals and photogrammetrically measured lengths. *The Journal of the Acoustical Society of America*, *130*(1), 568-573. http://dx.doi.org/10.1121/1.3578455

Harrell, F. E. (2015). *Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis* (Vol. 3). New York: springer. https://doi.org/10.1007/978-3-319-19425-7_13

Hartman, K. L., Visser, F., & Hendriks, A. J. (2008). Social structure of Risso's dolphins (Grampus griseus) at the Azores: a stratified community based on highly associated social units. *Canadian Journal of Zoology*, *86*(4), 294-306. https://doi.org/10.1139/Z07-138

Hauser, D. D., Van Blaricom, G. R., Holmes, E. E., & Osborne, R. W. (2006). Evaluating the use of whalewatch data in determining killer whale (Orcinus orca) distribution patterns. *Journal of Cetacean Research and Management*, *8*(3), 273. https://faculty.washington.edu/eeholmes/Files/Hauseretal2007.pdf

Hipp, R. D. (2020). *SQLite*. Retrieved from https://www.sqlite.org/index.html

Hirotsu, R., Yanagisawa, M., Ura, T., Sakata, M., Sugimatsu, H., Kojima, J., & Bahl, R. (2010). Localization of sperm whales in a group using clicks received at two separated short baseline arrays. *The Journal of the Acoustical Society of America*, *127*(1), 133-147. https://doi.org/10.1121/1.3268593

Hoyt, E. (2021). Whale and Dolphin Watching in Europe. Under Pressure, 86. UNDER-PRESSURE_Chapter07_WW_low-res_web.pdf (oceancare.org)

IUCN. 2021. The IUCN Red List of Threatened Species. Version 2021-3. https://www.iucnredlist.org. Accessed on 22-02-2022

IWC (2002). 34th Annual meeting transcript. *IWC, Verbatim record*. https://archive.iwc.int/?r=424

IWC (2018a). *PORTUGAL: THE AZORES Whalers turn their skills to whale watching*. Whale watching handbook. Downloaded in 2021. Portugal: The Azores | Whale Watching Handbook (iwc.int)

IWC (2018b). *SPECIES INFORMATION: Sperm whales*. Whale watching handbook. Downloaded in 2022. Sperm Whale | Whale Watching Handbook (iwc.int)

Jacobson, E. K., Yack, T. M., & Barlow, J. (2013). Evaluation of an automated acoustic beaked whale detection algorithm using multiple validation and assessment methods. http://dx.doi.org/10.13140/RG.2.1.1234.3923

Jaquet, N., Dawson, S., & Douglas, L. (2001). Vocal behavior of male sperm whales: Why do they click?. *The Journal of the Acoustical Society of America*, *109*(5), 2254-2259. https://doi.org/10.1121/1.1360718

Kawakami, T. A. K. E. H. I. K. O. (1980). A review of sperm whale food. *Sci. Rep. Whales Res. Inst*, *32*, 199-218. SC032199-218.pdf (icrwhale.org)

Lai, E. (2003). Converting analog to digital signals and vice versa. *Practical Digital Signal Processing*, 14-49. https://doi.org/10.1016/B978-075065798-3/50002-3

Madsen, P. T., Payne, R., Kristiansen, N. U., Wahlberg, M., Kerr, I., & Møhl, B. (2002a). Sperm whale sound production studied with ultrasound time/depth-recording tags. Journal of Experimental Biology, 205(13), 1899-1906. https://doi.org/10.1242/jeb.205.13.1899

Madsen, P. T., Wahlberg, M. & Møhl, B. (2002b) Male sperm whale (*Physeter macrocephalus*) acoustics in a high-latitude habitat: implications for echolocation and communication. *Behav Ecol Sociobiol* **53,** 31–41. https://doi.org/10.1007/s00265-002-0548-1

