

Fall 12-20-2017

Underwater Acoustic Signal Analysis Toolkit

Kirk Bienvenu Jr

University of New Orleans, New Orleans, kdbienve@uno.edu

Follow this and additional works at: <https://scholarworks.uno.edu/td>



Part of the [Numerical Analysis and Scientific Computing Commons](#), [Other Applied Mathematics Commons](#), [Other Physics Commons](#), [Signal Processing Commons](#), and the [Software Engineering Commons](#)

Recommended Citation

Bienvenu, Kirk Jr, "Underwater Acoustic Signal Analysis Toolkit" (2017). *University of New Orleans Theses and Dissertations*. 2398.

<https://scholarworks.uno.edu/td/2398>

This Thesis is protected by copyright and/or related rights. It has been brought to you by ScholarWorks@UNO with permission from the rights-holder(s). You are free to use this Thesis in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you need to obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/or on the work itself.

This Thesis has been accepted for inclusion in University of New Orleans Theses and Dissertations by an authorized administrator of ScholarWorks@UNO. For more information, please contact scholarworks@uno.edu.

Underwater Acoustic Signal Analysis Toolkit

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

Kirk Douglas Bienvenu, Jr.
B.S. Aeronautics, Embry-Riddle Aeronautical University, 2013
December, 2017

Acknowledgements

First and foremost, I want to thank my wife, Chelsea. You have always been my true inspiration and the driving force behind my pursuit to better myself. To you, and the rest of my family and friends, thank you all for being so patient with me over the last several years. I know it has been a long time with many ups and down, but you have never shown even the slightest impression of doubt. You stuck by me through past hardships and reassured me that there would be more opportunities, if only I continued to improve myself. You always knew that I would find my way. It was a process to say the least, and I am so thankful for your love and support. None of this would have been possible without you.

I would like to thank Dr. Juliette loup for the opportunities she has afforded me. My family thanks you for taking me under your vast wings and showing me a new world of opportunity. Without you, the last few years of my life would have been far more difficult, both financially and academically. I have been searching my entire life for a calling, and your passion for studying the physical world has most certainly taken root in me. I will do my best to carry the lessons you have taught me and aid others in their pursuit of knowledge.

There are so many classmates and colleagues I would like to thank as well, but there is always two of you that come specially to mind. Matt Firreno, thank you for the last four years, your friendship has made this time so much more enjoyable. I will always remember the bourbon fueled late night study sessions. You really have kept me on the straight and narrow and I'd like to believe that some of that laser focus you have with regard to your studies has rubbed off on me, and maybe some of the insomnia as well. Britt Aguda, I have learned much from you, and wasted equal amounts of time arguing with you as well. I have become a much more accomplished physicist and programmer than I had ever thought possible, and I owe that in large part to you. All readers should be aware that much of the code written for this project was supplied either by Britt, or with his aid. I could never have come so far without your special brand of burst work.

I have had many professors over the years, but those of special note who have prepared me so well for my next stage in life deserve to be recognized. I was first introduced to the wonderful world of calculus by Mr. Steven Shalit. Always a stickler for details, I have been commended for my understanding of fundamental calculus, and every bit of that is thanks to you. Dr. Joel Webb, your lectures were never dull, to say the least. You have a way of capturing the students' interest and refining their talents. Dr. Gregory Seab and Dr. Kevin Stokes, you have both been excellent sources of support during my academic career since my very first day in PHYS 1061. Thank you for your counsel and thank you for stealing me away from the engineering department. Dr. Linxiong Li, much of my research relies on large data sets and your statistics course was vital to giving me the knowledge necessary to understand much of the higher level published literature in signal analysis. I was a physics student lost in a sea of mathematicians, but you helped me stay right along with the rest of them. Thank you for understanding and guidance.

Dr. Chris Tiemann, we have only met a couple of times, but I know your work very well. Thank you for your help in my research. It should also be noted that much of my understanding of MATLAB programming was developed by studying and reverse engineering your code.

The work presented here was developed over the course of a year and a half under the Littoral Acoustic Demonstration Center for Gulf Ecological Monitoring and Modeling (LADC-GEMM) oversight with funding by the Gulf of Mexico Research Initiative (GOMRI). The course work required to understand the theory and applications of such an undertaking was all made possible by these organizations and their personnel. Thank you for the opportunity.

