Fall 12-20-2019

Detection of Whale Acoustic Signals in the Northern Gulf of Mexico LADC-GEMM Database

Yingxue Gao
University of New Orleans, New Orleans, ygao3@uno.edu

Detection of Whale Acoustic Signals in the Northern Gulf of Mexico LADC-GEMM Database

A Thesis

Submitted to the Graduate Faculty of the
University of New Orleans
in partial fulfillment of the
requirements for the degree of

Master of Science
in
Applied Physics

by

Yingxue Gao

M.S. Southwest University of Science and Technology of China, 2011
B.S. Southwest University of Science and Technology of China, 2008

December 2019

我还在寻找那只会唱歌的鲸.....

Table of Contents

List of Figures

List of Tables

Abstract

Low-pass Fourier filter, wavelet filter, as well as matched filter detection methods were used to detect baleen whale signals in northern Gulf of Mexico data collected by the Littoral Acoustic Demonstration Center (LADC) consortium. Some potential low frequency signals appeared on the matched filter output figure. The shape of the signals is in line with one of the typical signal shapes of fin whales--vertical down-sweeps with 18s-time interval. Another shape of the signals is in line with one of the call type shapes of Bryde's whales--down-sweeps with 7s-time interval. A high-pass Fourier filter was also used to find toothed whale high frequency sounds in the Gulf of Mexico data. The sounds featuring click trains and codas belonging to sperm whales have been clearly identified.

Keywords: LADC, Gulf of Mexico, Signal, Detection, Baleen whale, Toothed whale, Frequency, Filter

Chapter 1.Introduction

1.1 Gulf of Mexico

1.1.1 Geographic and geological characteristics

The Gulf of Mexico is a coastal area of southeast North America, partly surrounded by land. It is connected to the Atlantic by the Florida straits between the Florida peninsula and the island of Cuba, and to the Caribbean Sea by the Yucatan strait between Cuba and the Yucatan peninsula. Both straits are about 160 kilometers wide. The longest east-west and north-south distances of the Gulf of Mexico are about 1,800 kilometers and 1,300 kilometers, respectively, with a total area of about 1.55 million square kilometers. Its northwest, north and northeast are the southern coast of the United States, and its west, south and southeast are the eastern coast of Mexico.

Figure 1.1 Satellite map of Gulf of Mexico(Google, 2019).

Figure 1.2 shows a map of the underwater topography at the bottom of the Gulf of Mexico. Shallow and intertidal areas (< 20 m deep) account for about 38% of the Gulf, continental shelves (20 to 180 m) 22%, continental slopes (180 to 3000 m) 20%, and deep-sea areas (> 3000 m) the remaining 20% (Love et al., 2013).

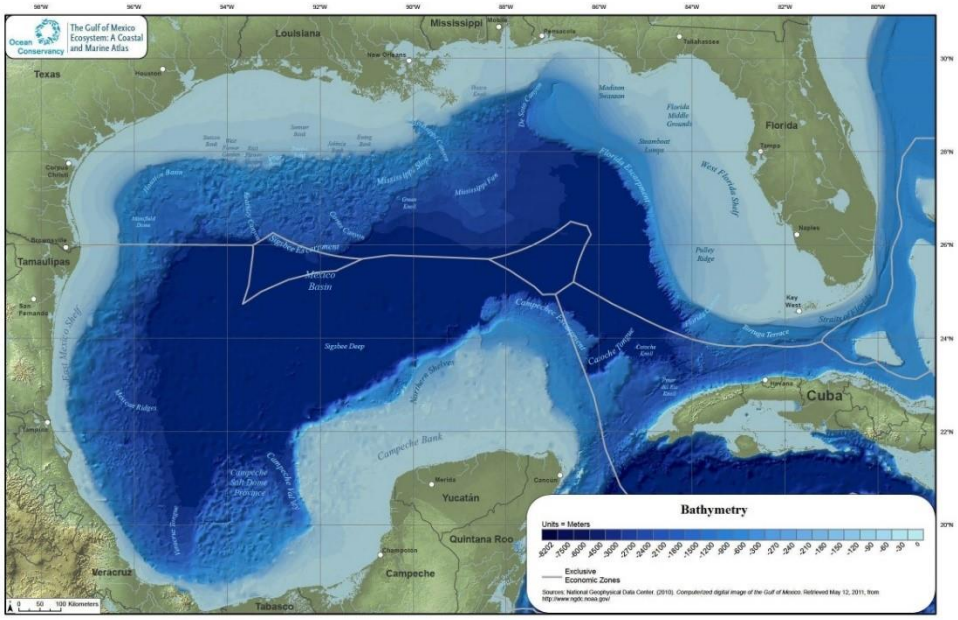


Figure 1.2 The underwater topography map of the Gulf of Mexico(Love et al., 2013).

Figure 1.3 shows the bottom sediments of Gulf of Mexico. In the deep waters of the central Gulf of Mexico and along the coasts of Texas, Louisiana and Mississippi, the bottom sediments are mostly mud (Balsam and Beeson, 2003), and those off Louisiana's barrier islands are mostly sand. Gravel and rock matrix are found on the outer edge of the southern continental shelf in Texas, Louisiana, Mississippi, Alabama and Florida (Love et al., 2013).

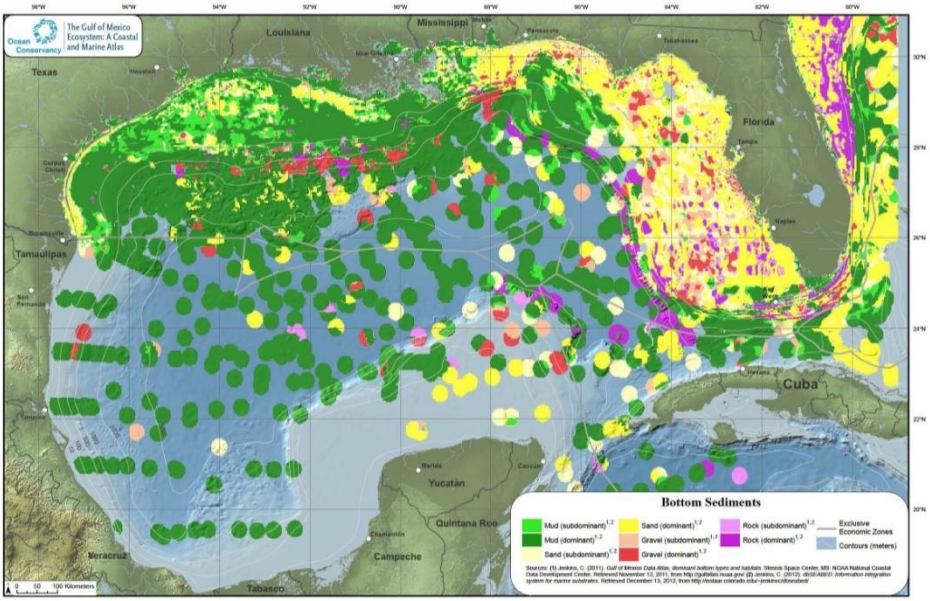


Figure 1.3 Bottom sediments map of Gulf of Mexico(Love et al., 2013).

1.1.2 Habitats and species

The Gulf of Mexico coast includes more than 5 million acres of wetlands, which provides important habitat for both native and migrating birds. Louisiana's wetlands attract migratory birds that fly over the Gulf of Mexico, and in the fall, many others, including the common loon from the Great Lakes region, migrate to the winter waters of the Gulf of Mexico, where they will stay until the spring of the next year.

Estuaries, the transitional zones between freshwater and marine environments, are habitats for shrimp, crabs, oyster and fish species. Coral, seagrasses and pelagic sargassum throughout the bay are also important habitats for fish, invertebrates and sea turtles (NOAA, 2019a).

Below the seabed, salt structures also form unique habitats that support chemo-synthetic communities and reef fish such as red snapper. Seafloor sediments provide habitats for a variety of organisms, mainly meiofaunal communities including nematodes, protozoans and diatoms (NOAA, 2019a).

In addition to the above species, one of the most fascinating marine species - marine mammals - can be found in the Gulf of Mexico. There are nearly 32 potential and known gulf marine mammal species described (Würsig, 2017), some of which are often found in the northern Gulf, while others are very rare. Table 1.1 lists possible species and other relevant information.

Table 1.1 Potential Marine Mammal Species in Gulf of Mexico(Würsig, 2017)

Species	**Sighting**	**General depth**
North Atlantic right whale	Rare sightings	
Blue whale		
Fin whale	Rare sightings	
Sei whale		
Humpback whale	Rare sightings	
Minke whale	Rare sightings	
Bryde’s whale	Quite common sightings	Upper slope
Sperm whale	Common sightings	Slope and deep ocean
Pygmy sperm whale	Common sightings	Slope and deep ocean
Dwarf sperm whale	Common sightings	Slope and deep ocean
Cuvier’s beaked whale	Occasional sightings	Deep ocean
Blainville’s whale	Occasional sightings	
Sowerby’s beaked whale		

Gervais' beaked whale	Occasional sightings	
Killer whale	Common sightings	Deep ocean
Short-finned pilot whale	Common sightings	Slope to deep ocean
Long-finned pilot whale		
False killer whale	Medium common sightings	Upper slope to deep ocean
Pysgmy killer whale	Medium common sightings	Deep ocean
Melon-headed whale	Common sightings	Deep ocean
Rough-toothed dolphin	Common sightings	Upper slope to deep ocean
Risso's dolphin	Common sightings	Upper slope to deep ocean
Common bottlenose dolphin	Common sightings	Upper slope and shallower waters
Pantropical spotted dolphin	Common sightings	Upper slope to deep ocean
Atlantic spotted dolphin	Common sightings	Upper slope and shallower waters
Spinner dolphin	Common sightings	Deep ocean to upper slope
Clymene dolphin	Common sightings	Deep ocean and slope
Striped dolphin	Common sightings	Deep ocean and slope
Short-beaked common dolphin		
Long-beaked common dolphin		
Fraser's dolphin	Occasional sightings	Deep ocean and slope
West Indian manatee	Common sightings	

Table 1.1 shows dolphins are often found in waters that are less than 200 meters deep on the shore/coast and continental shelf, while toothed whales mainly occur on continental slopes and deep waters. The Gulf of Mexico does not have only species that are unique to the Gulf. All species are found in other oceans except for the Caribbean Atlantic cows and Clymene and Atlantic spotted dolphins that are unique to the North Atlantic (Burger et al., 2017).

Efforts have been made to classify and collect the species and populations of whales in the Gulf of Mexico, but in fact it is a difficult task for any part of the ocean in the world. The population dispersion and quantity estimate of cetaceans produce different results in different years and even different seasons. Although sound censuses and visual censuses are popular(Davis et al., 2002), there are not enough sound data measurements to describe accurately the approximate number of marine animals except sperm whales.

As the earth and ocean environment change, factors such as climate, water temperature, salinity, ocean currents, noise, etc., the description and statistical information of

whale migration, especially for the baleen whales with migratory habits may change. This means that the observation (sighting) probability of some species of whales may be different compared with Table 1.1.

1.2 Significance of research and listening to marine mammals

The investigation and study of various whale populations around the world will help humans to have a better understanding of the distribution and living habits of this species, and thus help humans to adjust their behaviors in the natural ecology to help prevent damage to the species and thus prevent the extinction of those endangered species. Marine mammals are sensitive to changes in the environment, so they often act as "ecosystem sentinels" around the world. Therefore, the study of marine mammals helps to protect the balance of the ecological environment.

Most marine mammals live in the deep ocean, which makes it difficult for them to be within human visual range, so it is necessary to use different methods to gain a deeper knowledge of marine mammals in addition to observing them visually. Acoustic detection is a method that is profitable for tracking marine mammals without visual tracking.

It is important to monitor the sounds of marine mammal in any ocean area because they do everything by sound, such as communicating with each other and exploring their surroundings. The acoustic data collected includes the characteristic sounds of various whales as well as information about the marine environment in which they live. Computer programs can decode this information and match changes in marine mammal populations with environmental factors (LADC-GEMM, 2017).

1.3 Passive Acoustic Methods

Visual observational research is expensive from a time and money perspective because it involves training the observer and spending a lot of time driving expensive ships or planes. In addition, visual surveys are limited by daylight and weather conditions. The passive acoustic method involves fixing automatic recording instruments on the ocean floor and leaving them there to record cetacean vocalizations (Mellinger et al., 2007). Passive acoustic methods are popular because they have lower costs relative to other methods of observational research and are capable of collecting data for long periods of time during any day or night, in any weather

or ocean conditions, and from any location. The temporal and spatial distribution of sound recorded by passive acoustic monitoring instruments helps to study the distribution, movement, behavior, relative abundance and population structure of whales (Mellinger et al., 2007).

In the Gulf of Mexico, the LACD-GEMM project is using several techniques to collect acoustic data on marine mammals in particular. These devices include Automatic Underwater Vehicle-gliders (AUVs), the Automatic Surface Vehicle (ASVs) and the Environmental acoustic recording system (EARS) buoys(LADC-GEMM, 2017).

Figure 1.4 shows the arrangement of the various data collection devices in the Gulf of Mexico by the LADC-GEMM project.

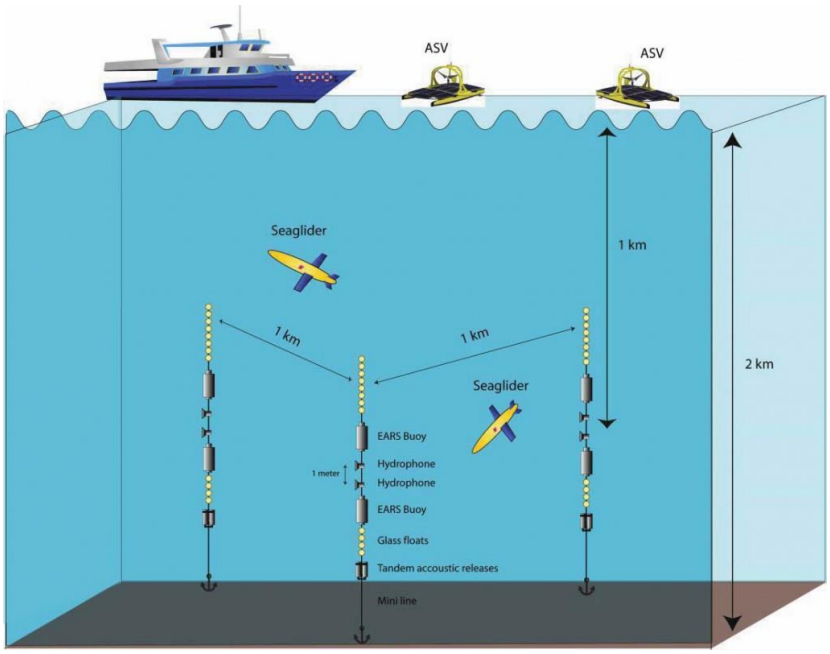


Figure 1.4 EARS, ASVs, Gliders arrangement in Gulf of Mexico(LADC-GEMM, 2017).

As shown in Figure 1.4, passive acoustic monitoring in an aquatic environment refers to the use of underwater microphones (hydrophones) to monitor and locate vocal marine mammals. Passive means simply listening without introducing noise into the aquatic environment, so this method should not be confused with active acoustic systems where a sound is projected into the medium (OSC, 2019).

1.4 Research purpose

Many toothed whale species live in the Gulf of Mexico for a long time as shown in Table 1.1, the most common of which is the Gulf of Mexico independent population of sperm whales. In this thesis, a one-minute audio sample is chosen from a mass of audio recordings containing toothed whales to try to find the sound signal from the sperm whale and recognize its sound signal characteristics and unique signal structure in the time and frequency domains.

Apart from toothed whales, baleen whales have also been reported in the Gulf of Mexico. This thesis will focus on locating sound records of fin whales and Bryde's whales in the LADC-GEMM database from 2001 to 2017 using an appropriate computer algorithm, and to further distinguish the sound characteristics of these baleen whales in the Gulf of Mexico from those in other oceans.

1.5 Biological background

This section contains a brief physical description and vocalization characteristics of several whales of interest in this thesis.

- Sperm whale

The sperm whale is the largest toothed creature on earth. The average male is 15 meters long and weighs 36,000 kilograms, while the average female is 11 meters long and weighs 20,000 kilograms. Juvenile' heads make up about a fifth to a quarter of their body length, while older adult males' heads make up about a third of their body length. The head has a large, waxy, oil-like structure called the spermatozoa(Burger et al., 2017).

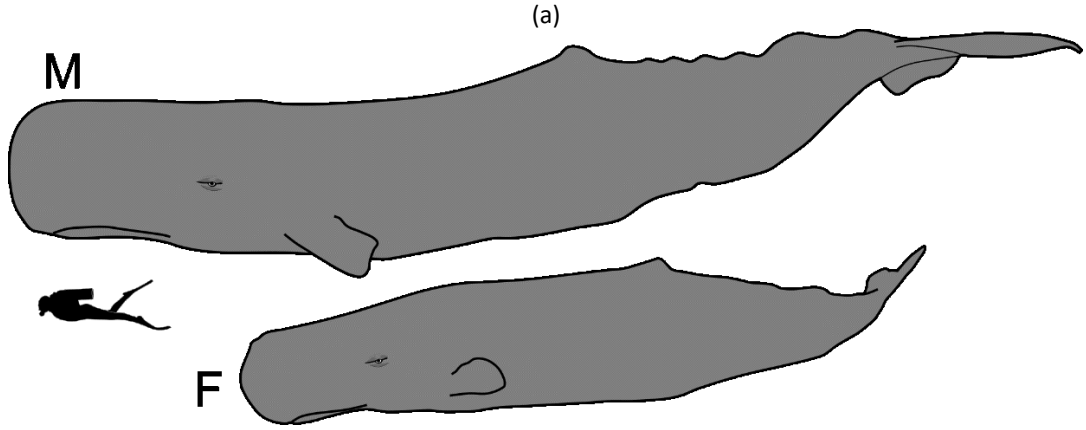
Sperm whales live in all the oceans of the world and usually feed in waters below 500 meters. The sperm whales in the Gulf of Mexico are mainly female, so they are on average 1.5-2.0 meters smaller than the average of all sperm whales elsewhere (Richter et al., 2008)(Jaquet and Gendron, 2009). Males travel and reproduce in the Gulf of Mexico and the North Atlantic as well as other oceans.

Sperm whales use echolocation in their directional clicks, which range in frequency from less than 100 hertz to 30 kilohertz. Most of these clicks have energy in the range of 5 to 25 kilohertz. Sperm whales also produce repeated click-through patterns in social situations called codas, which may be used to attract females, compete for mates, show aggression and

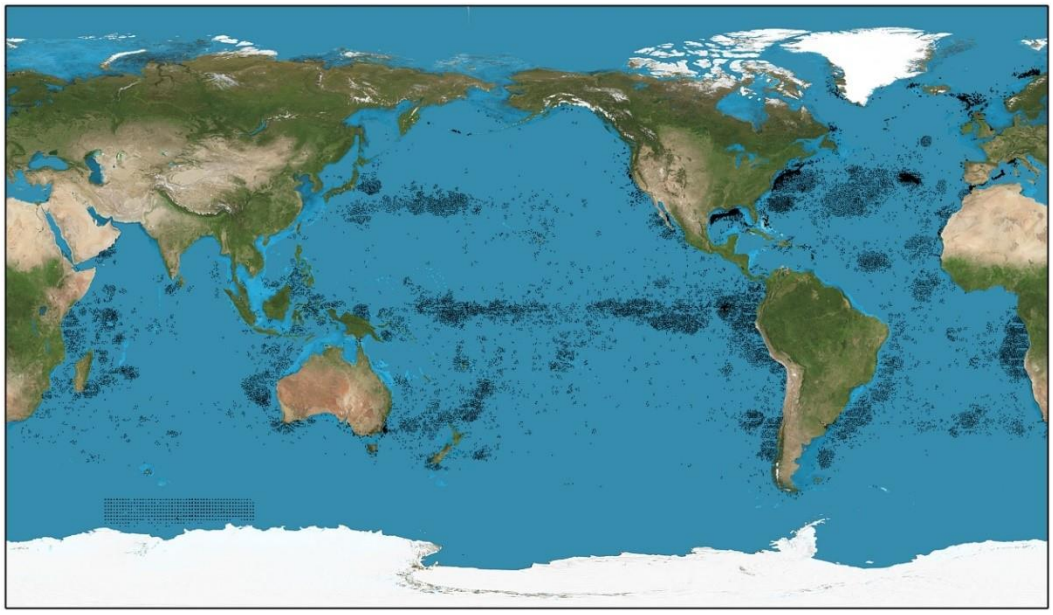
maintain group cohesion. During foraging, sperm whales make regular, spaced clicks, accompanied by very fast clicks in the final stages of capturing their prey (Wahlberg, 2002).

The following are pictures showing the appearance, whale size compared with human size, and world-wide distribution of sperm whales.

(a)



(b)



(c)

Figure 1.5 (a) Mother sperm whale and her baby; (b) Size comparison of a male and female sperm whale with a human (16m, 11m, 1.75m respectively); (c) Map of the distribution of sightings of sperm whales(Wikipedia contributors, 2019a)

- Fin whale

Fin whales belong to the parvorder of baleen whales, which are the second largest species on earth after blue whales. The maximum measured length is 25.9 meters and the maximum recorded weight is close to 74 tons (Lockyer, 1976). Fin whales exist in all major oceans from polar to tropical waters, but the areas with the highest population density are temperate and cool waters (National Marine Fisheries Service, 2010).

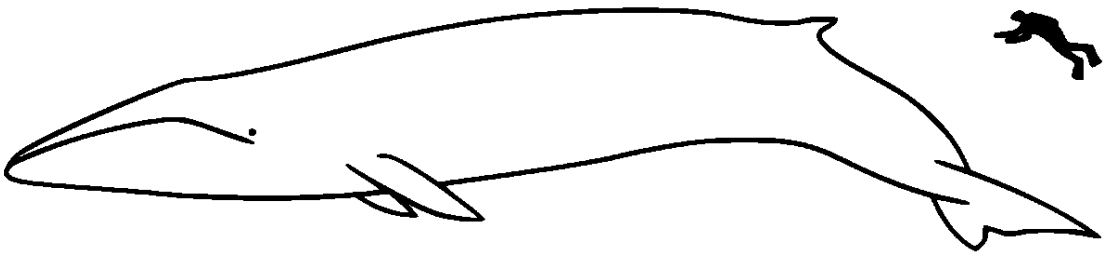
Male Fin whales produce long, loud, low-frequency sounds. Most of the sound is a down-swept frequency modulated (FM) pulse, between 16 and 40Hz, centered at approximately 20Hz. Down-swept frequencies above 50Hz and even up-swept higher than 100Hz have occurred but are not common or only appear in specific sea areas (Castellote et al., 2012). Each sound pulse lasts one to two seconds, and different combinations of sound pulses form sequences that last from a few to a dozen minutes each. Repeated sequences form the whale song, which usually lasts two hours or even several days. Usually the term “note” is used for the 20Hz pulses that compose a song. When 20Hz pulses are in irregular patterns, they are referred to as a “call” (Weirathmueller et al., 2017).