Mathias, D., Thode, A. M., Straley, J., Calambokidis, J., Schorr, G. S., & Folkert, K. (2012). Acoustic and diving behavior of sperm whales (Physeter macrocephalus) during natural and depredation foraging in the Gulf of Alaska. *The Journal of the Acoustical Society of America*, *132*(1), 518-532. http://dx.doi.org/10.1121/1.4726005

Matthews, J. N., Steiner, L., & Gordon, J. (2001). Mark-recapture analysis of sperm whale (Physeter macrocephalus) photo-id data from the Azores (1987-1995). Journal of cetacean research and management, 3(3), 219-226. 50 doc..50 chapter .. Page219 (psu.edu)

McClain, C. R., Balk, M. A., Benfield, M. C., Branch, T. A., Chen, C., Cosgrove, J., ... & Thaler, A. D. (2015). Sizing ocean giants: patterns of intraspecific size variation in marine megafauna. PeerJ, 3, e715. http://doi.org/10.7717/peerj.715

Michael Hahn (2021, November). *What is Clipping Audio and How to Fix It*. LANDR blog. What is Clipping Audio and How to Fix It | LANDR Blog

Miller, B. S. (2010). "Acoustic growth rates and three-dimensional localisation of sperm whales (Physeter macrocephalus) in Kaikoura," Ph.D. thesis., University of Otago, New Zealand, pp. 17–33.

Miller, B. S., Growcott, A., Slooten, E., & Dawson, S. M. (2013). Acoustically derived growth rates of sperm whales (Physeter macrocephalus) in Kaikoura, New Zealand. *The Journal of the Acoustical Society of America*, *134*(3), 2438-2445. https://doi.org/10.1121/1.4816564

Møhl, B. (2001). Sound transmission in the nose of the sperm whale Physeter catodon. A post mortem study. *Journal of Comparative Physiology A*, *187*(5), 335-340. https://doi.org/10.1007/s003590100205

Møhl, B., Wahlberg, M., Madsen, P. T., Heerfordt, A., and Lund, A. (2003). "The monopulsed nature of sperm whale clicks." J. Acoust. Soc. Am. 114(2), 1143–1154. https://doi.org/10.1121/1.1586258

Møhl, B., Wahlberg, M., Madsen, P. T., Miller, L. A., & Surlykke, A. (2000). Sperm whale clicks: Directionality and source level revisited. *The journal of the Acoustical Society of America*, *107*(1), 638-648. https://doi.org/10.1121/1.428329

Morato, T., Machete, M., Kitchingman, A., Tempera, F., Lai, S., Menezes, G., ... & Santos, R. S. (2008a). Abundance and distribution of seamounts in the Azores. *Marine Ecology Progress Series*, *357*, 17-21.

Morato, T., Varkey, D. A., Damaso, C., Machete, M., Santos, M., Prieto, R., ... & Pitcher, T. J. (2008b). Evidence of a seamount effect on aggregating visitors. *Marine Ecology Progress Series*, *357*, 23-32.

Mullin, K. D., Steiner, L., Dunn, C., Claridge, D., García, L. G., Gordon, J., & Lewis, T. (2022). Long-Range Longitudinal Movements of Sperm Whales (Physeter macrocephalus) in the North Atlantic Ocean Revealed by Photo-Identification. *Aquatic Mammals*, *48*(1), 3-8. http://doi.org/10.1578/AM.48.1.2022.3

NAMCO (2018, September 25). *Abundance Surveys – Counting Whales. NAMCO*. Last update: April 2020. Abundance Surveys - Counting Whales - NAMMCO

NAMCO (2021, May 28). *Sperm whale*. *NAMCO*. Sperm whale - NAMMCO. Last update: July 2021

Neves-Graça, K. (2004). Revisiting the Tragedy of the Commons: Ecological Dilemmas of Whale Watching in the Azores. Human Organization, 63(3), 289–300. http://www.jstor.org/stable/44127307