The time interval between pulses of fin whales varies in different seas and even in the same ocean area but in different seasons from 6 seconds to 46 seconds. Different types of fin whale songs are formed according to different pulse intervals. There are two main types of fin whale songs recorded. One is a singlet song that consists of a single, repeated note sequence with a fixed pulse interval and frequency; the other is a doublet song that has two primary notes, alternating between low frequency, high pulse interval notes and higher frequency, low pulse interval notes (Weirathmueller et al., 2017).

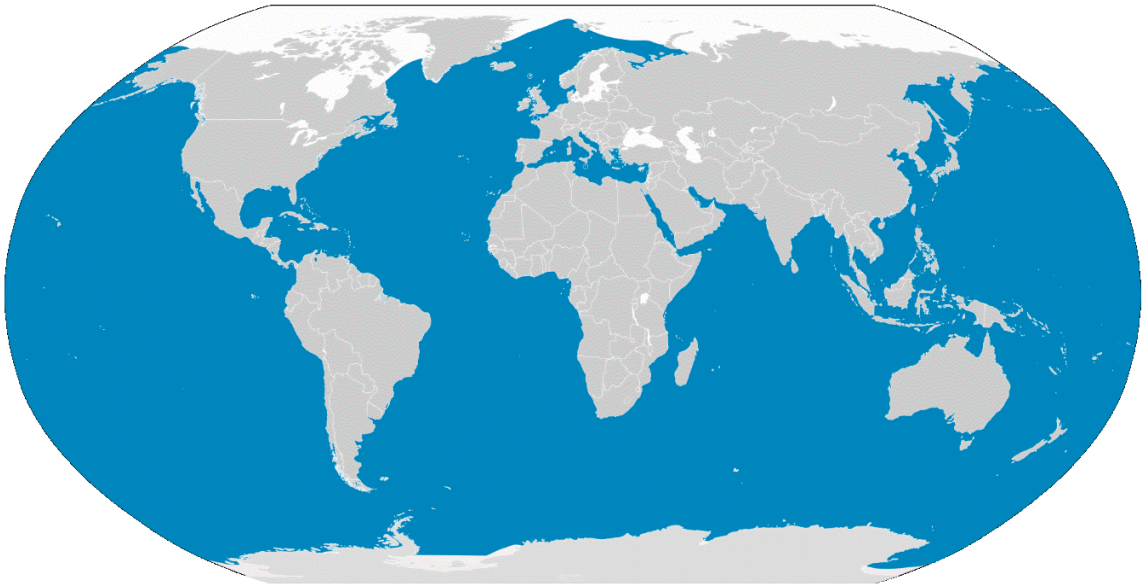
These vocalizations are directly related to the breeding season of the species. Only male vocalizations have been recorded so far (Croll et al., 2002), so these vocalizations may be a sign of mating calls.

The following are pictures showing the appearance, whale size compared with human size, and world-wide distribution of fin whales.

(a)



(b)



(c)

Figure 1.6 Fin whale (a) appearance; (b) Size comparison with a human; (c) distribution range(Wikipedia contributors, 2019b)

- Bryde's whale

An adult Bryde's whale is about 13 meters long and weighs about 26,455 pounds, and the average length of a calf is 7 meters. Females tend to be slightly larger than males. The Bryde's whales are also known as tropical whales who are found mainly in tropical, subtropical and warm temperate waters (61° to 72°F) around the world, from 40 degrees south latitude to 40 degrees north latitude. They are not highly gregarious(Burger et al., 2017).

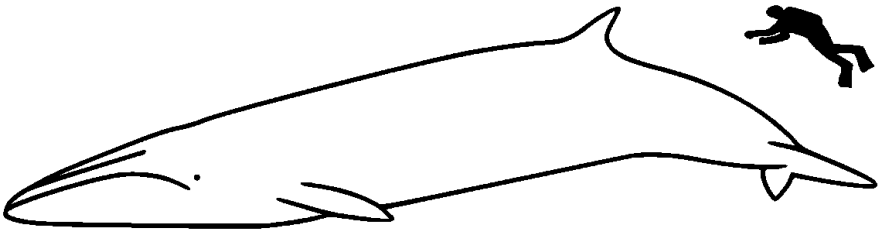
Gulf of Mexico Bryde's whales are members of the baleen whale family and a subspecies of the Bryde's whale. They live in the northeastern Gulf of Mexico, along continental shelf fractures 100 to 400 meters deep. The Gulf of Mexico Bryde's Whale is one of the few non-migrating baleen whales that live in the Gulf of Mexico throughout the year. Bryde's whales are vulnerable to stress and threats, such as ocean noise and whaling. Subspecies in the Gulf of Mexico are threatened by oil and gas activity, as well as oil spills and cleanup. Scientists believe there are fewer than 100 of them in the Gulf of Mexico, with 50 or fewer being mature individuals(NOAA, 2019b).

Bryde's whale can generate powerful sounds containing low frequencies. The calls/sounds are mainly divided into four types: down-sweep, tonal sequence, long-moan and pulse. The call/sound duration range is between 10μs and 7s, and the frequency is usually

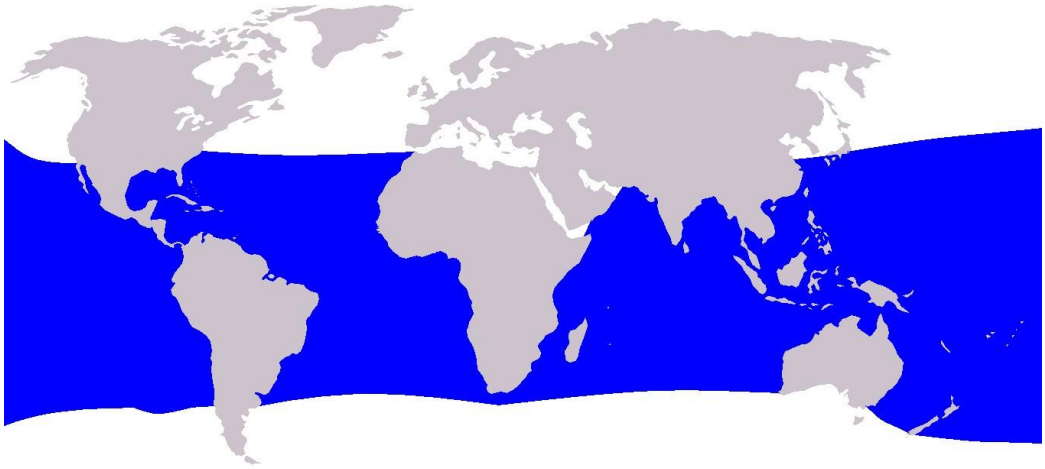
below 900 Hz. The inter-call/phase interval varies from 6.25ms to 40mins(Rice et al., 2014a). Most of the calls/sounds made by Bryde's whales include several types of calls simultaneously. They repeat these calls every few minutes, many of which are produced as the whales move, and the type of call varies with the size of the group (Heimlich et al., 2005).

The following are pictures showing the appearance, whale size compared with human size, and world-wide as well as Gulf of Mexico distribution of Bryde's whales.

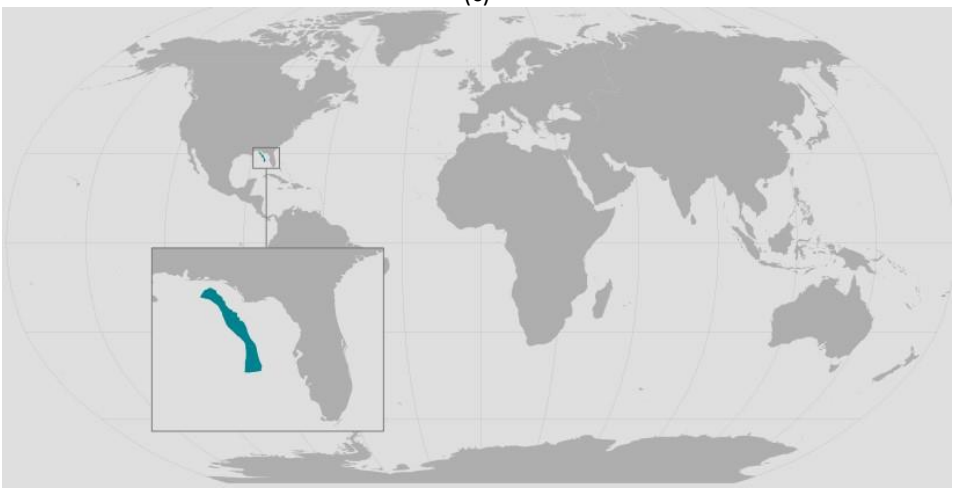
(a)



(b)



(c)



(d)

Figure 1.7 Bryde's whale (a) appearance showing faint lateral ridges; (b) Size comparison with a human;(NOAA, 2019b)*; (c)living area all over the world; (d)distribution of Gulf of Mexico Bryde's whales*(LaBrecque et al., 2015)

1.6 Ocean noise

Sound travels five times faster through water than through air, and many times farther than through air. Marine mammals use acoustic media to find food, mates and predators. However, undersea noise pollution is destroying the environment for marine life. Noise sounds obscure the whale's sense of hearing and being heard and disrupt critical behavior and communication between whales.

The primary man-made sources of ocean noise are commercial, industrial and military. Cavitation noise produced by a propeller and the engine itself make huge rumble; the petrochemical industry uses seismic air guns to emit extremely loud pulses to the seabed; military sonar emits sound frequencies ranging from very low (infrasound) to very high

(ultrasonic) to detect a target (Noise, 2003). Sounds produced by human activities including communications, navigation, defense, research, exploration and fishing are also part of ocean noise. During seismic exploration, acoustic explosions can last for days or weeks. Human activities generate a wide range of noise frequencies, ranging from a few hertz to hundreds of kilohertz (Studds and Wright, 2007).

Some physical processes including rainfall, sea ice rupture, undersea earthquakes and undersea volcanic eruptions also contribute to ocean noise. Rainstorms, for example, can raise the noise level of bubbles and sprays by more than thirty decibels underwater over frequencies from a few hundred hertz to 20,000 hertz (Andrew et al., 2002).

Figure 1.8 summarizes the background sounds in the ocean, showing typical sound levels at different frequencies.

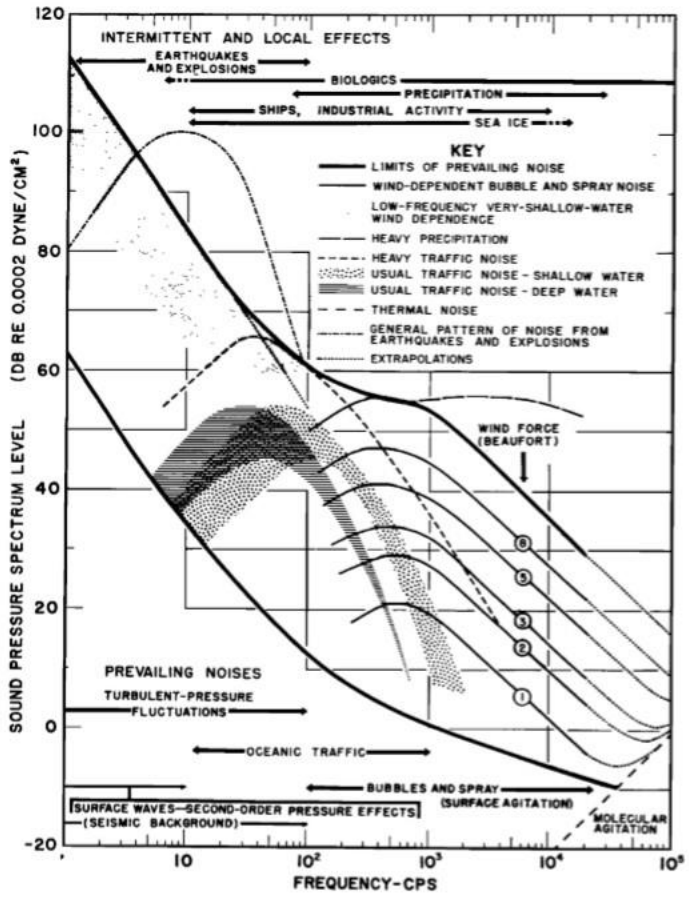


Figure 1.8 The typical sound levels of ocean background noises at different frequencies (Wenz, 1962).

Figure 1.8 is referred to as the Wenz curves. In the range of 20-500 Hz, environmental noise is mainly caused by the noise generated by long-distance transportation. At frequencies

ranging from 500 to 100,000 Hz, the bubbles associated with the spray and the breaking waves cause the main environmental noise, and the noise increases as the wind speed increases. At frequencies over 100,000 Hz, thermal noise from the random movement of water molecules dominates (Wenz, 1962).

Most toothed whales produce sounds with frequencies above the noise frequency or have a higher sound power than the noise power at that frequency, which makes the sound of the toothed whale relatively easy to detect. But for baleen whales, which produce very low-frequency sounds of low power, the sounds are just mixed in with the noise, which makes it harder to detect the sound of baleen whales. This is the primary difficulty of finding baleen whale signal in the Gulf of Mexico measured data used in this thesis.

Chapter 2. Data Background

A consortium of scientists from universities, government laboratories, and scientific companies established the Littoral Acoustic Demonstration Center (LADC)- Gulf Ecological Monitoring and Modeling (GEMM) in 2001 to study shallow water ambient noise, propagation and marine mammal acoustics using Environmental Acoustic Recording System (EARS) buoys in the northern Gulf of Mexico.

When LADC was formed, the buoys could measure signals with frequencies up to 1000Hz. When LADC added listening to the sounds of sperm whales to its noise and propagation measurement mission, the buoys were modified, and the measurement frequency raised to 5,859Hz within 36 days. These buoys were moored in the depths of the northern Gulf of Mexico from 550 to 950 meters, producing clear recordings of sperm whale echolocation and coda clicks and other whale recordings. EARS Generation 2 buoys are now able to record one channel to 96kHz, or 4 channels to 25kHz, over 13 days on four 120 GB laptop disk drives (Ioup et al., 2016).

Beginning with the 2001 experiment, LADC scientists studied the click and click behavior of sperm whales. In 2007, the study extended to the clicks of beaked whales. By 2019, baleen whales with very low-frequency sounds, such as Fin whales and Bryde's whales, are considered for study.

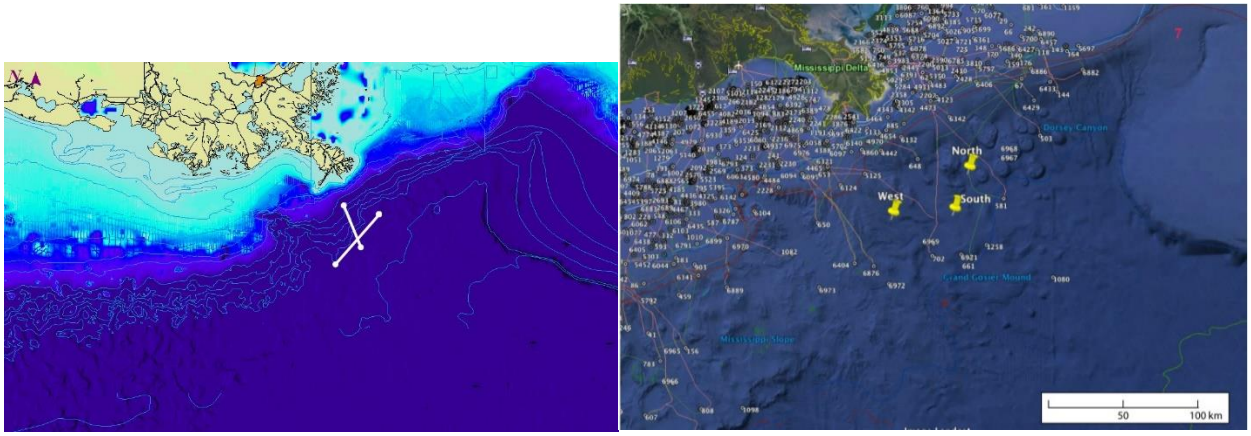
There were some marine mammal acoustic experiments performing by LADC in the northern Gulf of Mexico from 2001 to 2017, and all experiments used EARS buoys. The following table shows the parameters of part of the experiments:

Table 2.1 The parameters of LADC experiments

Buoys	*Time*	*Location in Gulf of Mexico*	*Depth(m)*	*Above bottom(m)*	*Sampling rate(kHz)*	*Other events*
LADC 01	7/16/2001-8/21/2001	Northeastern	600,800, 1000	50	11.7	Tropical Storm Barry passed within 100 nmi
LADC 02	8/19/2001-9/15/2001	North central	600,800, 1000	50	11.7	Tropical Storm Isidore passed with 73 nmi; Hurricane Lili passed within 116 nmi

LADC 07	7/3/2007-7/14/2007	South of Atchafalaya Bay	750-800		192	
	2015,2017	North	1000-2000	289(300m mooring) 537(500m mooring)	192	

Figure 2.1(a) shows the location of the LADC 01 and LADC 02 experiments. The white cross lines represent the position of LADC 01 and LADC 02 experiments. The three yellow pins in Figure 2.1(b) show the location of 2017 experiments sites. They were located between the shallow sea and the continental slopes in the northern Gulf of Mexico, and the bottom of the water is dominated by muddy sediments.



(a) (b)

Figure 2.1 Location of (a)2001 LADC 01 and LADC 02 experiments(Ioup et al., 2016)*; (b)2017 experiments sites*(Sidorovskaia et al., 2017)

Chapter 3. Methods and Techniques

Much mathematical knowledge and many algorithms are used in signal processing. This section of this chapter will theoretically elaborate and discuss the methods of calculation and problem solving used in this thesis.

3.1 Fourier Transform

The Fourier transform is a linear integral transform used to transform a signal between the time domain (or space domain) and the frequency domain. It has many applications in physics and engineering. In the time domain, the intensity of signal distribution over time is presented, while in the frequency domain, the frequency distribution of the signal can be intuitively reflected.

The function $\mathrm{F(s)}$ generated by the Fourier transform, also known as the spectrum, is called the Fourier transform of the original function $\mathrm{f(x)}$. In many cases, the Fourier transform is invertible, that is, the original function f(x) can be obtained from $\mathrm{F(s)}$. Both f(x) and F(s) are complex functions, but often f(x) is a real function. $\mathrm{F(s)}$ is always a complex function, which uses a complex number to represent amplitude and phase or real and imaginary parts.

The Fourier transform includes the Continuous Fourier Transform (CFT) and the Discrete Fourier Transform (DFT). The definition of the continuous Fourier transform in integral form(Bracewell, 2000) is,

$$F(s) = \int_{-\infty}^{\infty} f(x) e^{-i2\pi xs}\, dx \tag{3.1}$$

$$f(x) = \int_{-\infty}^{\infty} F(s) e^{i2\pi xs} ds \tag{3.2}$$

Equation (3.1) shows $\mathrm{F(s)}$, the forward Fourier transform of $\mathrm{f(x)}$. The independent variable x represents time (in seconds), and the transformation variable s represents frequency (in Hertz). Equation (3.2) shows that $\mathrm{f(x)}$ can be recovered from $\mathrm{F(s)}$ by the inverse Fourier transform.

Data processing on a digital computer uses the DFT, which is the discrete form of the Fourier Transform in both the time domain and frequency domain, transforming the time domain samples of signal into the frequency domain samples of the Discrete Time Fourier Transform (DTFT). Because digital systems can only handle discrete signals of finite length, both

x and s must be discretized and of finite length, and the corresponding Fourier transform must be established. For the N-point sequence $\{x_k\}=x_0, x_1, \ldots, x_{N-1}$, its DFT is (Brigham, 1988),

$$X_n = \sum_{k=0}^{N-1} x_k e^{\frac{-i2\pi kn}{N}}, \ \mathrm{n} = 0, 1, \ldots, \mathrm{N}-1 \tag{3.3}$$

$\{X_n\}=X_0, X_1, \ldots, X_{N-1}$. The inverse transform of the Discrete Fourier Transform (IDFT) is(Brigham, 1988),

$$x_k = \frac{1}{N}\sum_{n=0}^{N-1} X_n e^{i\frac{2\pi kn}{N}}, k = 0, 1, \ldots, \mathrm{N}-1 \tag{3.4}$$

In practical computing, the Fast Fourier Transform (FFT) replaces the Discrete Fourier Transform (DFT). The FFT is a method of rapidly computing the DFT or its inverse Transform of a sequence. The FFT quickly calculates such a transformation by decomposing the DFT matrix into a product of sparse (mostly zero) factors(Van Loan, 1992). Using FFT calculations will yield the same results as the DFT calculations, but it can reduce the complexity of DFT calculations from N^2, which is required for DFT definition calculations as shown in equation (3.3), to $N\backslash\log_2 N$(Higgins, 1976). Take $N=1024$ for example: the DFT will compute 1048596 times, while the FFT will compute 10240 times, that is, the FFT computes 2^7 times faster than the DFT. The FFT takes advantage of the symmetry of the Fourier Transform so there is no loss of accuracy.

3.2 Properties and theorems of the Fourier transform

In practical applications of the Fourier transform, its properties and theorems - both in the mathematical and graphical sense - can better explain changes in the time or frequency domains. The following are some properties and theorems of Fourier transforms which are widely used (Bracewell, 2000).

- Similarity theorem

  If $f(x)$ has the Fourier transform $F(s)$, then $f(ax)$ has the Fourier transform $|a|^{-1}F(s/a)$.

$$\int_{-\infty}^{\infty} f(ax)e^{-i2\pi xs}\,dx = \frac{1}{|a|}F(\frac{s}{a}) \tag{3.5}$$

This theorem shows that when the time scale of the left-side is compressed, it corresponds to an expansion of the frequency scale of the right-side. In addition, when one member of the transformation pair expands horizontally, the other element not only contracts horizontally but also grows vertically, so that the area underneath it remains unchanged.

- Shift theorem

If $f(x)$ has the Fourier transform $F(s)$, then $f(x-a)$ has the Fourier transform $e^{-i2\pi as}F(s)$.

$$\int_{-\infty}^{\infty} f(x-a)e^{-i2\pi xs}\,dx = e^{-i2\pi as}F(s) \tag{3.6}$$

This theorem indicates that if a given function moves an amount a in the positive direction in the function domain, the amplitude of the Fourier component does not change but the phase changes. According to this theorem, the phase delay of each component is proportional to s, that is, the higher the frequency, the greater the change in phase angle.

- Convolution theorem

The convolution of f and g is written as $f * g = \int_{-\infty}^{\infty} f(\tau)g(t-\tau)\,d\tau$. It is defined as the integral of the product of the two functions after one is reverse and shifted.

If $f(s)$ has the Fourier transform $F(s)$ and $g(x)$ has the Fourier transform $G(s)$, then $f(x)$ * $g(x)$ has the Fourier transform $F(s)G(s)$; that is, convolution of two functions means (complex) multiplication of their transforms.

$$\int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} f(x')g(x-x')\,dx'\right]e^{-i2\pi xs}\,dx = F(s)G(s) \tag{3.7}$$

This theorem and its inverse theorem play an important role in transforming a function, which can be thought of either as the convolution of two other functions, or the product of two other functions.

- Autocorrelation theorem

Given a signal $f(t)$, the autocorrelation is defined as the integral of $f(t)$ with itself, at lag τ. The autocorrelation is written as $f \star f = \int_{-\infty}^{\infty} f(t+\tau)\overline{f(t)}dt$, where $\overline{f(t)}$ represents the complex conjugate of $f(t)$. If $f(x)$ has the Fourier transform $F(s)$, then its autocorrelation function $\int_{-\infty}^{\infty} f^*(u)f(u+x)du$ has the Fourier transform $|F(s)|^2$.

$$\int_{-\infty}^{\infty}|F(s)|^2 e^{i2\pi xs}ds = \int_{-\infty}^{\infty} f^*(u)f(u+x)du \tag{3.8}$$

The autocorrelation theorem is a special case of the cross-correlation theorem, and the autocorrelation function of a signal is the Fourier transform of its power spectrum

For the DFT, the expressions of the theorems which are widely used are (Bracewell, 2000):

Shift theorem:

$$f(\tau - T) \supset e^{-i2\pi T\left(\frac{\nu}{N}\right)}F(\nu) \tag{3.9}$$

(Cyclic) Convolution theorem:

$$f_1(\tau) * f_2(\tau) \supset NF_1(\nu)F_2(\nu) \tag{3.10}$$

In the practical cyclic convolution computing, padding with zeros is necessary to wraparound and aliasing, possible in either domain.

Autocorrelation theorem:

$$\sum_{\tau'=0}^{N-1} f_1(\tau')f_2(\tau' + \tau) \supset N|F_1(\nu)|^2 \tag{3.11}$$

which is also the expression of the power (energy) spectrum.

Comparing the equations from (3.5) to (3.11), the arguments used in the CFT and DFT formulas are different, but the meaning of the transformation is the same.

3.3 Fourier filter

In measured physical data, the signal is often mixed with noise. There are many ways to remove noise from the signal, one of which is to filter out unwanted frequencies in the frequency domain.

Set $\mathrm{W(f)}$ as the filter factor, $\mathrm{f_1(t)}$ as the input and $\mathrm{f_2(t)}$ as the output functions in the time domain. Filtering the signal in the frequency domain is essentially analyzing $\mathrm{f_1(t)}$ into its spectrum, multiplying each spectrum component by the corresponding transfer function W to get the spectrum of $\mathrm{f_2(t)}$, and then synthesizing $\mathrm{f_2(t)}$ from its spectrum(Bracewell, 2000). Thus

$$F_2(\mathrm{f}) = \mathrm{W(f)}F_1(\mathrm{f}) \tag{3.12}$$

where $\mathrm{F_1(f)}$ and $\mathrm{F_2(f)}$ are the Fourier transform of $\mathrm{f_1(t)}$ and $\mathrm{f_2(t)}$, respectively. Then

$$f_2(t) = \int_{-\infty}^{\infty} W(f)F_1(f)e^{i2\pi ft}df \tag{3.13}$$

Equations (3.12) and (3.13) show the Fourier filter $\mathrm{W(f)}$ can first work in the frequency domain with $\mathrm{F_1(f)}$, and then the inverse Fourier transform can obtain the output function $\mathrm{f_2(t)}$ in the time domain (the filter of the original input function $\mathrm{f_1(t)}$). Figure 3.1 gives an example of how the Fourier filter works on data.

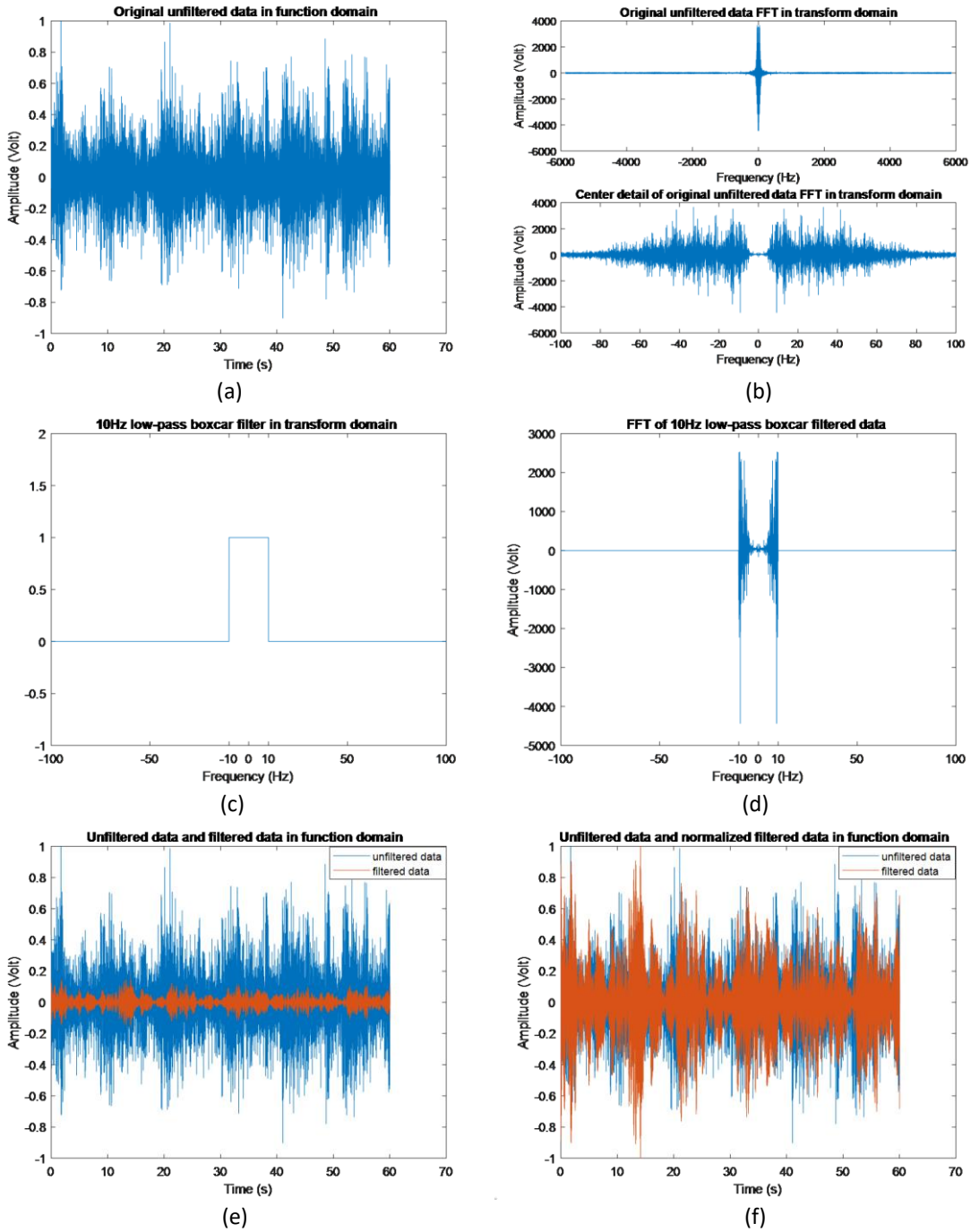


Figure 3.1 Fourier filter applied to measured data; (a) original unfiltered data; (b) FFT of original unfiltered data and detail; (c) 10Hz low-pass boxcar filter in transform domain; (d) FFT of filtered data; (e) comparison of unfiltered data and filtered data; (f) comparison of unfiltered data and normalized filtered data.

Figure 3.1(a) shows 60 seconds sound signal $f_1(t)$ in the time domain, and Figure 3.1(b) shows $F_1(f)$ which is the Fourier transform of $f_1(t)$ in the frequency domain. Here the filter factor $W(f)$ is a 10 Hertz low-pass rectangular (boxcar) function shown in Figure 3.1(c). Because of the boxcar filter, the signal with frequencies below magnitude 10 Hertz will be removed.

Figure 3.1(d) shows $\mathrm{F}_2(\mathrm{f})$ which is the result of filtered data where only signals below 10 Hertz were retained in the frequency domain. Figure 3.1(e) and Figure 3.1(f) shows the comparison between original unfiltered data $\mathrm{f}_1(\mathrm{t})$ and after applying the inverse Fourier transform (IFFT) to the filtered data ($\mathrm{f}_2(\mathrm{t})$ in time domain). As can be seen, it is obvious that many of the original signals were filtered out, all of which were above 10 Hertz.

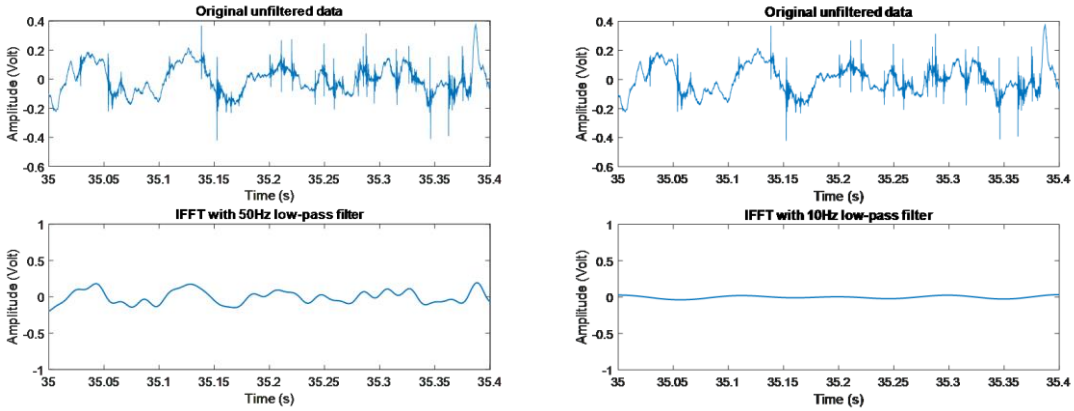
According to the convolution theorem mentioned in section 3.2, transformation multiplication corresponds to the convolution of the original function, so f_2 can be directly derived from f_1, that is,

$$f_2(t) = w(t) * f_1(t) \tag{3.14}$$

where $\mathrm{w}(\mathrm{t})$ is the Fourier transform of $\mathrm{W}(\mathrm{f})$. The two procedures are shown schematically as follows:

$$\begin{array}{ccc} \mathrm{f1(t)} & \supset & \mathrm{F1(f)} \\ \mathrm{conv} & & \mathrm{multi} \\ \mathrm{w(t)} & \supset & \mathrm{W(f)} \\ \| & & \| \\ \mathrm{f2(t)} & \supset & \mathrm{F2(f)} \end{array}$$

Moreover, the rectangular function and the sinc function are a transform pair, that is, in the time domain the sinc function is convolved with the original data to produce the filtered data. According to the similarity theorem mentioned in section 3.2, the narrower the rectangular function in the transform domain (the narrower the passband), the wider the time and the lower the amplitude of sinc function in time domain, so that the wider sinc function makes the filtered data smoother.



(a) (b)

Figure 3.2 Comparison between original data and (a) IFFT with 50Hz low-pass filter in small amounts; (b) IFFT with 10Hz low-pass filter in small amounts

Figure 3.2 is a detailed comparison between a 50Hz and a 10Hz low-pass filter. The character and amplitude of the 10Hz low-pass filtered data are smoother and smaller than those of the 50Hz low-pass filtered data because the 10Hz low-pass band used is narrower than the 50Hz low-pass band in the transform domain.

3.4 Power spectral density

The power spectral density (PSD) is often used when the energy of a signal is concentrated over a limited time interval, especially when its total energy is limited. It refers to the spectral energy distribution per unit time, which is applicable to signals that exist at all times or for a sufficiently large period, or within an infinitely long-time interval (Petre and Moses, 2005).

For continuous signals in time, the PSD describes how the power of a signal or time series is distributed over frequency. It shows the strength of the energy as a function of frequency. In other words, it shows at which frequencies the energy is strong and at which frequencies it is weak. The power can be the actual physical power, and the abstract signal that is more convenient to process can be simply represented by the squared value of the signal (the energy).

For some signals of interest, the truncated Fourier transform is used, that is, the signal is only integrated over a finite interval to calculate the Fourier transform $\mathrm{F(s)}$ to analyze the frequency content of the signal $\mathrm{f(t)}$.

$$\mathrm{F(s)} = \frac{1}{\sqrt{T}} \int_0^T f(t) e^{-iwt} dt \tag{3.15}$$

This is the amplitude spectral density and the power spectral density can be defined as(Miller and Childers, 2012):

$$\mathrm{P(f)} = \lim_{T\to\infty} E[|F(s)|^2] \tag{3.16}$$

here E is the expected value. Equation (3.16) can be written as(Miller and Childers, 2012):

$$E[|F(s)|^2] = \frac{1}{T}\int_0^T \int_0^T E[f^*(t)f(t')]e^{iw(t-t')}dtdt' \tag{3.17}$$

Equation (3.17) shows the PSD is a Fourier transform pair with the autocorrelation function of the signal. This is the same result as described in equation (3.11), which is the autocorrelation theorem.

3.5 Window function

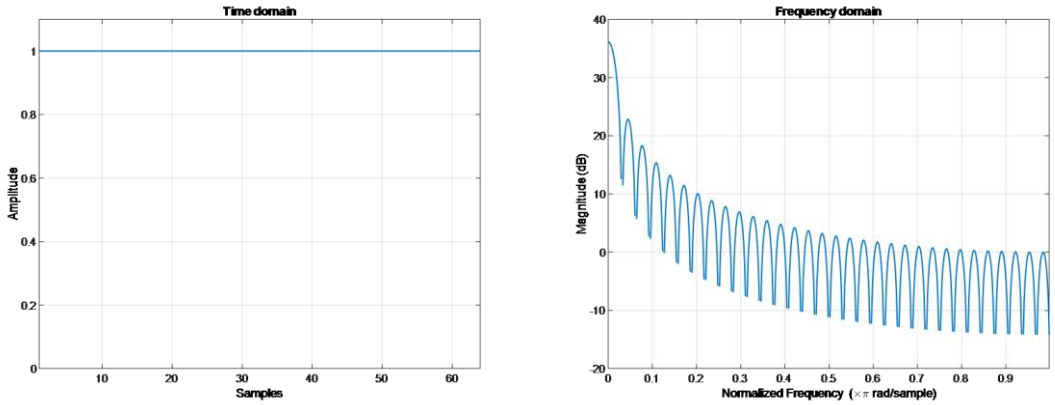
A window function is zero outside a selected interval and is usually symmetric and close to the maximum in the middle of the interval. When a function is multiplied by a window function, the product is almost zero outside the interval, that is, only the overlaps are left, that is, the "window view" (Weisstein, 2003).

Since the FFT can only transform finite-length time domain data, it is sometimes necessary to perform signal truncation on the time domain signal. Even for periodic signals, if the length of the truncation is not an integer multiple of the period, there will be a leak in the intercepted signal. In order to minimize this leakage error, a weighting function, also called a window function, is required. Windowing can make the time domain signal better meet the periodic requirements of FFT processing and reduce leakage(National Instruments, 2019).

Different window function weightings are different, and the time domain shape and frequency domain characteristics of different time windows are different. Different window functions can be used for signal interception based on signal type and analysis purpose. The most commonly used window functions are the Rectangular, Hann, Hamming, Blackman, Kaiser and so on. The following are the descriptions of the Rectangular window, Hann window and Hamming window, where M is used for the length of the impulse response which is equal to 64.

- Rectangular window

$$\mathrm{w(n)} = \begin{cases} 1, & 0 \le n \le M, \\ 0, & otherwise \end{cases} \tag{3.18}$$



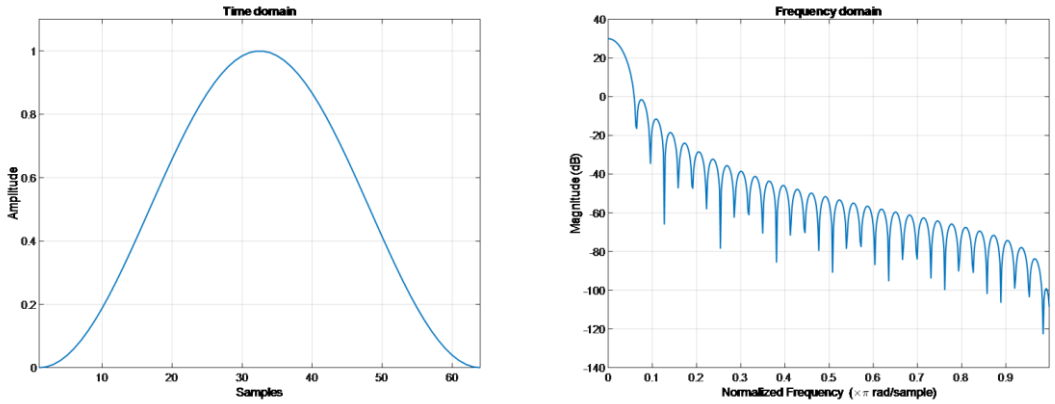
(a) (b)

Figure 3.3 Rectangular window in (a) time domain and (b) frequency domain

The rectangular window is the simplest window, and all data values except M are 0. Passing a rectangular window is equivalent to giving the signal of length M no window. The advantage of this kind of window is that the main lobe is relatively concentrated, that is, the Rectangular window has the narrowest main lobe, and thus, for a given length, it should yield the sharpest transitions at a discontinuity. The disadvantage is that, as Figure 3.3(b) shows, the first side lobe is only about 13 dB below the main peak, resulting the oscillations of considerable size around discontinuities. The side lobes are high and have negative side lobes, which leads to high frequency interference and leakage, and even a possible negative spectrum phenomenon(Oppenheim et al., 1998)(Wikipedia contributors, 2019c).

- Hann window

$$\mathrm{w(n)} = \begin{cases} 0.5(1-\cos(\frac{2\pi n}{M})), 0 \le \mathrm{n} \le \mathrm{M} \\ 0, otherwise \end{cases} \tag{3.19}$$



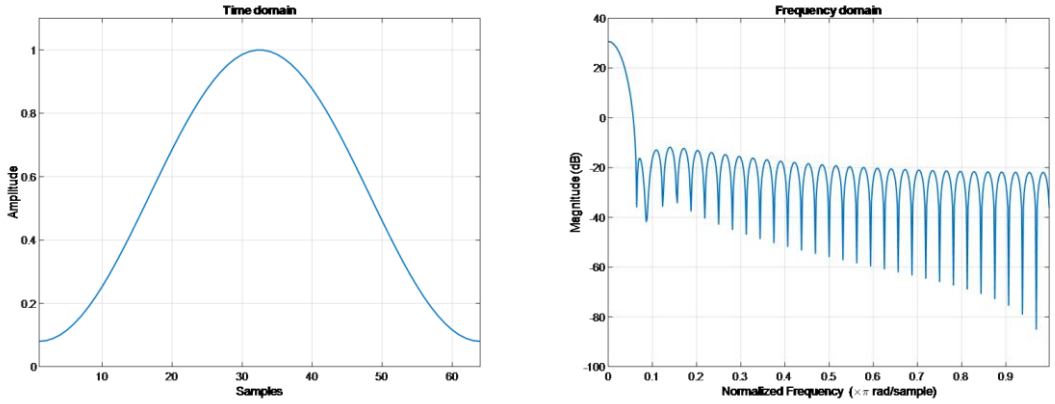
(a) (b)

Figure 3.4 Hann window in (a) time domain and (b) frequency domain

The Hann window is also called a raised cosine window. The Hann window can be viewed as the sum of the spectra of the three rectangular time windows, or the sum of the three $\mathrm{sinc}(\mathrm{t})$ functions, which cancel out the side lobes, eliminating high-frequency interference and leakage energy. The Hann window is superior to the rectangular window in reducing leakage view, but the widening of the Hann window main lobe is equivalent to the broadening of the analysis bandwidth and the decline of frequency resolution(Harris, 1978).

- Hamming window

$$\mathrm{w(n)} = \begin{cases} 0.54 - 0.46\cos(\frac{2\pi n}{M}), 0 \le \mathrm{n} \le \mathrm{M} \\ 0, otherwise \end{cases} \tag{3.20}$$



(a) (b)

Figure 3.5 Hamming window in (a) time domain and (b) frequency domain

The Hamming window causes the first sidelobe of Hann window to be largely eliminated, resulting in a sidelobe only 1/5 the height of the Hann window.

The Table 3.1 shows some parameters of the three windows, the values were given by MATLAB.

Table 3.1 Parameters of windows

	Leakage Factor (%)	Relative sidelobe attenuation (dB)	Main lobe width (-3dB)
Rectangular window	9.14	-13.3	0.027344
Hann window	0.05	-31.5	0.042969
Hamming window	0.03	-42.5	0.039063

The main purpose of windowing is to replace the rectangular window function present on all the signal samples with a relatively smooth window function, so that the mutations at both ends of the truncated time domain waveform become smoother or smaller. Because the

leakage of the first sidelobe is the largest, the first sidelobe of the spectrum window should be lowered to reduce the leakage (Tan, 2017). Table 3.1 shows the most attenuation of side lobes is the Hamming window, and its corresponding leakage is the least, as shown in the Table 3.1.

The main lobe width primarily affects signal energy distribution and frequency resolution. The actual resolution in the frequency domain is the product of the effective noise bandwidth and the frequency resolution. Therefore, the wider the main lobe, the wider the effective noise bandwidth, and the worse the frequency resolution is in the case of the same frequency resolution(National Instruments, 2012). Comparing the three windows in Table 3.1, a Rectangular window has the narrowest main lobe, while a Hann window has the widest. The narrower the main lobe, the more accurate the frequency resolution, so the frequency resolution of a Rectangular window is higher than in the other two windows.

3.6 Short Time Fourier Transform

Another form of Fourier transform is the Short Time Fourier Transform (STFT), also known as a Windowed Fourier Transform or Short Distance Fourier Transform, which is used to determine the frequency and phase of a local part of a signal that varies with time. In fact, the process of calculating the STFT is to divide the long signal into several short isometric signals, and then calculate the Fourier transform of each short segment. The reason for using this method is that for signals whose amplitude and frequency do not change too quickly, short fragments can be represented by signals whose frequency is close to constant (Zimmer, 2011). The Fourier transform only provides information on frequency component amplitudes but does not provide time information; the STFT clearly provides both kinds of information. This method of time-frequency analysis is extremely helpful in analyzing signals whose frequency changes with time.