Niemeijer, S. C., Libotte, J., Hartman, K. L., Steiner, L., & Azevedo, J. (2013). The presence of the sperm whale (Physeter macrocephalus) around the Azores: a study of re-sightings from 2003-2012. In *27th Annual Conference of the European Cetacean Society*. Microsoft PowerPoint - 20130404_ECS_Poster_AD25_time_1705 [Compatibiliteitsmodus] (uac.pt)

Norris, K. S. & Harvey, G. W. A theory for the function of the spermaceti organ of the sperm whale. *NASA SP* **262**, 397–416 (1972). https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19720017437.pdf

O'Connor, S., Campbell, R., Cortez, H., & Knowles, T. (2009). Whale Watching Worldwide: tourism numbers, expenditures and expanding economic benefits, a special report from the International Fund for Animal Welfare. *Yarmouth MA, USA, prepared by Economists at Large*, *228*. 26 March 2004 (ecolarge.com)

Oliveira, C., Gonçalves, J., Magalhães, S., Prieto, R., Silva, M. A., & Santos, R. S. (2007). *Whale watching management in the Azores: An updated review of the regulations. International Whaling Commission. Anchorage, Alaska*. SC/59/WW7: 4pp. (19) (PDF) Whale watching management in the Azores: An updated review of the regulations (researchgate.net)

Pace, D. S., Arcangeli, A., Mussi, B., Vivaldi, C., Ledon, C., Lagorio, S., Giacomini, G., Pavan, G., & Ardizzone, G. (2018). Habitat suitability modeling in different sperm whale social groups. *The Journal of Wildlife Management*, 82(5), 1062-1073. https://doi.org/10.1002/jwmg.21453

Pavan, G., Hayward, T. V., 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. *The Journal of the Acoustical Society of America*, *107*(6), 3487-3495. https://doi.org/10.1121/1.429419

Pavan, G., Priano, M., Manghi, M., & Fossati, C. (1997). Software tools for real-time IPI measurements on sperm whale sounds. *Proceedings-Institute of Acoustics*, *19*, 157-164. SOFTWARE_TOOLS_FOR_REAL-TIME_IPI_MEASURE20170530-6530-1gmo1gk-with-cover-page-v2.pdf (d1wqtxts1xzle7.cloudfront.net)

Paxton, C. G. M. (2016). Unleashing the Kraken: on the maximum length in giant squid (Architeuthis sp.). *Journal of Zoology*, *300*(2), 82-88. https://doi.org/10.1111/jzo.12347

Peng, C., Zhao, X., & Liu, G. (2015). Noise in the sea and its impacts on marine organisms. *International journal of environmental research and public health*, *12*(10), 12304-12323. https://doi.org/10.3390/ijerph121012304

Poupard, M., Ferrari, M., Best, P., & Glotin, H. (2022). Passive acoustic monitoring of sperm whales and anthropogenic noise using stereophonic recordings in the Mediterranean Sea, North West Pelagos Sanctuary. *Scientific reports*, *12*(1), 1-13. https://doi.org/10.1038/s41598-022-05917-1

Prieto, R., Pham, C. K., Brito, C., & Morato, T. (2013). *Biomass removal from shore-based whaling in the Azores*. Fisheries research, *143*, 98-101. http://dx.doi.org/10.1016/j.fishres.2013.02.001

Proceedings of the Symposium on Time Series Analysis, Wiley, New York. Chapter 15, 209-243.