The amplitude $\mathrm{X}(\mathrm{t}, \mathrm{f})$ of the Short Time Fourier Transform is used to present a standard procedure for the time-varying signal, that is, a spectrogram. The definition of the STFT is:

$$\mathrm{X}(\mathrm{t}, \mathrm{f}) = \int_{-\infty}^{\infty} w(t-\tau)\, x(\tau) e^{-i2\pi f\tau} d\tau \tag{3.21}$$

Here equation (3.21) shows to start in the time domain, take the one-dimensional Fourier transform of a function $\mathrm{x}(\mathrm{t})$ multiply by a window function $\mathrm{w}(\mathrm{t})$ (that is not zero for only a period of time), and then move this window function along the time axis to get a set of

Fourier transforms $X(t,f)$ that are arranged in a two-dimensional representation. The result is a complex function representing the magnitude and phase of the signal as a function of time and frequency (Allen, 1977).

3.7 Spectrogram

The two-dimensional spectrum obtained by the Fourier transform decomposes a complex wave into spectral components, but it cannot simultaneously reflect their changes over time. The Short Time Fourier Transform is capable of simultaneously analyzing the time variable and the frequency distribution of the wave, but the three-dimensional image is not easy to observe and analyze on paper. The spectrogram can describe how the various frequency components of the wave change over time, and based on the time-frequency analysis method, the third-dimensional value is usually represented by color in the form of a heat map(Smith, 2011).

Time-frequency analysis uses the STFT to analyze the signal. The DFT is performed after multiplying by a window function, so the frequency component of the signal can be obtained in this short period of time (Allen and Rabiner, 1977). Finally, time-frequency analysis of the entire signal can be used to obtain the time-frequency distribution of the signal. This is the principle of spectrogram generation, and this process essentially corresponds to computing the squared magnitude of the STFT of the signal s(t), that is, for a window w, $\text{spectrogram}\ (t, w) = |\text{STFT}\ (t, w)\ |^2$(National Instruments, 2008).

Take a small piece of a measured signal for example, as shown in the figures below of a 3-millisecond signal containing 35 sample points. A 10-point Hann window function is used with the sample data, with a different number of zeros padded before and after the window. The spectrogram is obtained by the STFT of the windowed signal. The figures below show the main process.

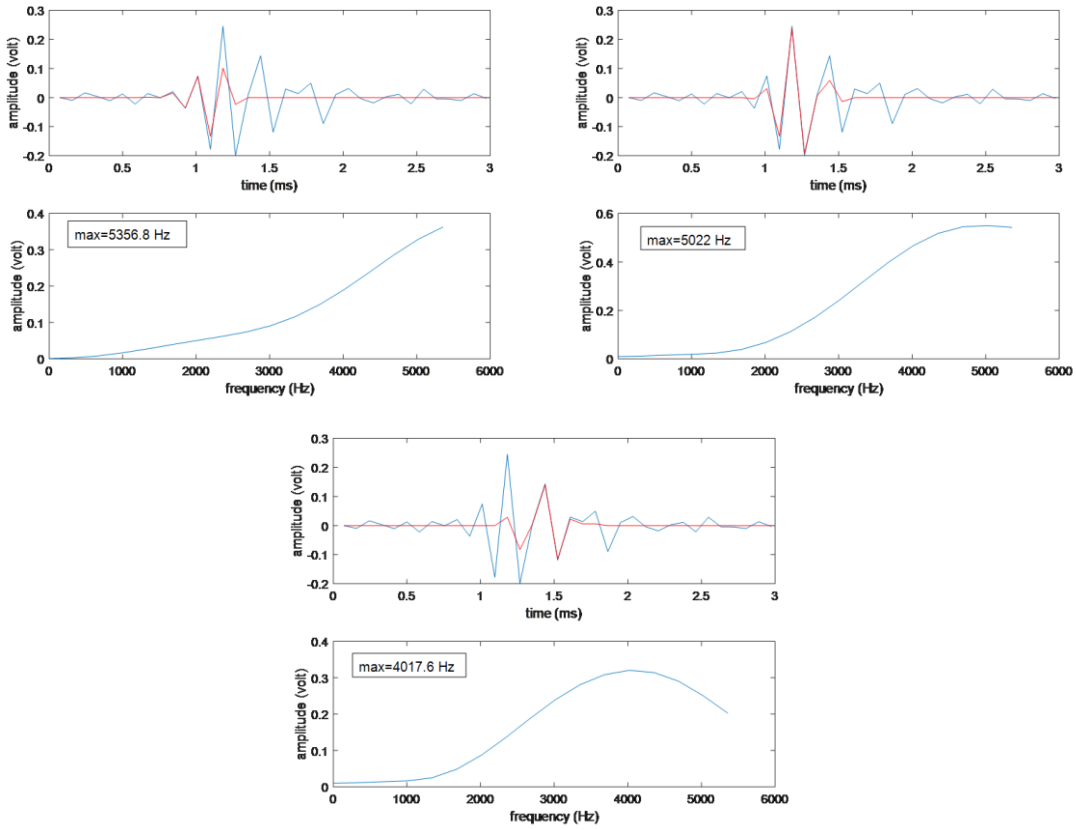


Figure 3.6 Signal with Hann window

Figure 3.6 shows the Hann window function applied to a piece of signal at different times of 1.0112 ms, 1.18187 ms and 1.43789 ms, and the red lines highlight the Hann window code fragment for the Fourier transform, whose results are shown below the time series. The corresponding maximum frequencies appear at 5356.8 Hz, 5022 Hz and 4017.6 Hz, respectively. These three frequencies with strong amplitudes appearing at specific time points appeared in brighter colors on the spectrogram as below.

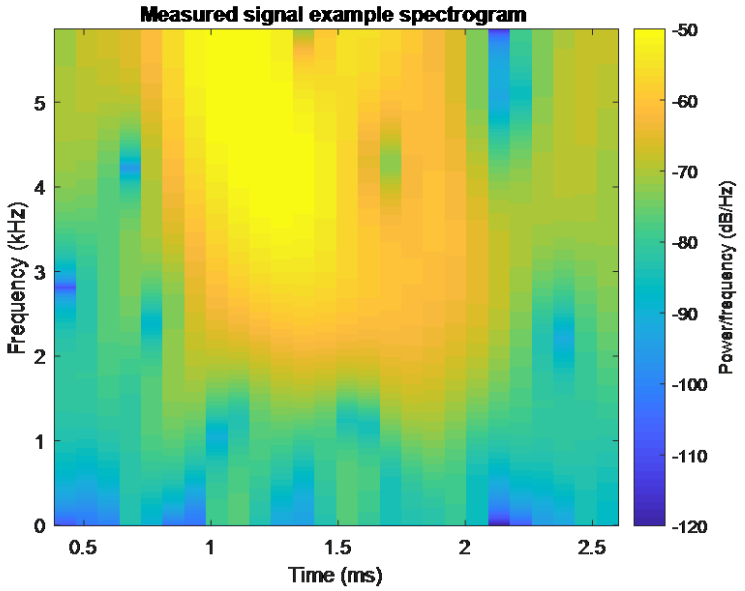


Figure 3.7 Signal spectrogram

The spectrogram shows time and frequency in two dimensions. Figure 3.7 shows the spectrogram of the same 3 milliseconds as in Figure 3.6. The main signal pulse occurs in a time range of 1 to 1.5ms, the same as in Figure 3.6, and with center frequency around 4250 Hz, and frequency range from 3000Hz to 5500Hz, which includes the three frequencies with strong amplitude shown in Figure 3.6.

3.8 Matched filter

Convolution is an overlapping relation between two functions, and the result reflects the overlap of the two functions. Convolution of a function of a given frequency band with a function of another frequency band is to filter the latter, where only the overlapped frequency band of the latter can pass through the filter well.

In signal processing, the correlation between a known signal and an unknown signal is used to detect the existence of the known signal in the unknown signal. In other words, the unknown signal is convolved with the conjugate time inversion of the known signal to obtain a matching filter (Turin, 1960).

A matched filter is the optimal linear filter to maximize the signal-to-noise ratio (SNR). The following section derives the matched filter (Wikipedia contributors, 2019d):

$$\mathrm{y}[n] = \sum_{k=-\infty}^{\infty} h[n-k]x[k] \tag{3.22}$$

where $x[k]$ is the input function and y[n] is the filtered output. Equation (3.22) is the discrete-time system expression. The continuous-time system expression is similar, with summations replaced with integrals.

The input signal x is composed of desirable signal s and additional noise v:

$$x = s + v \tag{3.23}$$

The output y is the inner product (convolution) of the filter h and the input signal x. It also consists of the y_s -outcome of the signal of interest through the filter system and y_v -outcome of the input signal's noise part through the filter system:

$$y = \sum_{k=-\infty}^{\infty} h^*[k]x[k] = h^H x = h^H s + h^H v = y_s + y_v \tag{3.24}$$

where h^H denotes the conjugate transpose of h. The conjugate transpose of a matrix A is formally defined by $A^H=(A^*)^T=(A^T)^*$, where A^T denotes the transpose and the A^* denotes the matrix with complex conjugated entries(Wikipedia contributors, 2019e).

The signal-to-noise ratio is the ratio of the output power of the desired signal to the output power of the noise:

$$\mathrm{SNR} = \frac{|y_s|^2}{E\{|y_v|^2\}} \tag{3.25}$$

where E denotes expectation. The normalized matched filter can be derived to be(Kay, 1993),

$$h = \frac{1}{\sqrt{s^H R_v^{-1} s}} R_v^{-1} s \tag{3.26}$$

$$R_v = E\{vv^H\} \tag{3.27}$$

where v^H denotes the conjugate transpose of v, and R_v is the covariance matrix of the noise.

For a given signal and system, amplitude-frequency characteristics more represent the frequency characteristics, while phase-frequency characteristics more represent the time characteristics. The matched filter makes the phase frequency characteristic of the filter conjugate with the phase frequency characteristic of the signal and makes all the frequency components of the output signal superimposed in the same phase at the output end to form the peak value. The matched filter can also make the amplitude-frequency characteristics of the signal weight the input waveform so as to receive the signal energy most effectively and

suppress the output power of the noise. Whether in the time domain or frequency domain, the matched filter fully ensures that the signal passes through as much as possible, and the noise passes through as little as possible, so as to obtain the maximum SNR output (Haykin, 2001).

In practical application of the matched filter, it is often necessary to construct a replica which is a signal waveform that can be used as input to the matched filter. A very high quality (noise free) real signal or an artificial signal that reflects the real signal frequency function as accurately as possible can be used as a replica. Between replica and signal data, the FFT-based finite impulse response (FIR) filtering using overlap-add method will filter the data (Zimmer, 2011).

3.9 Cross-correlation

The cross-correlation function in signal analysis represents the degree of correlation between the two-time series, that is, the degree of correlation between the values of two signals at any two different times. When describing the correlation between two different signals, the two signals can be either random signals or well-known signals.

The cross-correlation $\mathrm{f(x)}$ of two real functions $\mathrm{g(x)}$ and $\mathrm{h(x)}$ is defined as (Bracewell, 2000):

$$\mathrm{f(x)} = \mathrm{g} \star \mathrm{h} = \int_{-\infty}^{\infty} \mathrm{g(u-x)h(u)du} \tag{3.28}$$

When the functions are complex it is customary to define the (complex) cross-correlation function as g* scans h, where g* is the complex conjugate of g

$$g^* \star \mathrm{h} = \int_{-\infty}^{\infty} g^*(u-x)h(u)du = \int_{-\infty}^{\infty} g^*(u)h(u+x)du \tag{3.29}$$

For discrete functions, the cross-correlation is defined as:

$$g \star h(\mathrm{n}) = \sum_{\mathrm{m}=-\infty}^{\infty} g^*(\mathrm{m}-\mathrm{n})\mathrm{h(m)} \tag{3.30}$$

Mathematically, cross-correlation can be considered to be figure out how much h should move on the x axis to be equal to g, that is, slide the h function along the x axis, calculate the integral of their product at each h position. When the functions match, the value of $\mathrm{g}\star\mathrm{h}$ is maximized. In practical applications, cross correlation is often used to search in longer signals for shorter known features. After calculating the cross-correlation of two signals, the maximum value of the cross-correlation function represents the time point for the optimal alignment of signals.

3.10 Wavelet denoising

Wavelets are basis function $w_{j,k}(t)$ in continuous time. A basis is a set of linearly independent functions that can be used to produce all admissible function $f(t)$(Strang and Nguyen, 1996).

$$f(t) = \sum_{j,k} b_{j,k} w_{j,k}\left(2^j - k\right) \tag{3.31}$$

The time series $f(t)$ has the characteristics of nonlinearity and high signal to noise ratio. Wavelet theory is developed according to the requirements of time-frequency localization. It has the characteristics of adaptive and "mathematical microscopy" and is especially suitable for the processing of non-stationary and nonlinear signals("7 The Mathematical Microscope: Waves, Wavelets, and Beyond | A Positron Named Priscilla: Scientific Discovery at the Frontier | The National Academies Press," n.d.).

The wavelet method used for time series denoising is nonlinear thresholding, best used when the signal-to-noise ratio is large. The energy of the useful signal after wavelet transform is concentrated in a small number of wavelet coefficients, while white noise is still scattered over a large number of wavelet coefficients in the wavelet transform domain. The value of the wavelet coefficient of the useful signal is relatively larger than the value of the wavelet coefficient of the noise with less energy and smaller amplitude. Therefore, from the magnitude of the coefficients in the spectrum, the useful signal and noise can be separated.

Select the appropriate orthogonal wavelet base $w_{j,k}(t)$ and the decomposition layer number j and transform the noisy signal into j wavelet layers by using wavelet transform. To perform threshold processing on the decomposed wavelet coefficients, two types of threshold can be used: hard and soft. The hard threshold method preserves larger wavelet coefficients and sets the smaller wavelet coefficients to zero; the soft threshold method sets the smaller wavelet coefficients to zero and contracts the larger wavelet coefficients toward zero. The estimated signal after denoising by the soft threshold method is an approximate optimal estimate of the original signal and often has wider applicability.

Chapter 4. Results

4.1 Known fin whale sound signals

Two audio recordings will be used as the kernels for matched filtering algorithm to detect fin whale sound signal in the Gulf of Mexico data. One audio file is from the National Oceanic and Atmospheric Administration (NOAA) (NOAA, 2019c) and the other audio file is from the Voices in the Sea(Voices In The Sea, 2018). Though it is not clear exactly which ocean the signals come from, the two audio files clearly indicate different properties of a fin whale calls/notes.

Figure 4.1 shows the signal amplitude versus time for a fin whale pulse train; Figure 4.2 shows the amplitude versus frequency; and Figure 4.3 shows the spectrogram (frequency versus time) for the pulse train.

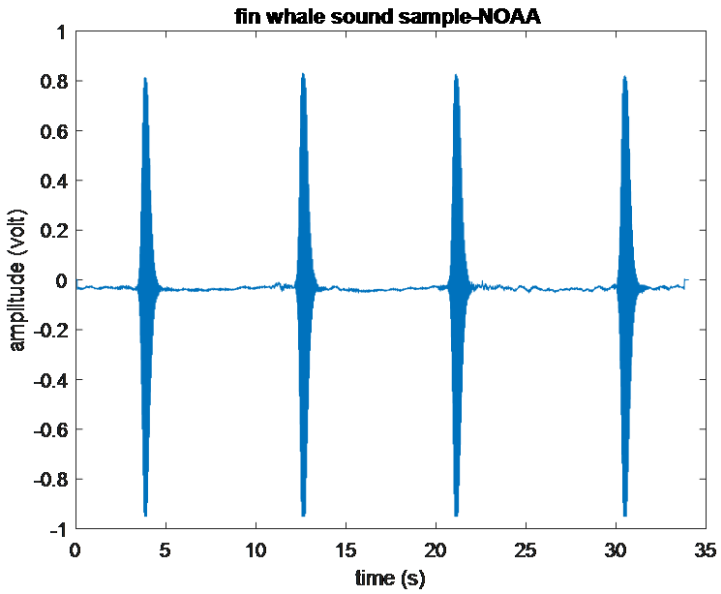


Figure 4.1 Pulse of fin whale sounds (audio file from NOAA)

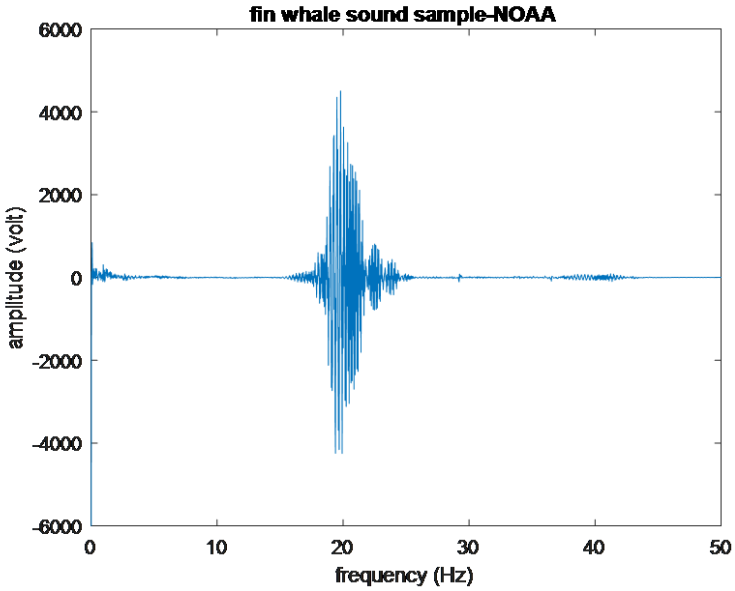
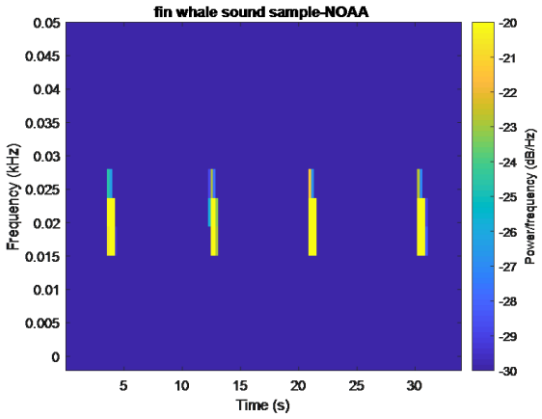
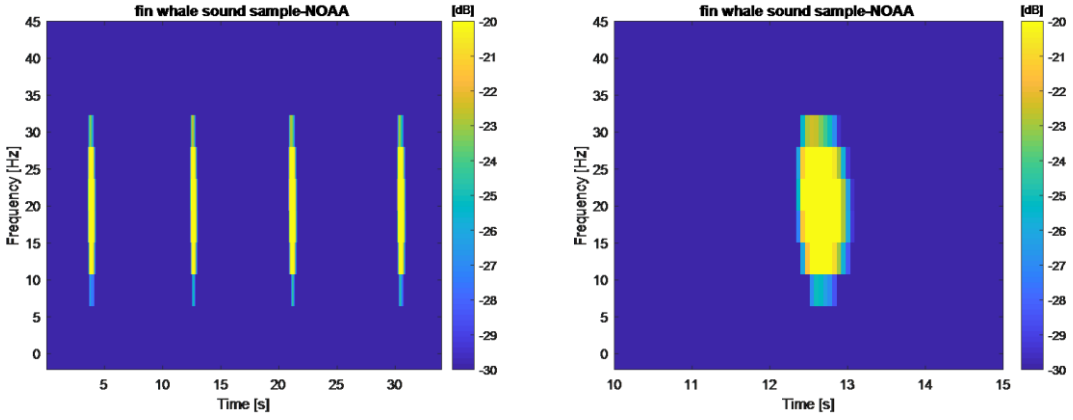


Figure 4.2 Fourier Transform of fin whale sounds (audio file from NOAA)



(a)



(b) (c)

Figure 4.3 Pulse train of fin whale sounds spectrogram(audio file from NOAA) (a)pulse train sample in kilohertz; (b) pulse train sample in hertz; (c)single down-sweep sample

Figure 4.1, Figure 4.2 and Figure 4.3 are fin whale sounds recorded by NOAA. This sound is a very simple form of a 27-15Hz down-sweep signal with stereotyped center frequency 20Hz and without back-beat, and the pulse duration is about 1.5s. The time recording is not long enough to determine the pulse is a call or note, but the graph shows this series of sounds has a regular average inter-pulse interval (IPI) of 9s, so it is probably a regular series pattern of notes that compose a fin whale singlet song of repeating sequences with a single IPI (Weirathmueller et al., 2017).

This recording could come from the central Mediterranean Sea where a fin whale type "A" call was recorded with a down-sweep from 23Hz to 17Hz, a 1s call duration and a 6-46s IPI (Sciacca et al., 2015). However, this recording could also be from the Northwest Atlantic where a fin whale song was recorded during September to January with a short pulse interval of 9.6s, a down-sweep from 25Hz to 17Hz, and a 1s pulse duration (Morano et al., 2012).

Figure 4.4, Figure 4.5 and Figure 4.6 are fin whale sounds recorded by Voices in the Sea. This series of sounds is more complicated than that recorded by NOAA. It is not just a series of single form signals but includes sounds with different amplitudes and frequency. The time recording is not long enough to determine the pulse is a call or note, but Figure 4.4 shows the pulse duration is close to 2s and the IPI of each signal is different. The IPI includes 8s, 17s, 12s, 14s, and 11s. According to the dominant IPI patterns, this could be classified as complex IPI (Weirathmueller et al., 2017). In addition, Figure 4.6(a) shows the center frequency of each signal is not exactly the same, so this may be an example of baleen whale low frequency sounds used for communication. If the recording time is long enough for a fin whale song, this could be an example of a double song including two primary notes, alternating between a low frequency, high IPI (note type A) and a higher frequency, lower IPI (note type B) (Weirathmueller et al., 2017), and relatively high frequency signals tend to have lower power than the center frequency. Figure 4.6(b) is a single pulse of Figure 4.6(a) which shows the typical down-sweep signal shape of a fin whale. In addition, Figure 4.5 indicates the sound frequency is approximately 15-40Hz. Usually the frequency of a mature fin whale does not exceed 30Hz, so one possibility is that the 40Hz frequency sound recording here is produced by a juvenile fin whale.

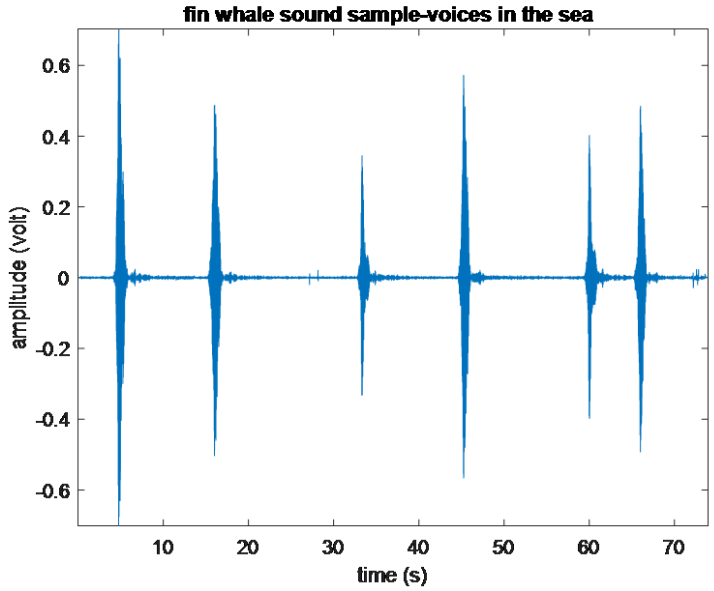


Figure 4.4 Pulse of fin whale sounds (audio file from Voices in the Sea)

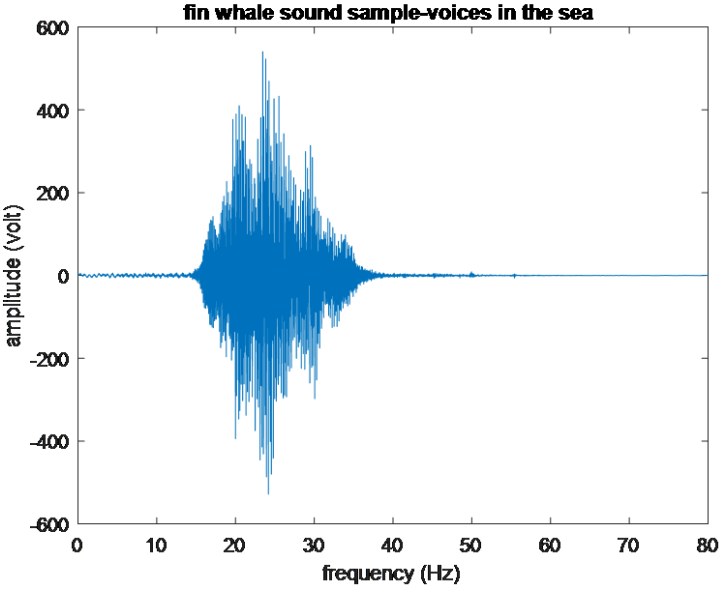


Figure 4.5 Fourier Transform of fin whale sounds (audio file from Voices in the Sea)

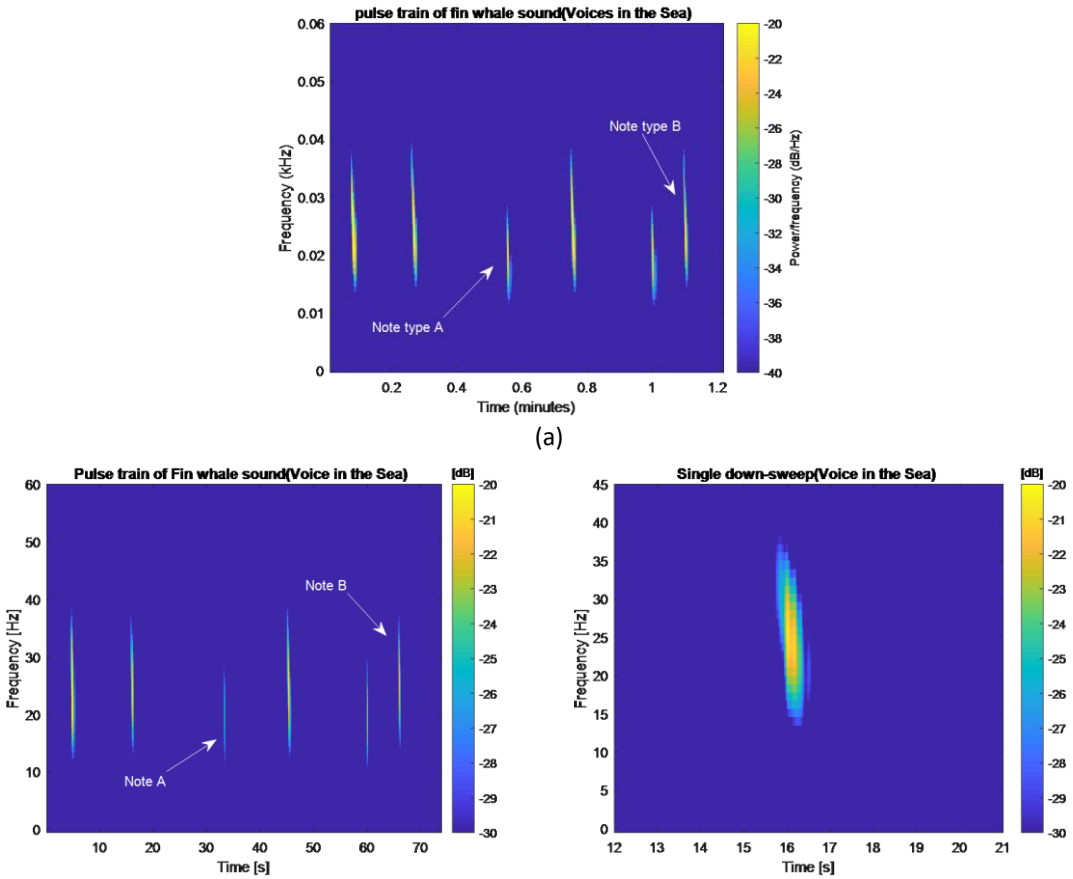


Figure 4.6 Fin whale sounds spectrogram (audio file from Voices in the Sea) (a)pulse train sample in kilohertz; (b) pulse train sample in hertz; (c)single down-sweep sample

According to the diversity and complexity of this recorded sound interval and frequency, this recording could come from the Northeast Pacific Ocean where irregular sequences with complex IPIs (Soule and Wilcock, 2013) produced by more than one fin whale have been recorded. Furthermore, this sound recording could also be from the Northwest Atlantic where additional recordings showed another five types of IPI including a short IPI around 9s in autumn and winter, a long IPI around 15s in spring and a transitional IPI during summer (Morano et al., 2012).

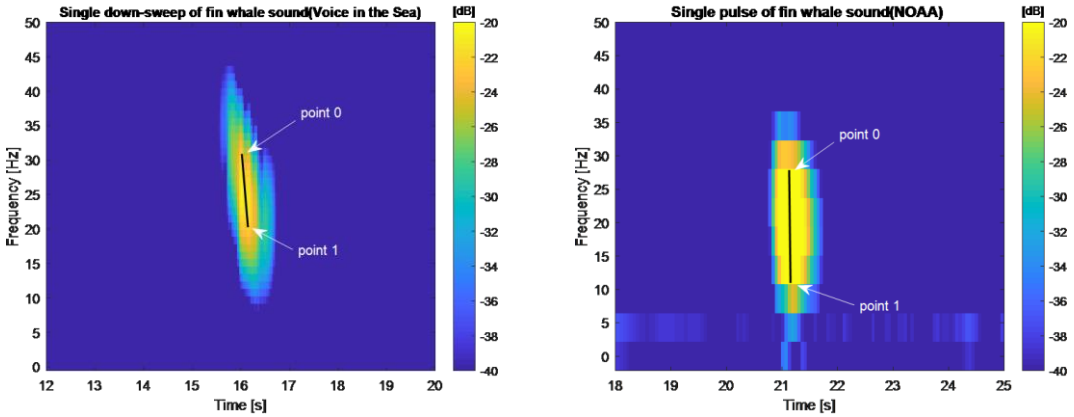
4.2 Fin whale sounds detection

In new records that are considered to contain the type of sounds of interest, the kernel is used to automatically detect the presence of the desired sound type. Detection is achieved by cross-correlation between the synthesized kernel and the new record. When the signal

characteristics in the new record closely match the characteristics of the nuclear sound type, the function will show significant peaks/spikes.

4.2.1 Synthesize fin whale matched filter kernel

Matched filtering is the best method to detect a known signal in white Gaussian noise(Van Trees, 2001). Figure 4.7 shows a spectrogram of a single fin whale vocalization with a line drawn through the maximum values (yellow part). The signal structure is an approximately-linear frequency down sweep with a steep negative slope in the left graph, and almost vertical in the right graph.



(a) (b)

Figure 4.7 A spectrogram of a single fin whale vocalization (a) single down-sweep (Voice in the Sea); (b) single pulse (NOAA)

There is only one-line segment with two points to describe this sound: the starting point 0 and the ending point 1. The values for time, frequency and amplitude should be measured at these two points. Selected points will be used to synthesize a signal representing this sound type in frequency-time space.

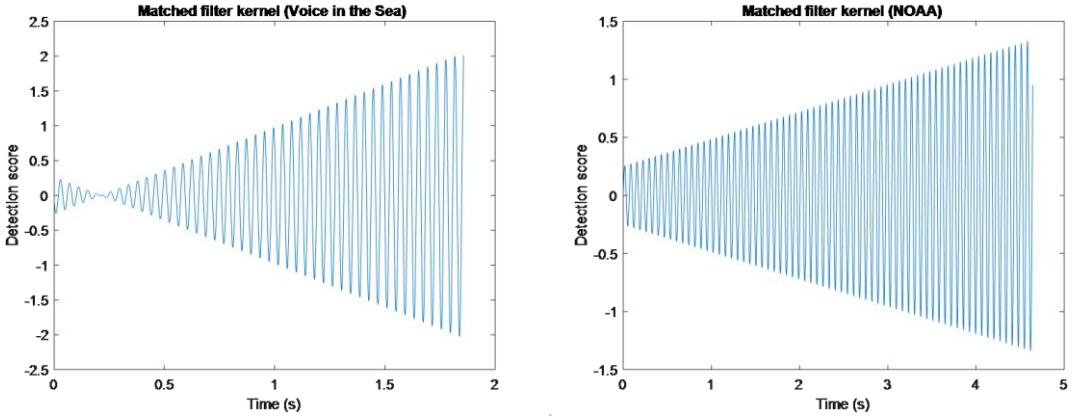
Define f_0 as the start frequency of the segment, f_1 as the end frequency, d the duration of the segment, t the time; then $f(t)$ is the frequency at a particular instant in time, a_0 is the start amplitude, a_1 is the end amplitude, $a(t)$ is the amplitude at time instant t, and $k(t)$ is the kernel (Mellinger and Clark, 1997).

$$f(t) = f_0 + \frac{1}{2}\frac{(f_1 - f_0)}{d}t \tag{4.1}$$

$$a(t) = a_0 + \frac{(a_1 - a_0)}{d}t \tag{4.2}$$

$$k(t) = a(t)\sin(2\pi \cdot f(t)) \tag{4.3}$$

In Figure 4.7(a), the values for the two end points are f_0=31.22 Hz, f_1=19.38 Hz, a_0=-0.39436 volt, a_1=-0.35572 volt, and d=0.24 second. In Figure 4.7(b), the values for the two end points are f_0=27.995 Hz, f_1=10.7665 Hz, a_0=0.1315 volt, a_1=0.21436 volt, and d=0.69 second. Figure 4.8 shows the synthetic matched filters $k(t)$ in the time domain.



(a) (b)

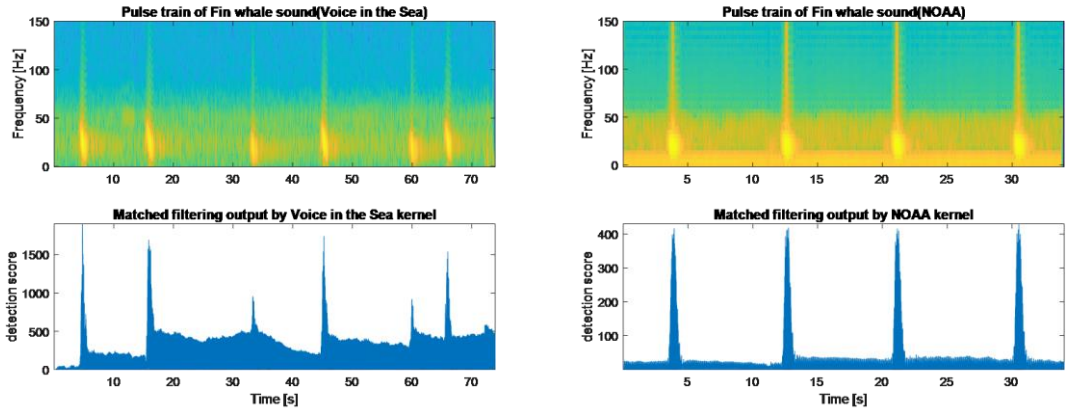
Figure 4.8 Matched filter in time domain (a) from Voice in the Sea; (b) from NOAA

The kernel is chosen from the spectrogram, but the above equations clearly show that matched filtering uses the sound's time-varying waveform in the actual computation.

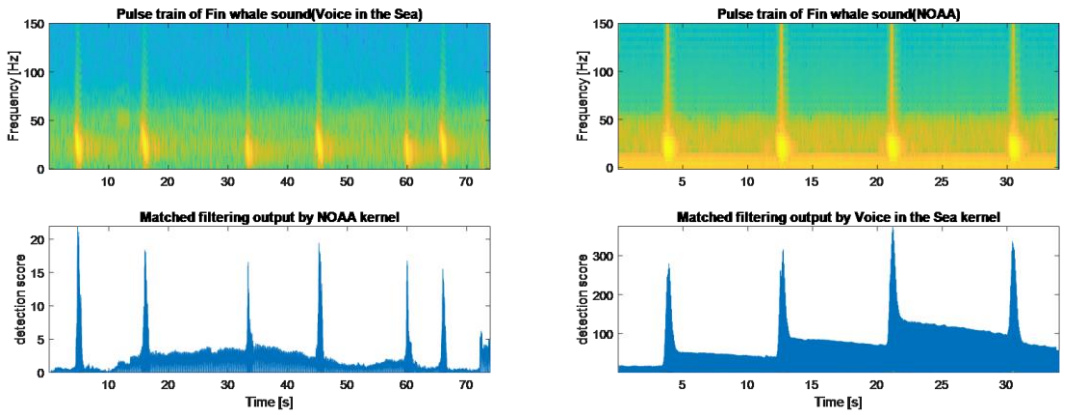
$x(t)$ is the signal which may contain the sound to be detected. Matched filtering simply performs the cross-correlation, with $d(t)$ the filter output(Mellinger and Clark, 1997).

$$d(t) = \sum_{t_0} x(t_0 - t)\, k(t_0) \tag{4.4}$$

Figure 4.9 shows a comparison of the spectrogram and the output of the matched filter.



(a) (b)



(c) (d)

Figure 4.9 Comparison of the spectrogram and the output of matched filter (a) Synthetic Voice in the Sea kernel exactly matched with the corresponding provenance Voice in the Sea data; (b) Synthetic NOAA kernel exactly matched with the corresponding provenance NOAA data; (c) Synthetic NOAA kernel exactly matched with the signal in Voice in the Sea data; (d) Synthetic Voice in the Sea kernel exactly matched with the signal in NOAA

To test whether the matched filter kernel was synthesized effectively, the output shown in Figure 4.9(a) is the cross-correlation result between $k(t)$ synthesized from Voice in the Sea data and the Voice in the Sea sample audio file $x(t)$; the output shown in Figure 4.9(b) is the cross-correlation result between $k(t)$ synthesized from NOAA data and the NOAA sample audio file $x(t)$; the output shown in Figure 4.9(c) and 4.9(d) are the cross-correlation results between $k(t)$ synthesized from NOAA data with the Voice in the Sea sample audio file $x(t)$ and between $k(t)$ synthesized from Voice in the Sea data with the NOAA sample audio file $x(t)$, respectively.

The peaks/spikes in the lower graphs correspond to the time position of fin whale sounds, and the relative heights of the vertical axis values are significant. Compare the top and bottom figures, synthetic Voice in the Sea kernel and NOAA kernel exactly matched with their corresponding provenance data, and synthetic NOAA kernel and Voice in the Sea kernel also exactly matched with the signal of the other. Therefore, this test indicates that the synthetic kernels are effective and can be used to detect fin whale sound in the Gulf of Mexico LADC data after being denoised by bandpass filtering and wavelet transform.

4.2.2 Fin whale matched filter kernel applied to Gulf of Mexico LADC data

The kernels from Voice in the Sea were applied to samples of LADC 2001, 2015 and 2017 data, respectively.

The pulses in Figure 4.10 as below were numbered, and six kernels were synthesized for each of the six pulses, and then matching the input signal 2001(file jd21311443) with 6 kernels, respectively.

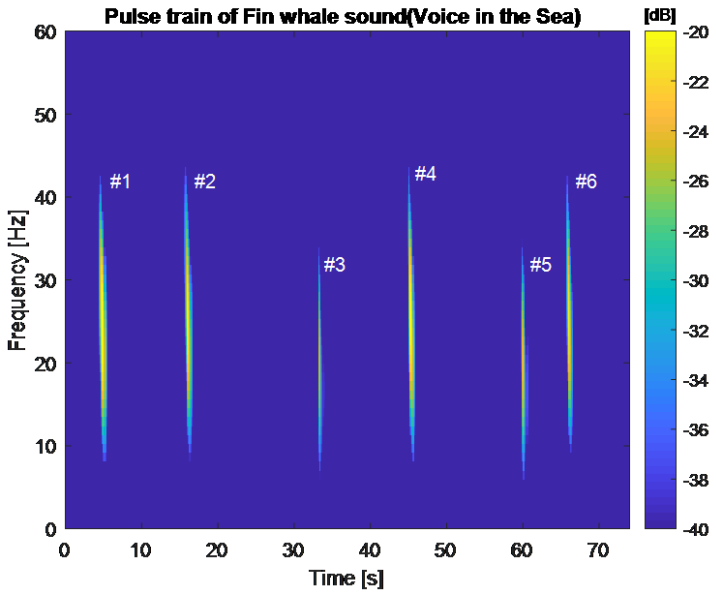
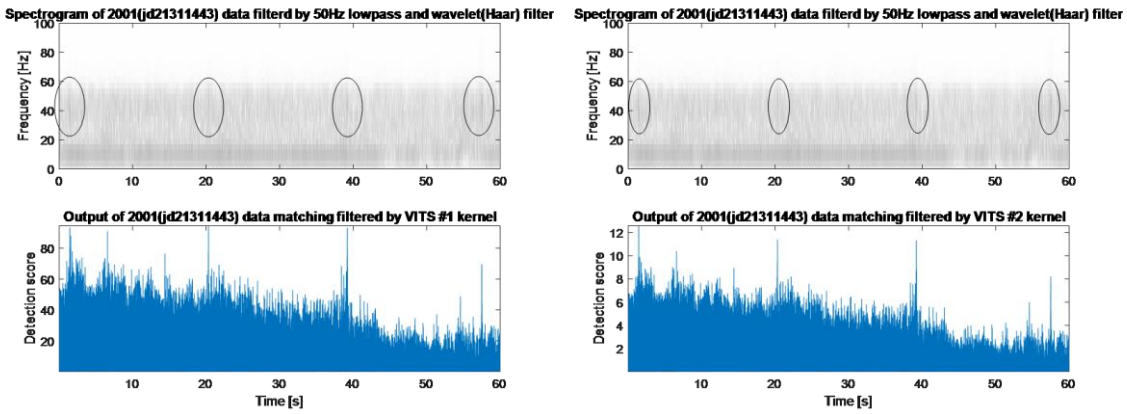


Figure 4.10 Numbering of a kernel sequence

After 50Hz lowpass and wavelet (Haar) filtering, the following six figures are the results of matching the six kernels with the input signal.



(a) (b)

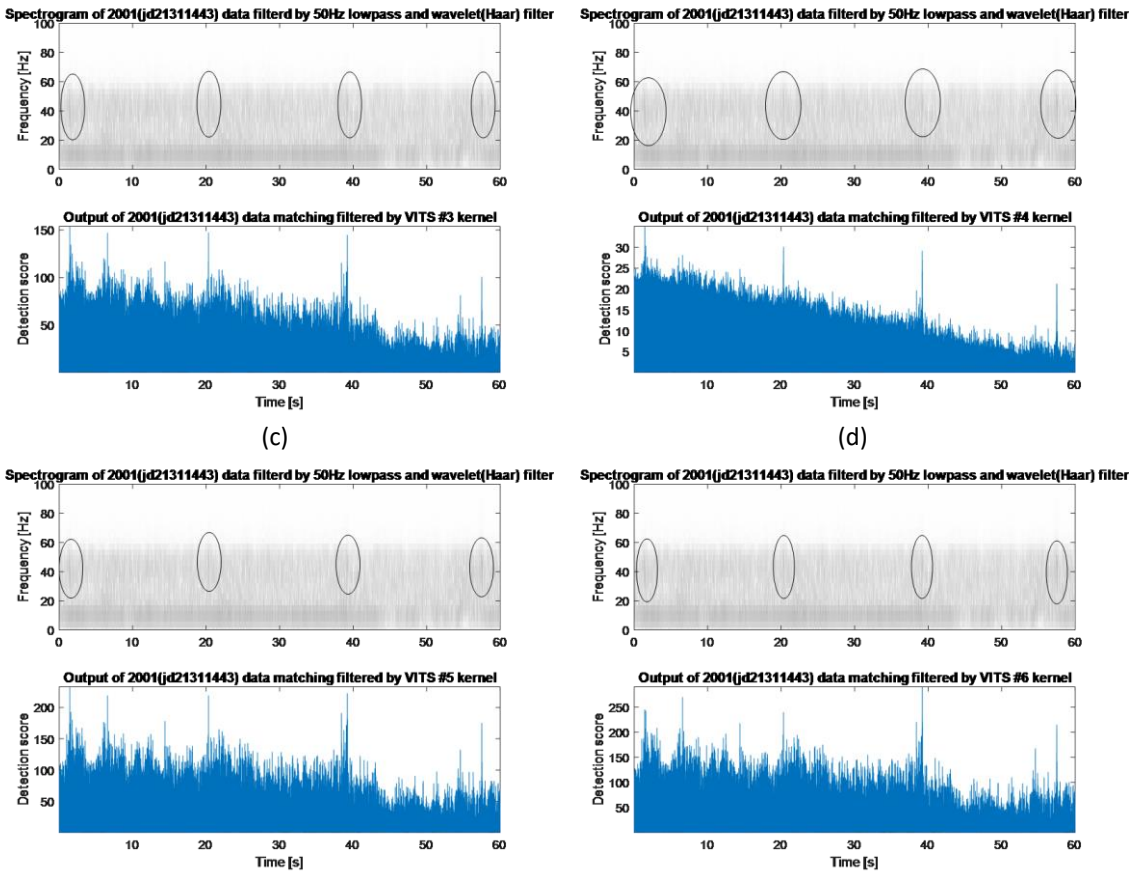
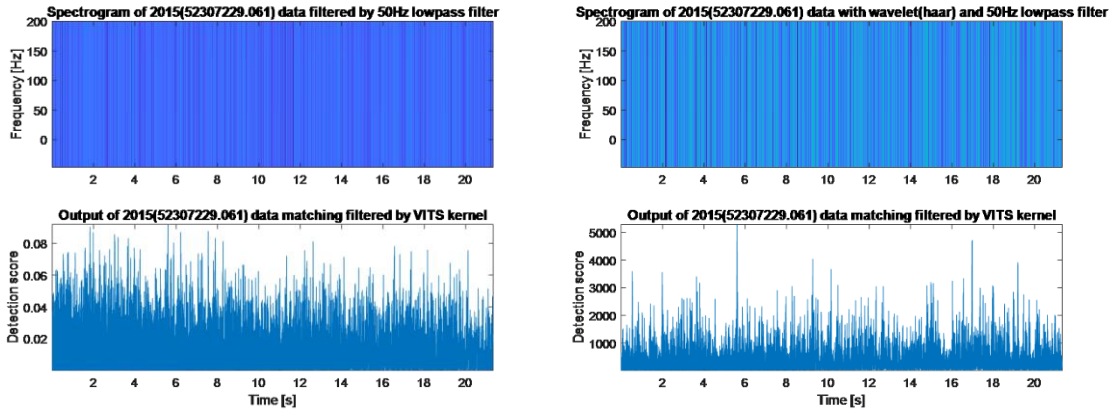


Figure 4.11 Comparison spectrograms of 50Hz lowpass and wavelet(Haar) filter with matching filtered by Voice in the Sea (a)#1 kernel; (b)#2 kernel; (c)#3 kernel; (d)#4 kernel; (e)#5 kernel; (f)#6 kernel

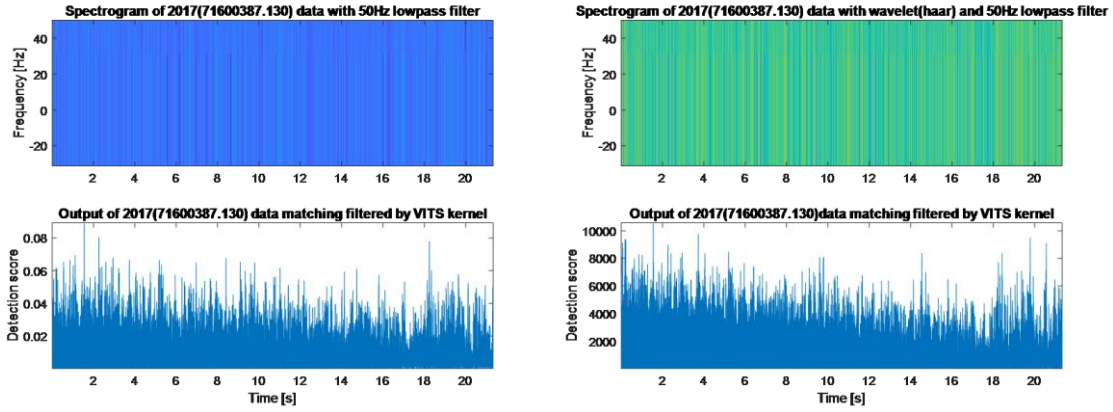
It can be seen in the spectrogram and matching figure that it seems that 4 signals will always appear; Figure 4.11(b) matched with the 4th kernel is most obvious, and the shape of the signal is in line with one of the typical signal shapes of fin whale--vertical down-sweep with 18s time interval. However, it is still not certain whether these signals are really fin whale sound signals because they are still not obvious/loud enough relative to background noise. Maybe further denoising is still necessary, or some other detection methods such as spectrogram correlation or neural networks are more suitable for finding low frequency baleen whale signals.