Rendell, L. E., & Whitehead, H. (2003). Vocal clans in sperm whales (Physeter macrocephalus). *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *270*(1512), 225-231. https://doi.org/10.1098/rspb.2002.2239

Rhinelander, M. Q., & Dawson, S. M. (2004). Measuring sperm whales from their clicks: Stability of interpulse intervals and validation that they indicate whale length. The Journal of the Acoustical Society of America, 115(4), 1826-1831. http://dx.doi.org/10.1121/1.1689346

Ridgway, S. H., & Hanson, A. C. (2014). Sperm whales and killer whales with the largest brains of all toothed whales show extreme differences in cerebellum. *Brain, Behavior and Evolution*, *83*(4), 266-274. https://doi.org/10.1159/000360519

Roman, J., Estes, J. A., Morissette, L., Smith, C., Costa, D., McCarthy, J., ... & Smetacek, V. (2014). Whales as marine ecosystem engineers. *Frontiers in Ecology and the Environment*, *12*(7), 377-385. https://doi.org/10.1890/130220

Rosel, P. E., Wilcox, L. A., Yamada, T. K., & Mullin, K. D. (2021). A new species of baleen whale (Balaenoptera) from the Gulf of Mexico, with a review of its geographic distribution. *Marine Mammal Science*, *37*(2), 577-610. https://doi.org/10.1111/mms.12776

RStudio Team. (2015). *RStudio: Integrated Development Environment for R*. Boston, MA. Retrieved from http://www.rstudio.com/

Santos RS, Hawkins S, Monteiro LR, Alves M, Isidro EJ (1995) Case studies and reviews. Marine research, resources and conservation in the Azores. Aquat Conserv 5:311–354. https://doi.org/10.1002/aqc.3270050406

Santos, R. S., Hawkins, S., Monteiro, L. R., Alves, M., & Isidro, E. J. (1995). Marine research, resources and conservation in the Azores. *Aquatic Conservation: Marine and Freshwater Ecosystems*, *5*(4), 311-354. https://doi.org/10.1002/aqc.3270050406

Scheiber, H. N. (2000). Historical memory, cultural claims, and environmental ethics: the jurisprudence of whaling regulation. In *Law of the Sea* (pp. 127-166). Brill Nijhoff. (PDF) Historical memory, cultural claims, and environmental ethics in the jurisprudence of whaling regulation | AGWANDA CHRISTOPHER - Academia.edu

Shabangu, F. W., & Andrew, R. K. (2020). Clicking throughout the year: sperm whale clicks in relation to environmental conditions off the west coast of South Africa. *Endangered Species Research*, *43*, 475-494. https://doi.org/10.3354/esr01089

Silva, L. (2015). How ecotourism works at the community-level: the case of whale-watching in the Azores. *Current Issues in Tourism*, *18*(3), 196-211. https://doi.org/10.1080/13683500.2013.786027

Silva, M. A., Prieto, R., Cascão, I., Seabra, M. I., Machete, M., Baumgartner, M. F., & Santos, R. S. (2014). Spatial and temporal distribution of cetaceans in the mid-Atlantic waters around the Azores. *Marine Biology Research*, *10*(2), 123-137. https://doi.org/10.1080/17451000.2013.793814

SREA (Serviço Regional de Estatística dos Açores). (2007). Estudo sobre os turistas que visitam os Açores 2005_2006.

Steiner, L., Lamoni, L., Plata, M. A., Jensen, S. K., Lettevall, E., & Gordon, J. (2012). A link between male sperm whales, Physeter macrocephalus, of the Azores and Norway. *Journal of the Marine Biological Association of the United Kingdom*, *92*(8), 1751-1756. http://doi.org/10.1017/S0025315412000793

Steiner, L., Pérez, M., der Van Linde, M., Freitas, L., Santos, R., Martins, V., & Gordon, J. (2015, December). Long distance movements of female/immature sperm whales in the North Atlantic. In *Poster presented at the Biennial Society for Marine Mammalogy Conference, San Francisco, USA*. (20) (PDF) Long distance movements of female/immature sperm whales in the North Atlantic (researchgate.net)

Taylor, B. L., Perrin, W. F., Reeves, R. R., Rosel, P. E., Wang, J. Y., Cipriano, F., ... & Brownell Jr, R. L. (2017). Why we should develop guidelines and quantitative standards for using genetic data to delimit subspecies for data-poor organisms like cetaceans. *Marine Mammal Science*, *33*(S1), 12-26. https://doi.org/10.1111/mms.12413