Applying the kernel to the 2015 and 2017 data, even after wavelet transform filtering, no signals of interest have been found so far.



(a) (b)

Figure 4.12 Comparison 2015 data spectrogram of (a)only 50Hz lowpass; (b)50Hz lowpass and wavelet filter with matching filtered by Voice in the Sea kernel

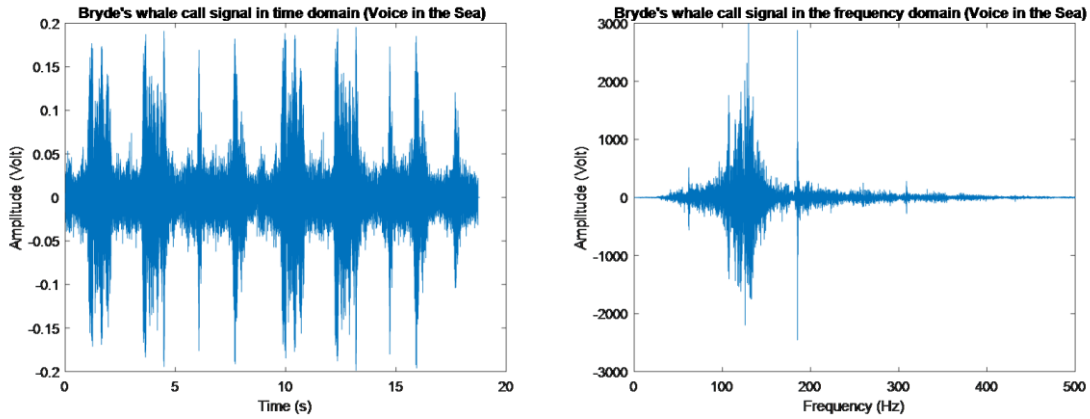


(a) (b)

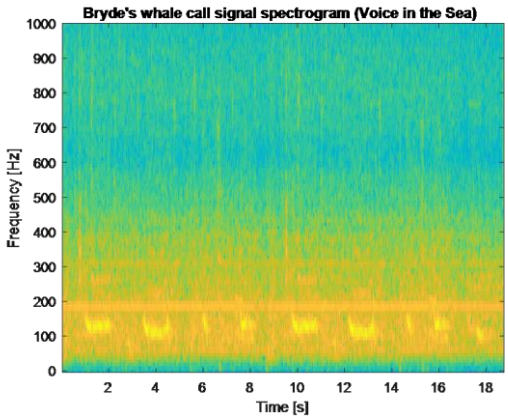
Figure 4.13 Comparison 2017 data spectrogram of (a)only 50Hz lowpass; (b)50Hz lowpass and wavelet filter with matching filtered by Voice in the Sea kernel

4.3 Known Bryde's whale call signals

A Bryde's whale call audio recording downloaded from Voice in the Sea, and it will be used to search for the Bryde's whale call in northern Gulf of Mexico LADC data.



(a) (b)



(c)

Figure 4.14 Bryde's whale call in (a)time domain; (b)frequency domain; (c)spectrogram

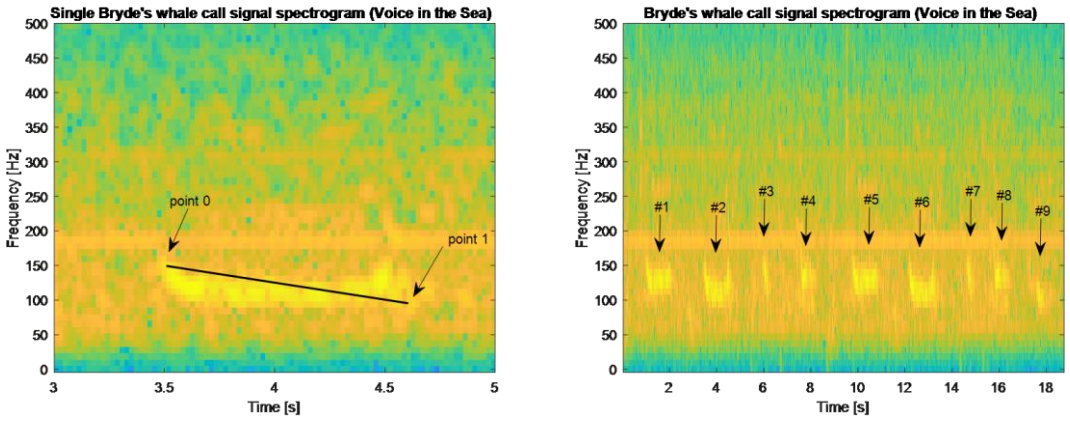
This sample audio file lasts about 20s. As can be seen from Figure 4.14(a), it consists of 9 pulses with different durations, but the time interval between them is almost the same. Figure 4.14(b) shows the distribution of signal frequency, and the frequency range is from 30 to 400Hz. But Figure 4.14(c) shows that this sample Bryde's whale sound frequency range actually is around 90-170Hz, and the center frequency is about 120Hz. The call types of the signal in this audio contain down-sweep and tonal similar to those from the Gulf of Mexico (Rice et al., 2014b) and Be2, Be6 and Be10 similar to those from the Gulf of California and eastern tropical Pacific (Oieson et al., 2003)(Viloria-Gómora et al., 2015).

4.4 Bryde's whale sounds detection

Similar with the fin whale sound detection, a kernel will be synthesized from known signals, and then be applied to the LADC data.

4.4.1 Synthesize Bryde's whale matched filter kernel

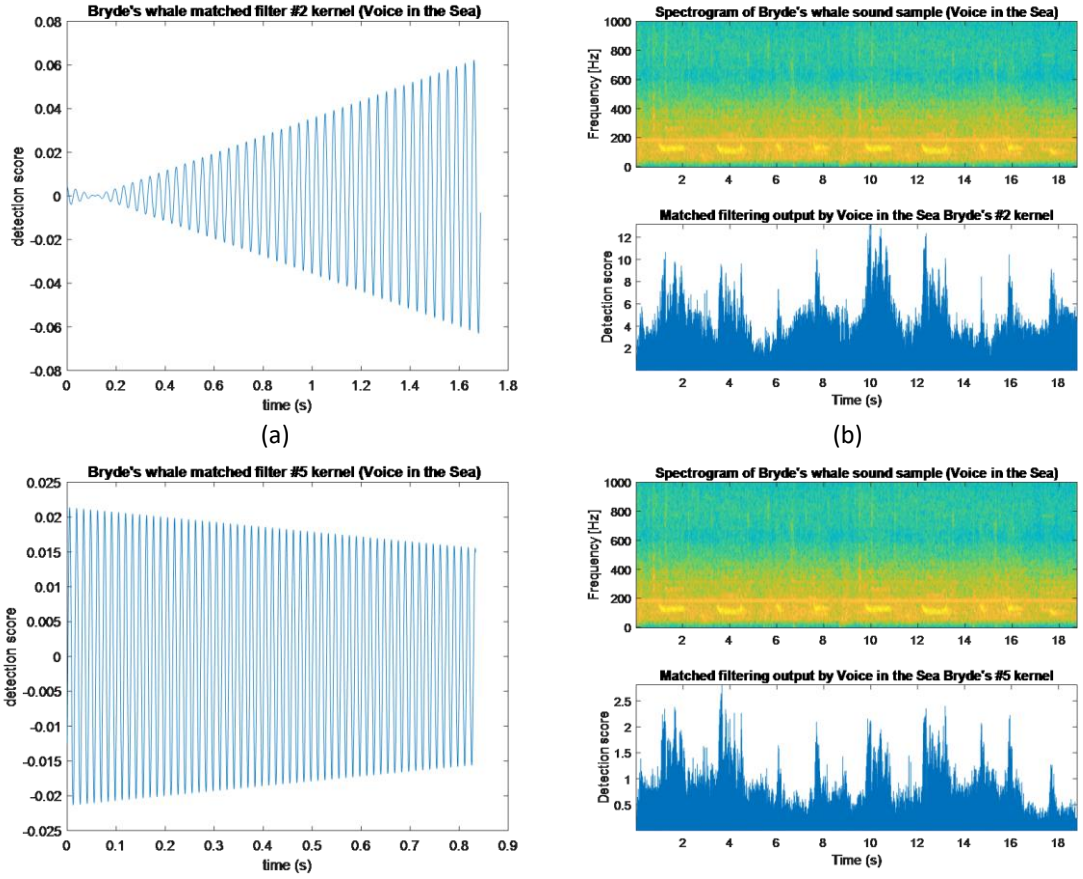
Single call was used to synthesize matched filter kernel, and similar to the fin whale matched filter synthesis, the information of initial and final frequencies, amplitude as well as the time duration are necessary to be collected.



(a) (b)

Figure 4.15 Bryde's whale (a)single call; (b)potential 9 kernels

Nine kernels were synthesized and applied to the original data where they came from. According to the test results, the 2nd and 5th kernels look better and can be used to find the Bryde's whale signal in LADC data.



(a) (b)

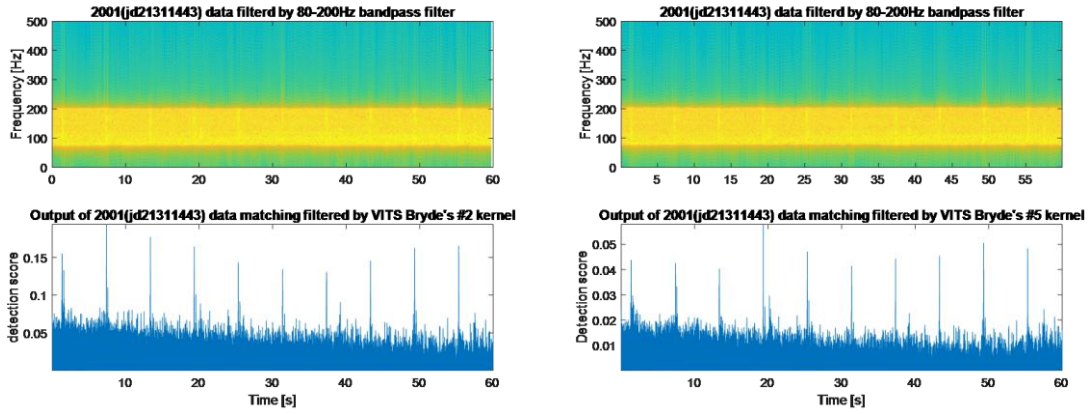
(c) (d)

Figure 4.16 Bryde's whale (a)#2 kernel in time domain; (b)#2 kernel matched filtering with Voice in the Sea data; (c)#5 kernel in time domain; (d)#5 kernel matched filtering with Voice in the Sea data

4.4.2 Bryde's whale matched filter kernel applied to Gulf of Mexico LADC data

According to older reference (Edds et al., 1993), the Bryde's whale call type in Gulf of Mexico were mainly "growl" (200-900Hz) and "discrete pulses" (400-610Hz) produced by calf, and "pulsed moan" (100-500Hz) produced by adult. However, in recent years reference (Rice et al., 2014b)(Širović et al., 2014), the Bryde's whale call type in Gulf of Mexico are mainly "down-sweep- sequences" (80-130Hz), "long-moans" (90-150Hz) and "tonal-sequences" (90-110Hz). Probably because of the ocean noise pollution, it seems that the Bryde's whale had to choose to reduce the frequency of their vocalizations in order to be able to transmit their sounds further in noise pollution. A similar reduction in the frequency of sound has been found in blue whale vocalizations over the last 60 years (McDonald et al., 2009).

Based on the frequency range of whale calls recorded in recent years in Gulf of Mexico, A bandpass filter(80-200Hz) was used to filter the data before the matched filtering operation.



(a) (b)

Figure 4.17 Bryde's whale (a)#2 kernel matched filtering with LADC 2001 sample data; (b)#5 kernel matched filtering with LADC 2001 sample data

The output figures show 10 peaks/spikes, and there are corresponding 10 bright vertical lines shown in spectrograms. According to the references, if they are potential Bryde's whale call, they are probably the "down-sweep sequences" once appeared in Gulf of Mexico in 2010, and also similar to the "down-sweep" type Be6 once appeared in eastern tropical Pacific. But the time interval between each call here is about 7s, which is within the duration range of growl call (0.5-50s), but longer than other call types.

The kernels were also used to apply to 2015 data.

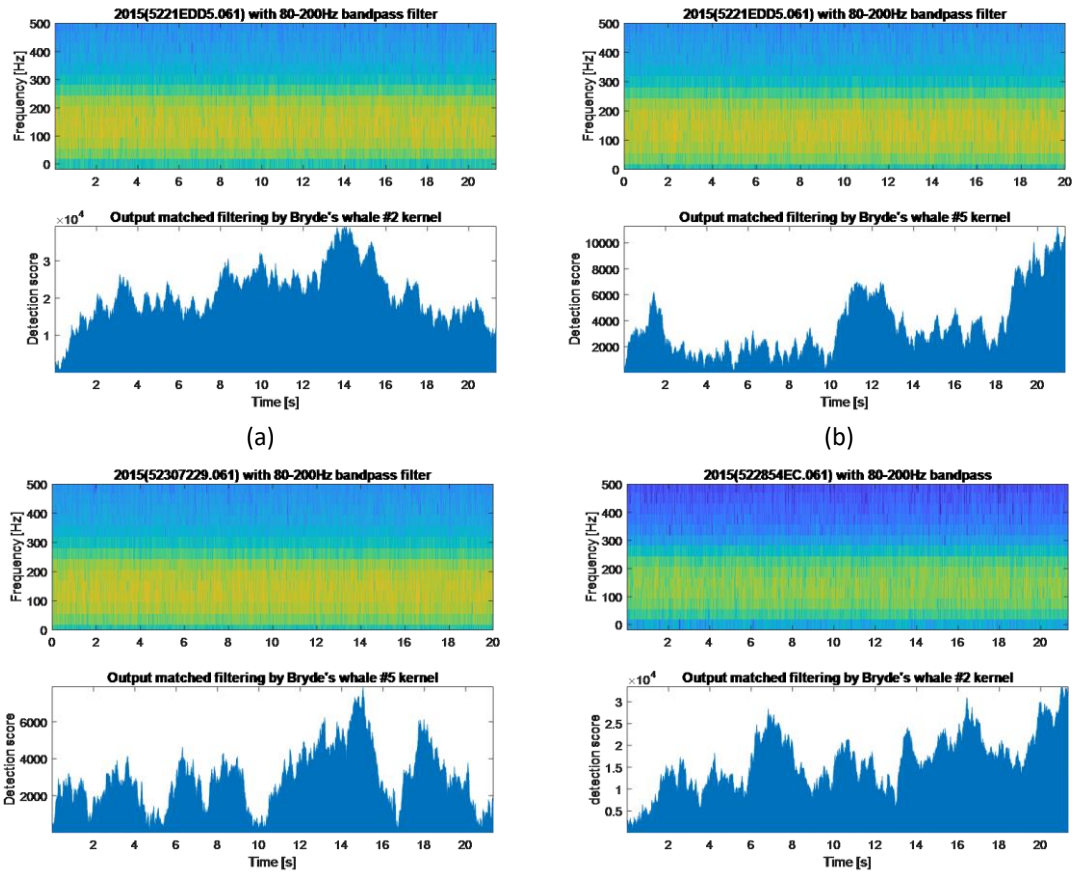


Figure 4.18 Bryde's whale (a)#2 kernel matched filtering 5221EDD5.061 sample data; (b)#5 kernel matched filtering 5221EDD5.061 sample data; (c)#5 kernel matched filtering 52307229.061 sample data; (d)#2 kernel matched filtering 522854EC.061 sample data

It looks like there are some spikes in output of matched filtering figures, but there is no sound signal(shape) similar to that of Bryde's whale calls in the corresponding time spectrogram.

Therefore, it is not yet possible to conclude that the sound signals of Bryde's whale have been found in the LADC data of 2001 and 2015.

4.5 Higher frequency whale sound recordings in the time domain

A one-minute unfiltered audio sample was recorded in the northern Gulf of Mexico in 2001 by LADC using EARS buoys which could record up to 5958 Hz (Ioup et al., 2016). It is easy to hear that this audio file includes noise and signals which sound like buzzes and clicks.

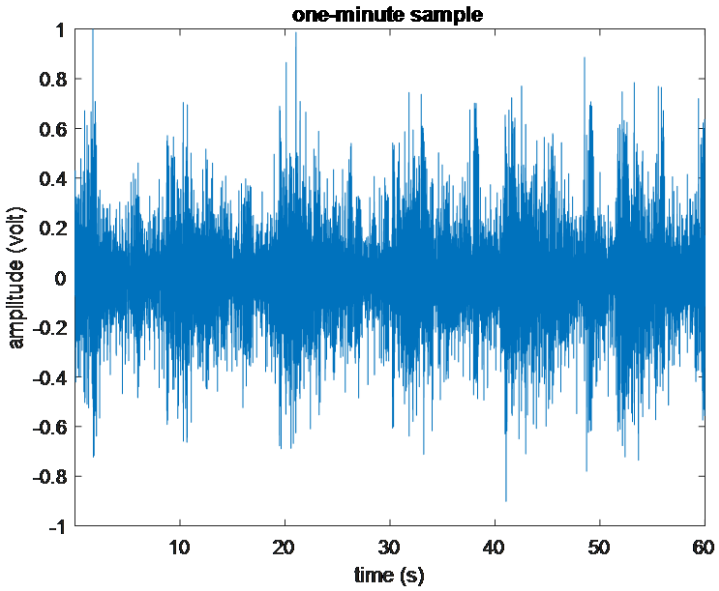
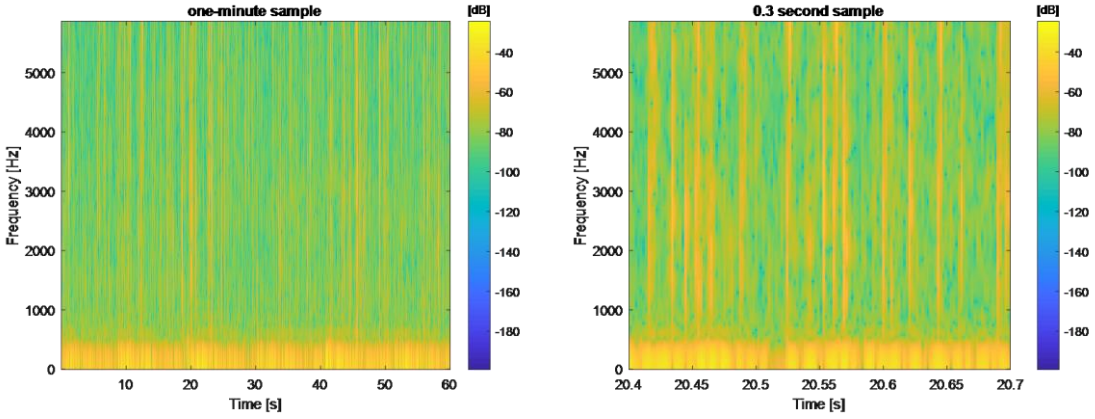


Figure 4.19 One-minute unfiltered sample

Figure 4.19 shows the original data for the one-minute unfiltered audio file without any detailed descriptions of the individual signals; there are many complex signals mixed together.



(a) (b)

Figure 4.20 Spectrogram of one-minute unfiltered sample (a)60 seconds sample; (b)0.3 second sample

Figure 4.20 shows the spectrogram of the unfiltered sample. The clicks shown on the right graph clearly have high frequencies up to 6kHz. Noise is also present in the graph, and its frequency is generally lower than 1000 Hz.

A 1000 Hz high-pass filter is applied as shown in Figure 4.21.

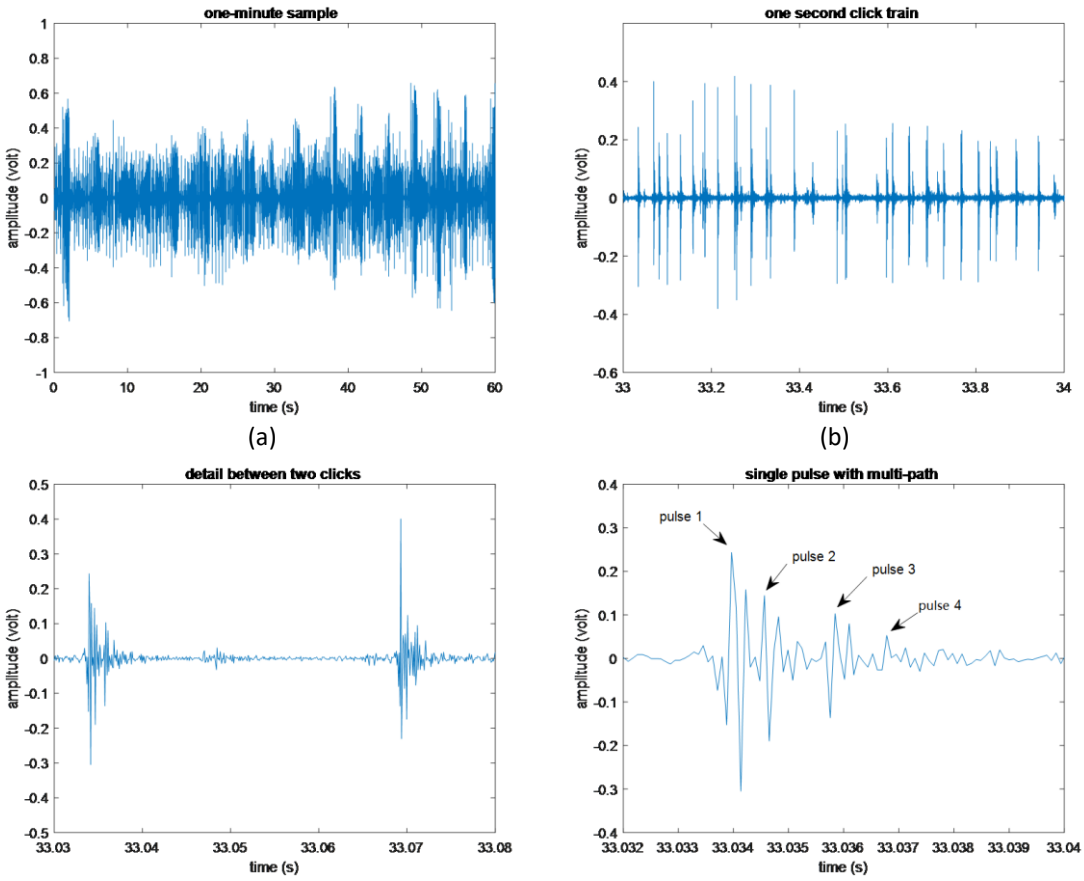


Figure 4.21 Detail of one-minute sample (a)one-minute sample with 1000Hz filter; (b)click train; (c)detail between two clicks; (d)single pulse with multi-path

Figure 4.21 shows some detail after a 1000 Hz high-pass filter is applied to the one-minute unfiltered sample. Comparing Figure 4.19 with Figure 4.21(a), the latter does seem to have less noise as the high-pass filter does remove frequencies below 1000 Hz.

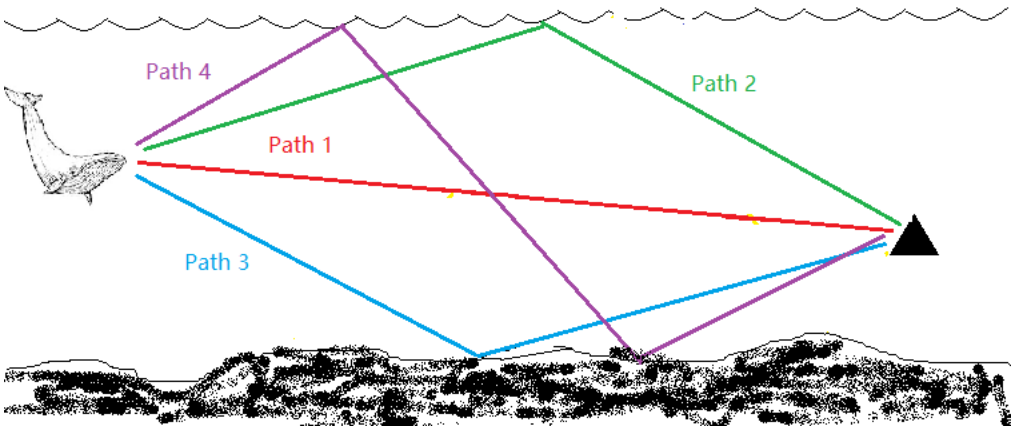
Figure 4.21(b) shows a representative click train very clearly. The time sequence in Figure 4.21(b) shows more than 30 clicks per second with varying amplitude, and there are some clicks with very short time intervals, almost overlapping. From this graph, it seems that the sound may be made by dolphins, whose frequency range and features are particularly similar to Amazon River dolphin. However, the Amazon River dolphin is a common species in South American countries and usually they do not swim freely into the Gulf of Mexico. Since the data collection includes many sperm whale's acoustic signals, so this sound probably produced by sperm whale. The overlapping clicks may also happen when more than one animal produces sounds at almost the same time. However, Figure 4.21(c) shows a single pulse followed by

multiple boundary reflections(Au, 1993). Some oscillating background noise higher than 1000 Hz appears between two single clicks, which may due to some other high-frequency noise such as from measuring or recording equipment, passing ships, and so on.

Figure 4.21(d) shows a short snippet of 8ms from Figure 4.21(b) and appears to be a single click. It includes a large amplitude pulse followed by a sequence of weaker pulses. The pulses are very short, lasting about 0.7ms with 2-4 oscillations. Meanwhile, the second pulse in Figure 4.21(d) is not only smaller in amplitude but also phase-inverted (Zimmer, 2011). In other words, the deflections of the first/main pulse are first negative and then positive. However, for the second pulse the deflections are first positive and then negative, etc. As we know, under certain circumstances, an interface may cause total reflection with phase shift. The phase of the third pulse is the same as the phase of the first/main pulse. Therefore, the original click is followed by the boundary reflection and the phase reversal may indicate a surface and bottom reflection.

According to the principle of optics, when light rays from a medium with a higher refractive index enter a medium with a lower refractive index, if the light incidence angle is less than the critical angel, the light is partially reflected, and the remaining sound energy enters the second medium as a transmitted wave. In this situation, the reflected wave does not phase shift because it is not totally reflected, but once the light incidence angle is equal or larger than the critical angle, that is with the onset of total reflection, a phase shift occurs in the reflected wave.

For a surface reflection, the interface is between the water and air; for a bottom reflection, the interface is between the water and the solid material of the bottom. The cartoon diagram below illustrates the geometry of the whale sound path with a direct path (red) and the first three multipaths (green, blue, purple) to the recorder.

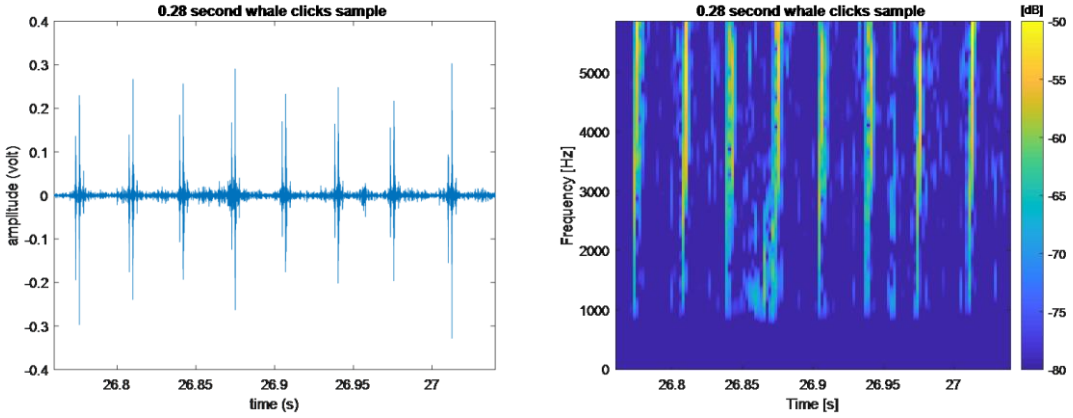


In the world of hydrography and marine science, backscatter is the reflection of a signal (such as sound waves or light waves) back in the direction from where it originated. Different bottom types “scatter” sound energy differently. Harder bottom types (like rock) reflect more sound than softer bottom types (like mud), and smoother bottom types (like pavement) reflect more sound than bumpier bottom types (like a coral reef)(NOAA, 2018).

The muddy bottom and the surface water will both scatter/reflect the sound produced by a whale at different level. The surface water totally reflects the incidence sound wave, so the sound phase is shifted. But the soft muddy bottom, which has a higher refractive index than water, does not totally reflect the incidence sound wave, so the phase of the sound wave passing through the mud is not reversed. Moreover, during the sound propagation, part of the sound energy will be absorbed by both mud and surface water, thus reducing the energy of the reflected sound, especially the soft bottoms mud is usually associated with high bottom losses (10 to 30 dB/bounce)(“PRINCIPLES OF UNDERWATER SOUND Chapter 8,” n.d.). This explains why the amplitude of the latter three pulses in Figure 4.21(d) gradually decreases.

Zooming in on the whole call train in Figure 4.21(b), the fact is not every main original click is followed by phase reversal reflections. The first reason may be that there is no total reflection. Another reason may be the presence of other vocal cetaceans. Furthermore, Figure 4.21 shows the signal from a particular animal; the particular animal is difficult to judge from the time domain only, and later frequency domain analysis may provide useful information.

As mentioned before, this audio does sound like it contains more than one kind of sound. It turns out that when zoomed in another time range of the sound sample, some new signal characteristics do appear, as shown in Figure 4.22.



(a) (b)

Figure 4.22 A series of whale clicks (a)time domain; (b)spectrogram

Figure 4.22 shows a click train produced by a whale having the same interval between each click. This may be a stereotyped click sequence produced by a sperm whale, called a coda. Sperm whale codas are often used to communicate between individuals, and the pattern of sperm whale codas is different in different geographical areas. In addition, Figure 4.22 shows that each click seems to contain a multi-path structure, and the difference in amplitude between successive clicks of the coda is small. The small amplitude difference is a characteristic of codas.

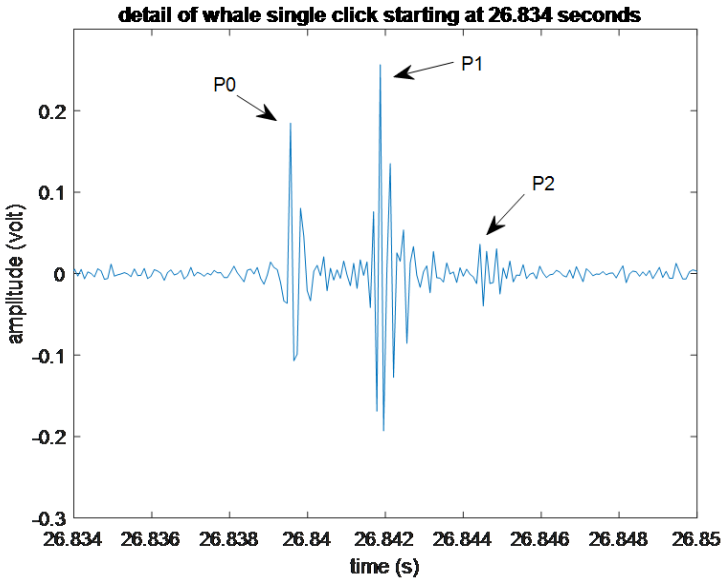


Figure 4.23 Detail of whale single click

Figure 4.23 is part of Figure 4.22. It shows a single click including three pulses separated by around 2.4ms. This multi-pulse structure is similar to that of sperm whales, and the inter-pulse interval (IPI) has a relationship with the length of the whale (Drouant and Ioup, 2017). When the IPI is shorter than 5ms for smaller whales, the formula (Gordon, 1987) for the whale length from the IPI is

$$L = 4.833 + 1.453(IPI) - 0.009(IPI)^2 \quad (4.5)$$

When the IPI is longer than 5ms for larger whales, the formula(Rhinelander and Dawson, 2004)for the whale length from the IPI is

$$L = 17.12 - 2.189(IPI) + 0.251(IPI)^2 \quad (4.6)$$

For an IPI of 2.4ms, the calculated length is 8.27m, which could be an immature sperm whale.

In contrast to Figure 4.21(d), the sequence of clicks produced because of reflection, the multi-pulse structure of Figure 4.23 is produced by the sperm whale itself. This structure has been called bent-horn sound generation (Zimmer et al., 2005). The bent-horn theory said "*a single pulse is generated below the blowhole at the tip of the nose. This original pulse is mainly backward-oriented and only some of the sound energy will leak directly into the water (called P0). Most of the sound energy will travel backwards towards the frontal air sac at the skull, where it is reflected forward. The major part of this forward reflected energy will leave the head through the junk, which is below the spermaceti organ, to form the powerful sonar pulse (called P1). The remaining part of the sound energy is reflected back inside the spermaceti organ, where it once again hits the forward boundary of the spermaceti organ (distal air sac) to be reflected again towards the skull and consequently result in secondary sonar pulses, P2, P3, etc., all with progressively decreasing intensity*." (Møhl, 2001)(Norris and Harvey, 1972). This bent-horn theory gives a biological reason to explain why a group of signals can contain several pulses.

The recording of the signal from different directions and orientations around the animal may produce different signal patterns. It is not clear which direction the data were recorded from the animal, but according to the bent-horn theory, the P1 should be the strongest pulse, the P0 should be the original pulse and the P2, P3 should be the weaker pulses. Therefore, in

the Figure 4.23, the left pulse could be the P0, the middle pulse could be the P1 and the right pulse could be the P2. This data sample probably was recorded in front of the animal (the whale is facing the EARS recorder) so that the receiver can first sense the original weak pulse P0, then the powerful pulse P1, and finally another weaker pulse P2 (Zimmer, 2011)

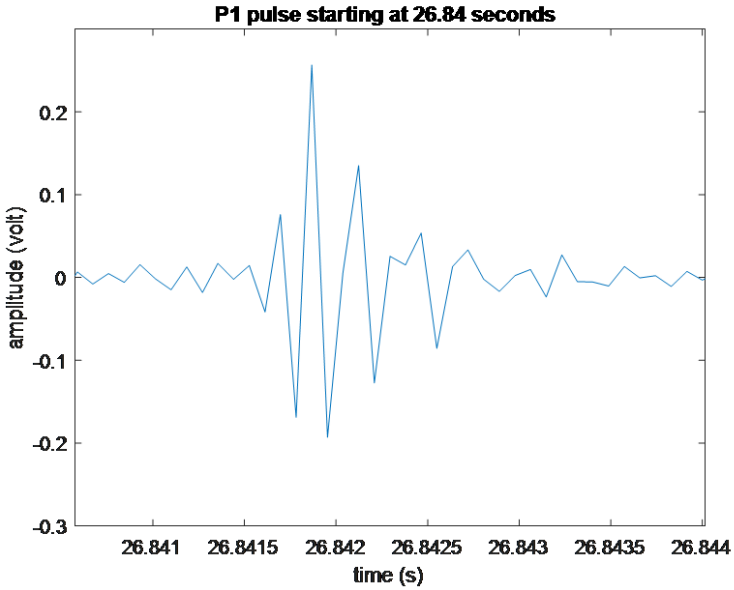


Figure 4.24 Detail of P1 pulse

Figure 4.24 is P1 pulse including about five oscillation lasting 1.5ms. The amplitude of the first two oscillations is higher and then decreases, similar to the sperm whale click.

4.6 Higher frequency whale sound recordings in the frequency domain

The time domain graphs show that the clicks of cetaceans are actually oscillating sound waves, and the frequency domain can give additional information about the frequency of these sound waves.

Figure 4.25 shows the Fourier Transform of a one-minute sample in the frequency domain. A 703250-point Fourier transform was used with frequencies below 1000Hz removed by a Rectangle or boxcar window in the frequency domain to reduce noise below 1000Hz.

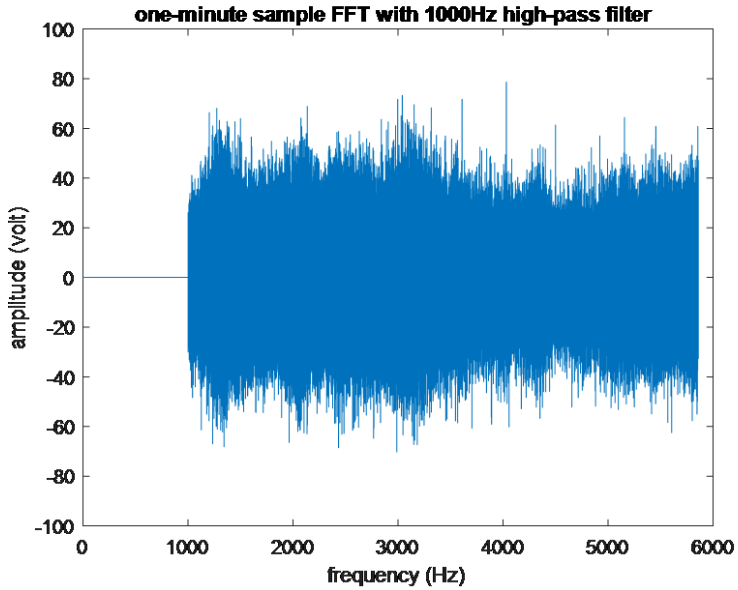
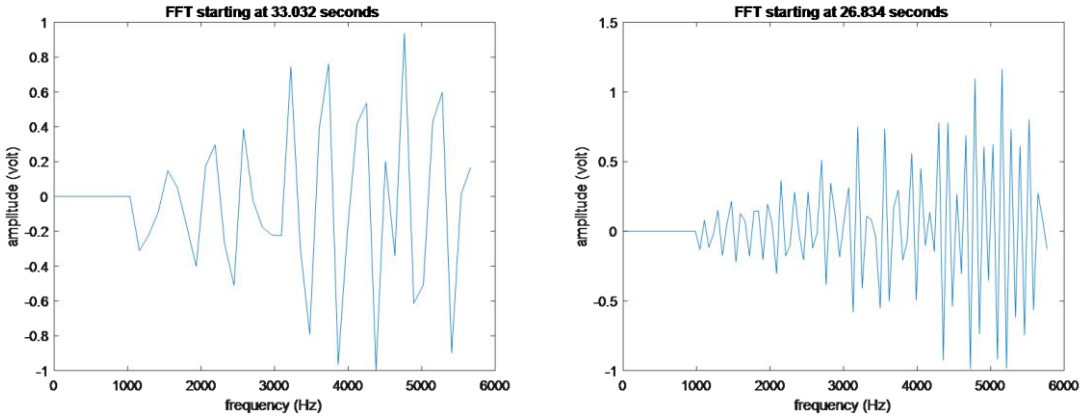


Figure 4.25 Fourier Transform of one-minute sample with 1000 Hz high-pass filter

The EARS receivers and recording equipment were unable to record signals above 6000Hz in 2001. The sounds produced by dolphins, which usually have high frequencies above 20kHz and up to 100 kHz would not be able to be recorded. It may not be correct to speculate the Figure 4.21(d) is a sound signal from a dolphin. In order to further determine what kind of animal produces the Figure 4.21(d) and Figure 4.23 signals, looking at the corresponding time signal in the frequency domain is very necessary, as described below.



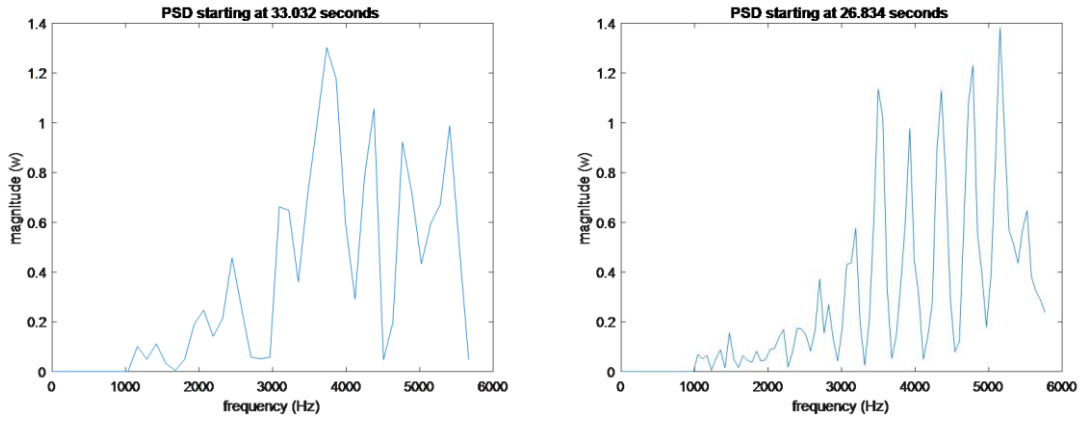
(a) (b)

Figure 4.26 Fourier Transform fragment (a) FFT from 33.032 to 33.04 seconds; (b) FFT from 26.834 to 26.85 seconds

Fourier analysis can show the most likely frequency from the largest peak in the Fourier spectrum. Figure 4.26 shows two frequency distributions from different times in the one-minute sample of Figure 4.25, Figure 4.21(d) and Figure 4.23 show that the average time

intervals between two consecutive oscillation is 0.0002 seconds, which means the largest signal frequency should be about 5000Hz, and it does show that the highest amplitude appears around 4800-5200Hz in Figure 4.26. This is toward the lower end of the normal range of dolphin frequencies.

The power spectrum of a time series describes the distribution of power of the frequency components composing that signal (Wikipedia contributors, 2019f), and is the square magnitude of the Fourier Transform. The power spectrum (power spectral density (PSD)) can thus give additional frequency information in the two-time ranges of interest.



(a) (b)

Figure 4.27 Power spectrum density fragment (a) PSD from 33.032 to 33.04 seconds; (b) PSD from 26.834 to 26.85 seconds

The amplitude of the frequency component is proportional to the signal intensity at that frequency, which in turn is proportional to the power carried by the signal(Giordano, 1997). Figure 4.27(a) shows the signal intensity or PSD at a frequency around 3800 Hz is very strong, which means that at this particular sampling frequency the signal carries the highest power at 3800Hz. Figure 4.27(b) shows the corresponding frequency carrying the highest power is around 5200Hz. Figure 4.27(b) shows the frequency that carries the highest signal power is the largest frequency in Figure 4.26(b) but Figure 4.27(a) and Figure 4.26(a) do not reflect this corresponding feature. This further uncertainty adds to the difficulty of determining what animal sounds are in Figure 4.21(d).

Therefore, the signal in Figure 4.23 could be thought to have been produced by a sperm whale, and the signal in Figure 4.21(d) could have been also produced by sperm whale or another indeterminate toothed whale.