Taylor, B.L., Baird, R., Barlow, J., Dawson, S.M., Ford, J., Mead, J.G., Notarbartolo di Sciara, G., Wade, P. & Pitman, R.L. (2019). Physeter macrocephalus (amended version of 2008 assessment). *The IUCN Red List of Threatened Species 2019*: e.T41755A160983555. Accessed on 23/02/2022. http://dx.doi.org/10.2305/IUCN.UK.2008.RLTS.T41755A160983555.en

Teloni, V., Zimmer, W. M., Wahlberg, M. and MADSEN, P. T. (2007). 127 Consistent acoustic size estimation of sperm whales using clicks recorded from unknown aspects. *J. Cetacean Res. Manage*, *9*(2), 127-136. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.463.1996&rep=rep1&type=pdf

Thode, A., Mellinger, D. K., Stienessen, S., Martinez, A., & Mullin, K. (2002). Depth-dependent acoustic features of diving sperm whales (Physeter macrocephalus) in the Gulf of Mexico. *The Journal of the Acoustical Society of America*, *112*(1), 308-321. https://doi.org/10.1121/1.1482077

Thode, A., Skinner, J., Scott, P., Roswell, J., Straley, J., & Folkert, K. (2010). Tracking sperm whales with a towed acoustic vector sensor. The Journal of the Acoustical Society of America, 128(5), 2681-2694. https://doi.org/10.1121/1.3495945

van der Linde, M. L., & Eriksson, I. K. (2020). An assessment of sperm whale occurrence and social structure off São Miguel Island, Azores using fluke and dorsal identification photographs. *Marine Mammal Science*, *36*(1), 47-65. http://doi.org/10.1111/mms.12617

Vieira, N., & Brito, C. (2009). *Past and recent sperm whales sightings in the Azores based on catches and whale watching information*. Journal of the Marine Biological Association of the United Kingdom, 89(5), 1067-1070. http://doi.org/10.1017/S0025315409000873

Vinding, K., Bester, M., Kirkman, S. P., Chivell, W., & Elwen, S. H. (2015). The use of data from a platform of opportunity (whale watching) to study coastal cetaceans on the southwest coast of South Africa. *Tourism in Marine Environments*, *11*(1), 33-54. http://dx.doi.org/10.3727/154427315X14398263718439

Wahlberg, M. (2002). The acoustic behaviour of diving sperm whales observed with a hydrophone array. *Journal of Experimental Marine Biology and Ecology*, *281*(1-2), 53-62. https://doi.org/10.1016/S0022-0981(02)00411-2

Watwood, S. L., Miller, P. J., Johnson, M., Madsen, P. T., & Tyack, P. L. (2006). Deep-diving foraging behaviour of sperm whales (Physeter macrocephalus). *Journal of Animal Ecology*, *75*(3), 814-825. https://doi.org/10.1111/j.1365-2656.2006.01101.x

Watwood, S. L., Patrick J. O. Miller, Johnson, M., Madsen, P. T., & Tyack, P. L. (2006). Deep-Diving Foraging Behaviour of Sperm Whales (Physeter macrocephalus). *Journal of Animal Ecology*, *75*(3), 814–825. http://www.jstor.org/stable/3838392

Web site: https://audacityteam.org/. It is free software distributed under the terms of the GNU General Public License. The name Audacity® is a registered trademark.

Weilgart, L. and Whitehead, H. (1988). Distinctive vocalizations from mature male sperm whales (Physeter macrocephalus). Can. J. Zool. 66,1931-1937. http://dx.doi.org/10.1139/z88-282

Werth, A. J. (2004). Functional morphology of the sperm whale (Physeter macrocephalus) tongue, with reference to suction feeding. *Aquatic Mammals*, *30*(3), 405-418. DOI 10.1578/AM.30.3.2004.405

Whitehead, H. (2003). SPERM WHALES. Social Evolution in the Ocean. *University of Chicago press*. ISBN: 9780226895185

Whitehead, H. (2017). SPERM WHALE. In B. Wursig, J.G.M. Thewissen, K. Kovacs, *Encyclopedia of Marine Mammals* (3rd edition, pp. 919-924). *Academic Press*. ISBN: 9780128043271

Whitehead, H., Dillon, M., Dufault, S., Weilgart, L., & Wright, J. (1998). Non-geographically based population structure of South Pacific sperm whales: dialects, fluke-markings and genetics. *Journal of Animal Ecology*, 253-262. https://www.jstor.org/stable/2647492

Zimmer, W. M. X., Tyack, P. L., Johnson, M. P., and Madsen, P. T. (2005b). "Three-dimensional beam pattern of regular sperm whale clicks confirms bent-horn hypothesis," J. Acoust. Soc. Am. **117**, 1473–1485. https://doi.org/10.1121/1.1828501

Zimmer, W. M., Madsen, P. T., Teloni, V., Johnson, M. P., & Tyack, P. L. (2005a). Off-axis effects on the multipulse structure of sperm whale usual clicks with implications for sound production. *The Journal of the Acoustical Society of America*, *118*(5), 3337-3345. https://doi.org/10.1121/1.2082707

Clapham, P. J., & Ivashchenko, Y. V. (2018). Whaling catch data are not reliable for analyses of body size shifts. Nature ecology & evolution, 2(5), 756-756. https://doi.org/10.1038/s41559-018-0534-2

Clements, C. F., Blanchard, J. L., Nash, K. L., Hindell, M. A., & Ozgul, A. (2017). Body size shifts and early warning signals precede the historic collapse of whale stocks. Nature ecology & evolution, 1(7), 1-6. https://doi.org/10.1038/s41559-017-0188

Azevedo, J. M. N.; Fernández, M.; González García, L.; (2021) MONICET: long-term cetacean monitoring in the Azores based on whale watching observations (2009-2020). https://doi.org/10.14284/546

Pace, D. S., Lanfredi, C., Airoldi, S., Giacomini, G., Silvestri, M., Pavan, G., & Ardizzone, D. (2021). Trumpet sounds emitted by male sperm whales in the Mediterranean Sea. Scientific reports, 11(1), 1-16. https://doi.org/10.1038/s41598-021-84126-8

Appendix

Recommendations and suggestion for the acoustic protocol:

- Hydrophone with higher sensitivity and, more important, with a higher frequency response, that can cover the frequency range of sperm whales (2-40 kHz).
- Record at 96 kHz. 16- or 24-bit depths depend on the methods for the acoustic analyses. 24-bit is recommended, as software convert automatically to 16 bits if needed.
- Always use STEREO mode with the Zoom recorder.
- Always record in WAV extension.
- Hydrophone always at the maximum depth. If possible, longer cables for possible interaction of thermocline.
- Test with more sinkers in order to have always a straight cable.
- Choose good sea state (Beaufort and Douglas $\leq$ 4).
- Record from the area behind the animals – merge with photo-identification needs.
- Put the hydrophone always before the fluke-up, e.g., when the animal is preparing to dive.
- Check the level of the recorder: avoid clipping by lowering the volume.
- Record as much as possible till 15 min after fluke-up. Avoid recordings shorter than 5 min.
- Collect all the related data: sea state, number of boats, recordings settings, time of the recordings, date and GPS coordinates. Number of individuals recorded and present in the area can be of interest too.
- Take photo-ID of the fluke.
- Rename the recordings when at the base. Choose one kind of nomenclature that allows automatic ordering process from software (e.g., excel).
- Write all the information on a datasheet. Avoid losing information.
- Train the operators.
- Choose automatic analyses of the IPI (e.g., PAMGuard or others not discussed).
- Link other approaches for validation of the data.

Acknowledgments