4.7 Higher frequency whale sound recordings in the time-frequency domain

For the sperm whale single pulse in Figure 4.24 for example, the time and frequency range can be presented clearly in a spectrogram.

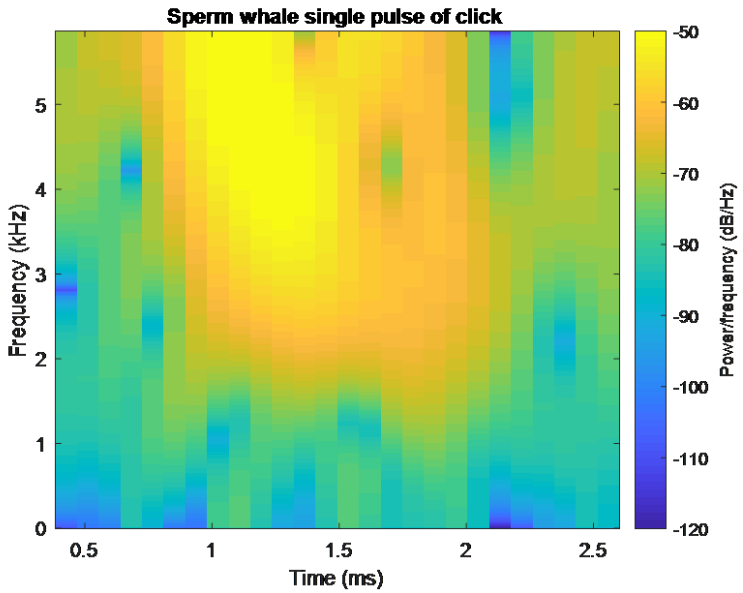


Figure 4.28 Spectrogram of sperm whale single pulse of click

The spectrogram shows signal amplitude versus both time and frequency. Figure 4.28 shows the spectrogram of the same 3 milliseconds as in Figure 4.24 (35 sample points of data). The main pulse of the click of a sperm whale occurs in a time range of 1 to 1.5ms with center frequency around 4250Hz, and the frequency range from 3000Hz to 5500Hz.

Chapter 5. Conclusions

There are a large number of marine mammals living in the Gulf of Mexico. The LADC database contains sound signals from many of them. After bandpass filtering low-frequency noise, the signals of the toothed whales are easily detected, and their sound characteristics are clearly expressed.

One characteristic of sperm whale clicks is that many clicks sequences appear, some of which can be called coda. The special construction of the sperm whale skulls gives the individual pulses of their sound signals distinct structural features.

It is not as easy to find the sound signal of the baleen whale, which has a low frequency signal. Their sounds are submerged in a lot of low frequency noise, so it is important to use the proper detection method for signals with low signal-to-noise ratio.

This paper uses the matched filter method to try to detect the sound of fin whales and Bryde's whale. Some low-frequency sound signals have been found in the 2001 LADC database, but it is not yet fully certain that they are the signals of fin whale or Bryde's whale.

In the future, other detection methods will be tried to continue to search for whale signals in the Gulf of Mexico LADC database, for example, using the spectrogram correlation and neural network methods. The latter two methods have a lower error rate than matched filter.

References

7 The Mathematical Microscope: Waves, Wavelets, and Beyond | A Positron Named Priscilla: Scientific Discovery at the Frontier | The National Academies Press, (n.d.). Available: https://www.nap.edu/read/2110/chapter/9, (date last viewed: 23-Oct-19). Retrieved October 23, 2019, from https://www.nap.edu/read/2110/chapter/9

Allen, J. (**1977**). "Short term spectral analysis, synthesis, and modification by discrete Fourier transform," IEEE Trans. Acoust., **25**, 235–238. doi:10.1109/TASSP.1977.1162950

Allen, J. B., and Rabiner, L. R. (**1977**). "A Unified Approach to Short-Time Fourier Analysis and Synthesis," Proc. IEEE, **65**, 1558–1564. doi:10.1109/PROC.1977.10770

Andrew, R. K., Howe, B. M., Mercer, J. A., and Dzieciuch, M. A. (**2002**). "Ocean ambient sound: Comparing the 1960s with the 1990s for a receiver off the California coast," Acoust. Res. Lett. Online, **3**, 65–70. doi:10.1121/1.1461915

Au, W. W. L. (**1993**). "Characteristics of Dolphin Sonar Signals," The Sonar of Dolphins, Springer New York, New York, NY, pp. 115–139. doi:10.1007/978-1-4612-4356-4_7

Balsam, W. L., and Beeson, J. P. (**2003**). "Sea-floor sediment distribution in the Gulf of Mexico," Deep Sea Res. Part I Oceanogr. Res. Pap., **50**, 1421–1444. doi:10.1016/j.dsr.2003.06.001

Bracewell, R. (**2000**). *The Fourier Transform And Its Applications*, McGraw-Hill Companies, Singapore.

Brigham, E. O. (**1988**). *The Fast Fourier Transform and Its Applications*, Prentice Hall,.

Burger, J., Overstreet, R. M., Chen, Y., Roberts, K. J., Hawkins, W. E., Valverde, R. A., Holzwart, K. R., et al. (**2017**). *Habitats and Biota of the Gulf of Mexico: Before the Deepwater Horizon Oil Spill*, (C. H. Ward, Ed.) Springer New York, New York, NY, Vol. 2. doi:10.1007/978-1-4939-3456-0

Castellote, M., Clark, C. W., and Lammers, M. O. (**2012**). "Fin whale (Balaenoptera physalus) population identity in the western Mediterranean Sea," Mar. Mammal Sci., **28**, 325–344. doi:10.1111/j.1748-7692.2011.00491.x

Croll, D. A., Clark, C. W., Acevedo, A., Tershy, B., Flores, S., Gedamke, J., and Urban, J. (**2002**). "Only male fin whales sing loud songs," Nature, **417**, 809–809. doi:10.1038/417809a

Davis, R. W., Ortega-Ortiz, J. G., Ribic, C. A., Evans, W. E., Biggs, D. C., Ressler, P. H., Cady, R. B., et al. (**2002**). "Cetacean habitat in the northern oceanic Gulf of Mexico," Deep. Res. Part I Oceanogr. Res. Pap., **49**, 121–142. doi:10.1016/S0967-0637(01)00035-8

Drouant, G., and Ioup, J. W. (**2017**). "Calculating sperm whale lengths in the Northern Gulf of Mexico," J. Acoust. Soc. Am., **142**, 2561–2562. doi:10.1121/1.5014367

Edds, P. L., Odell, D. K., and Tershy, B. R. (**1993**). "Vocalizations of a Captive Juvenile and Free-Ranging Adult-Calf Pairs of Bryde'S Whales, Balaenoptera Edeni," Mar. Mammal Sci., **9**, 269–284. doi:10.1111/j.1748-7692.1993.tb00455.x

Giordano, N. J. (**1997**). *Computational Physics*, Prentice Hall, Upper Saddle River.

Google (**2019**). *Google Maps*, Available: https://www.google.com/maps/dir///@23.1835253,-86.5707854,1915186m/data=!3m1!1e3!4m2!4m1!3e0, (date last viewed: 01-Aug-19). Retrieved August 1, 2019, from https://www.google.com/maps/dir///@23.1835253,-86.5707854,1915186m/data=!3m1!1e3!4m2!4m1!3e0

Gordon, J. C. D. (**1987**). *The behaviour and ecology of sperm whales off Sri Lanka.* University of Cambridge.

Harris, F. J. (**1978**). “On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform,” Proc. IEEE, **66**, 51–83. doi:10.1109/PROC.1978.10837

Haykin, S. (**2001**). *Communication Systems*, John Wiley & Sons, INC.

Heimlich, S. L., Mellinger, D. K., Nieukirk, S. L., and Fox, C. G. (**2005**). “Types, distribution, and seasonal occurrence of sounds attributed to Bryde’s whales (*Balaenoptera edeni*) recorded in the eastern tropical Pacific, 1999–2001,” J. Acoust. Soc. Am., **118**, 1830–1837. doi:10.1121/1.1992674

Higgins, R. J. (**1976**). “Fast Fourier transform: An introduction with some minicomputer experiments,” Am. J. Phys., **44**, 766. doi:10.1119/1.10128

Ioup, G. E., Ioup, J. W., Sidorovskaia, N. A., Tiemann, C. O., Kuczaj, S. A., Ackleh, A. S., Newcomb, J. J., et al. (**2016**). “Environmental Acoustic Recording System (EARS) in the Gulf of Mexico,” pp. 117–162. doi:10.1007/978-1-4939-3176-7_6

Jaquet, N., and Gendron, D. (**2009**). “The social organization of sperm whales in the gulf of California and comparisons with other populations,” J. Mar. Biol. Assoc. United Kingdom, **89**, 975–983. doi:10.1017/S0025315409001507

Kay, S. M. (**1993**). *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory (v. 1)*, Prentice Hall, Upper Saddle River. Retrieved from http://users.isr.ist.utl.pt/~pjcro/temp/Fundamentals Of Statistical Signal Processing--Estimation Theory-Kay.pdf

LaBrecque, E., Curtice, C., Harrison, J., Van Parijs, S. M., and Halpin, P. N. (**2015**). “Biologically important areas for cetaceans within US waters - Gulf of Mexico region,” Aquat. Mamm., **41**, 30–38. doi:10.1578/AM.41.1.2015.30

LADC-GEMM (**2017**). *RESEARCH – LADC-GEMM*, Available: http://www.ladcgemm.org/research-2/, (date last viewed: 01-Aug-19). Retrieved August 1, 2019, from http://www.ladcgemm.org/research-2/

Van Loan, C. (**1992**). *Computational Frameworks for the Fast Fourier Transform*, Front. Appl. Math., Society for Industrial and Applied Mathematics, 287 pages. doi:doi:10.1137/1.9781611970999

Lockyer, C. (**1976**). “Body weights of some species of large whales,” ICES J. Mar. Sci., **36**, 259–273. doi:10.1093/icesjms/36.3.259

Love, M. S., Baldera, A., Yeung, C., and Robbins, C. (**2013**). *The Gulf of Mexico Ecosystem- a coastal and marine atlas* New Orleans.

McDonald, M., Hildebrand, J., and Mesnick, S. (**2009**). “Worldwide decline in tonal frequencies of blue whale songs,” Endanger. Species Res., **9**, 13–21. doi:10.3354/esr00217

Mellinger, D. K., and Clark, C. W. (**1997**). “Methods for automatic detection of mysticete sounds,” Mar. Freshw. Behav. Physiol., **29**, 163–181. doi:10.1080/10236249709379005

Mellinger, D. K., Stafford, K. M., Moore, S. E., Dziak, R. P., and Matsumoto, H. (**2007**). “an overview of fixed passive acoustic observation methods for cetaceans,” Oceanography, **20**, 36–45. Retrieved from https://tos.org/oceanography/assets/docs/20-4_mellinger.pdf

Miller, S. L., and Childers, D. (**2012**). *Probability and Random Processes*, Elsevier Inc., Elsevier. Retrieved from https://linkinghub.elsevier.com/retrieve/pii/B9780123869814500217

Møhl, B. (**2001**). “Sound transmission in the nose of the sperm whale Physeter catodon A post mortem study,” J. Comp. Physiol. A Sensory, Neural, Behav. Physiol., **187**, 335–340. doi:10.1007/s003590100205

Morano, J. L., Salisbury, D. P., Rice, A. N., Conklin, K. L., Falk, K. L., and Clark, C. W. (**2012**). "Seasonal and geographical patterns of fin whale song in the western North Atlantic Ocean," J. Acoust. Soc. Am., **132**, 1207–1212. doi:10.1121/1.4730890

National Instruments (**2008**). *STFT Spectrograms VI - NI LabVIEW 8.6 Help*, Available: http://zone.ni.com/reference/en-XX/help/371361E-01/lvanls/stft_spectrogram_core/, (date last viewed: 01-Aug-19). Retrieved August 1, 2019, from http://zone.ni.com/reference/en-XX/help/371361E-01/lvanls/stft_spectrogram_core/

National Instruments (**2012**). *Characteristics of Different Smoothing Windows - LabWindows/CVI 2012 Help - National Instruments*, Available: http://zone.ni.com/reference/en-XX/help/370051V-01/cvi/libref/analysisconcepts/characteristics_of_different_smoothing_windows/, (date last viewed: 01-Aug-19). Retrieved August 1, 2019, from http://zone.ni.com/reference/en-XX/help/370051V-01/cvi/libref/analysisconcepts/characteristics_of_different_smoothing_windows/

National Instruments (**2019**). *Understanding FFTs and Windowing* Instrum. Fundam., Instrument Fundamentals,Austin. Retrieved from https://www.ni.com/en-us/innovations/white-papers/06/understanding-ffts-and-windowing.html

National Marine Fisheries Service (**2010**). *Final Recovery Plan for the Fin Whale* Natl. Mar. Fish. Serv.,Silver Spring, 121 pages. Retrieved from https://repository.library.noaa.gov/view/noaa/4952

NOAA (**2018**). *How does backscatter help us understand the sea floor?*, Available: https://oceanservice.noaa.gov/facts/backscatter.html, (date last viewed: 01-Aug-19). Retrieved August 1, 2019, from https://oceanservice.noaa.gov/facts/backscatter.html

NOAA (**2019**). *Gulf of Mexico Data Atlas*, Available: https://www.ncddc.noaa.gov/website/DataAtlas/atlas.htm?plate=Bathymetry - Gulf, (date last viewed: 01-Aug-19). Retrieved August 1, 2019, from https://www.ncddc.noaa.gov/website/DataAtlas/atlas.htm?plate=Bathymetry - Gulf

NOAA (**2019**). *Gulf of Mexico Bryde's Whale | NOAA Fisheries*, Available: https://www.fisheries.noaa.gov/species/gulf-mexico-brydes-whale#overview, (date last viewed: 01-Aug-19). Retrieved August 1, 2019, from https://www.fisheries.noaa.gov/species/gulf-mexico-brydes-whale#overview

NOAA (**2019**). *NOAA Sounds in the Ocean*, Available: https://www.nefsc.noaa.gov/psb/acoustics/sounds.html, (date last viewed: 01-Aug-19). Retrieved August 1, 2019, from https://www.nefsc.noaa.gov/psb/acoustics/sounds.html

Noise, O. (**2003**). *Ocean Noise and Marine Mammals*, National Academies Press, Washington, D.C., 220 pages. doi:10.17226/10564

Norris, K. S., and Harvey, G. W. (**1972**). "Theory for the Function of spermaceti organ of the sperm whale," Anim. Orientat. Navig.,.

Oieson, E. M., Barlow, J., Gordon, J., Rankin, S., and Hildebrand, J. A. (**2003**). "Low Frequency Calls of Bryde'S Whales," Mar. Mammal Sci., **19**, 407–419. doi:10.1111/j.1748-7692.2003.tb01119.x

Oppenheim, A. V., Schafer, R. W., and Buck, J. B. (**1998**). *Discrete Time Signal Processing*, Prentice Hall, Upper Saddle River.

OSC (**2019**). *Passive Acoustic Monitoring (PAM)*, Available:

http://www.passiveacousticmonitoring.com/, (date last viewed: 01-Aug-19). Retrieved August 1, 2019, from http://www.passiveacousticmonitoring.com/

Petre, S., and Moses, R. (**2005**). *Spectral Analysis of Signals*, Prentice Hall, Prentice Hall, Upper Saddle River. Retrieved from http://ieeexplore.ieee.org/document/4049928/

PRINCIPLES OF UNDERWATER SOUND Chapter 8, (n.d.). Available: https://fas.org/man/dod-101/navy/docs/fun/part08.htm, (date last viewed: 24-Oct-19). Retrieved October 24, 2019, from https://fas.org/man/dod-101/navy/docs/fun/part08.htm

Rhinelander, M. Q., and Dawson, S. M. (**2004**). "Measuring sperm whales from their clicks: Stability of interpulse intervals and validation that they indicate whale length," J. Acoust. Soc. Am., **115**, 1826–1831. doi:10.1121/1.1689346

Rice, A. N., Palmer, K. J., Tielens, J. T., Muirhead, C. A., and Clark, C. W. (**2014**). " Potential Bryde's whale (Balaenoptera edeni) calls recorded in the northern Gulf of Mexico ," J. Acoust. Soc. Am., **135**, 3066–3076. doi:10.1121/1.4870057

Rice, A. N., Palmer, K. J., Tielens, J. T., Muirhead, C. A., and Clark, C. W. (**2014**). "Potential Bryde's whale (Balaenoptera edeni) calls recorded in the northern Gulf of Mexico," J. Acoust. Soc. Am., **135**, 3066–3076. doi:10.1121/1.4870057

Richter, C., Gordon, J., Jaquet, N., and WÜrsig, B. (**2008**). "Social structure of sperm whales in the northern Gulf of Mexico," Gulf Mex. Sci., **26**, 118–123. doi:10.18785/goms.2602.03

Sciacca, V., Caruso, F., Beranzoli, L., Chierici, F., De Domenico, E., Embriaco, D., Favali, P., et al. (**2015**). "Annual acoustic presence of fin whale (Balaenoptera physalus) offshore Eastern Sicily, central Mediterranean Sea," PLoS One, , doi: 10.1371/journal.pone.0141838. doi:10.1371/journal.pone.0141838

Sidorovskaia, N., Comeaux, D., Greenhow, D., Griffin, S., Heimlich, S., Pierpoint, C., Richter, C., et al. (**2017**). *PROCEEDINGS OF THE LADC-GEMM 2017 GULF OF MEXICO EXPERIMENT*.

Širović, A., Bassett, H. R., Johnson, S. C., Wiggins, S. M., and Hildebrand, J. A. (**2014**). "Bryde's whale calls recorded in the Gulf of Mexico," Mar. Mammal Sci., **30**, 399–409. doi:10.1111/mms.12036

Smith, J. O. (**2011**). "Spectral Audio Signal Processing," Cent. Comput. Res. Music Acoust.,. Retrieved from https://ccrma.stanford.edu/~jos/sasp/%5Cnhttp://books.w3k.org/

Soule, D. C., and Wilcock, W. S. D. (**2013**). "Fin whale tracks recorded by a seismic network on the Juan de Fuca Ridge, Northeast Pacific Ocean," J. Acoust. Soc. Am., **133**, 1751–1761. doi:10.1121/1.4774275

Strang, G., and Nguyen, T. (**1996**). *Wavelets and Filter Banks*, Wellesley College.

Studds, G. E., and Wright, A. J. (**2007**). "A Brief Review of Anthropogenic Sound in the Oceans Publication Date," Int. J. Comp. Psychol., Retrieved from https://escholarship.org/content/qt5cj6s4r9/qt5cj6s4r9.pdf. Retrieved from https://escholarship.org/content/qt5cj6s4r9/qt5cj6s4r9.pdf

Tan, X. (**2017**). *怎样用通俗易懂的方式解释窗函数？ - 知乎*, Available: https://www.zhihu.com/question/50402321, (date last viewed: 01-Aug-19). Retrieved August 1, 2019, from https://www.zhihu.com/question/50402321

Van Trees, H. L. (**2001**). *Detection, Estimation, and Modulation Theory, Part I*, J. Chem. Inf. Model., John Wiley & Sons, Inc., New York, USA, Vol. 53, 1689–1699 pages. doi:10.1002/0471221082

Turin, G. (**1960**). "An introduction to matched filters," IEEE Trans. Inf. Theory, **6**, 311–329.

doi:10.1109/TIT.1960.1057571
Viloria-Gómora, L., Romero-Vivas, E., and Urbán R., J. (**2015**). “Calls of Bryde’s whale (Balaenoptera edeni) recorded in the Gulf of California,” J. Acoust. Soc. Am., **138**, 2722–2725. doi:10.1121/1.4932032
Voices In The Sea (**2018**). *Fin Whale : Baleen Whales : Voices in the Sea*, Available: http://cetus.ucsd.edu/voicesinthesea_org/species/baleenWhales/fin.html, (date last viewed: 01-Aug-19). Retrieved August 1, 2019, from http://cetus.ucsd.edu/voicesinthesea_org/species/baleenWhales/fin.html
Wahlberg, M. (**2002**). “The acoustic behaviour of diving sperm whales observed with a hydrophone array,” J. Exp. Mar. Bio. Ecol., **281**, 53–62. doi:10.1016/S0022-0981(02)00411-2
Weirathmueller, M. J., Stafford, K. M., Wilcock, W. S. D., Hilmo, R. S., Dziak, R. P., and Tréhu, A. M. (**2017**). “Spatial and temporal trends in fin whale vocalizations recorded in the NE Pacific Ocean between 2003-2013,” PLoS One, **12**, 1–24. doi:10.1371/journal.pone.0186127
Weisstein, E. W. (**2003**). *CRC CONCISE ENCYCLOPEDIA OF MATHEMATICS*, CRC Press.
Wenz, G. M. (**1962**). “Acoustic Ambient Noise in the Ocean: Spectra and Sources,” J. Acoust. Soc. Am., **34**, 1936–1956. doi:10.1121/1.1909155
Wikipedia contributors (**2019**). “Sperm Whale,” Wikipedia,. Retrieved from https://en.wikipedia.org/wiki/Sperm_whale
Wikipedia contributors (**2019**). “Fin whale,” Wikipedia,. Retrieved from https://en.wikipedia.org/wiki/Fin_whale
Wikipedia contributors (**2019**). “Window function,” Wikipedia,. Retrieved from https://en.wikipedia.org/wiki/Window_function
Wikipedia contributors (**2019**). “Matched Filter,” Wikipedia,. Retrieved from https://en.wikipedia.org/wiki/Matched_filter
Wikipedia contributors (**2019**). “Conjugate transpose,” Wikipedia,. Retrieved from https://en.wikipedia.org/wiki/Conjugate_transpose
Wikipedia contributors (**2019**). “Spectral density,” Wikipedia,. Retrieved from https://en.wikipedia.org/wiki/Spectral_density
Würsig, B. (**2017**). “Marine Mammals of the Gulf of Mexico,” Habitats Biota Gulf Mex. Before Deep. Horiz. Oil Spill, Springer New York, New York, NY, Vol. 2, pp. 1489–1587. doi:10.1007/978-1-4939-3456-0_5
Zimmer, W. M. . (**2011**). *passive acoustic monitoring of cetaceans*, Cambridge University Press, New York, NY.
Zimmer, W. M. X., Tyack, P. L., Johnson, M. P., and Madsen, P. T. (**2005**). “Three-dimensional beam pattern of regular sperm whale clicks confirms bent-horn hypothesis,” J. Acoust. Soc. Am., **117**, 1473–1485. doi:10.1121/1.1828501

VITA

Yingxue Gao was born in Chengdu, Sichuan, China. In 2004, she attended Southwest University of Science and Technology in Mianyang, China, where she received her Bachelor's and Master's degree in Optical Information Science and Technology and Condensed Matter Physics in 2008 and 2011, respectively. She joined the Master's program of Applied Physics of Department of Physics, University of New Orleans, in 2018. During her Master's program, she worked on acoustic signal processing of marine mammals in the Gulf of Mexico.