Table of Contents

List of Figures.....	iv
Abstract	v
Chapter 1: Introduction.....	1
1.1 History of LADC-GEMM	1
1.2 The Data	1
Chapter 2: Applications of Transform in Signal Analysis	4
2.1 Bioacoustics	4
2.2 The Fourier Transform.....	5
2.3 Power Spectral Density	6
2.4 Bandpass Filtering.....	7
2.5 The Wavelet Transform	8
Chapter 3: The Analysis Toolkit	12
3.1 Load Sequence	12
3.2 Time Series:	13
3.3 Spectrogram	14
3.4 Power Spectral Density	14
3.5 Bandpass Filter.....	15
3.6 Fine Scale Frequency Analysis	16
Results and Conclusion	17
References	18
Vita	19

List of Figures

Figure 1: EARS configuration schematic	2
Figure 2: Moving-Coil Transducer. Image thanks to AspenCore, Inc.	2
Figure 3: Schematic of a BM024 cylindrical hydrophone (Jones et. al., 1992).....	3
Figure 4: NOAA illustration of oceanographic zones by depth	4
Figure 5: Sperm whale cranial schematic (Madsen, et. al., 2002)	4
Figure 6: Frequency Content of Beaked Whale Signals	5
Figure 7: Rect function plot	7
Figure 8: Gibbs oscillations in the PSD as a result of windowing the time series data	7
Figure 9: Sinc function plot.....	8
Figure 10: Example of $H(t)$ with values $kc = 10$ & $2n = 4$	8
Figure 11: The Fourier transform does not localize frequencies with the time of occurrence	9
Figure 12: Localizing better in time sacrifices resolution in frequency, and vice-versa	9
Figure 13: Wave vs. Wavelet (Mann and Haykin, 1995)	9
Figure 14: Uniform resolution STFT vs. multiresolution wavelet transform	9
Figure 15: Compression and Dilation of a sinc wavelet	10
Figure 16: Examples of wavelets $\psi(\omega)$	11
Figure 17: Thresholding example with $\delta = 0.5$	11
Figure 18: Image of the analysis tab of the Banshee program	12
Figure 19: Time series of an EARS file captured in the Northern Gulf of Mexico generated by Banshee .	13
Figure 20: Time series of blown up click using start time and duration edit boxes.....	13
Figure 21: Spectrogram of entire EARS file.....	14
Figure 22: Spectrogram of smaller selection of data from EARS file.....	14
Figure 23: Power Spectral Density (PSD) generated by Banshee	14
Figure 24: Catalog Tab of Banshee	15
Figure 25: (a) original time series; (b) original PSD; (c) time series with $kc = 20kHz$ (d) PSD with $kc = 20kHz$	16
Figure 26: Fourier vs. Wavelet analysis	16

Abstract

This project started early in the summer of 2016 when it became evident there was a need for an effective and efficient signal analysis toolkit for the Littoral Acoustic Demonstration Center Gulf Ecological Monitoring and Modeling (LADC-GEMM) Research Consortium. LADC-GEMM collected underwater acoustic data in the northern Gulf of Mexico during the summer of 2015 using Environmental Acoustic Recording Systems (EARS) buoys. Much of the visualization of data was handled through short scripts and executed through terminal commands, each time requiring the data to be loaded into memory and parameters to be fed through arguments. The vision was to develop a graphical user interface (GUI) that would increase the productivity of manual signal analysis. It has been expanded to make several calculations autonomously for cataloging and meta data storage of whale clicks. Over the last year and a half, a working prototype has been developed with MathWorks matrix laboratory (MATLAB), an integrated development environment (IDE). The prototype is now very modular and can accept new tools relatively quickly when development is completed. The program has been named Banshee, as the mythical creatures are known to “wail”. This paper outlines the functionality of the GUI, explains the benefits of frequency analysis, the physical models that facilitate these analytics, and the mathematics performed to achieve these models.

Keywords: Physics; Signal Processing; Acoustics; Software Engineering; Numerical Analysis and Scientific Computation; Bioacoustics

Chapter 1: Introduction

1.1 History of LADC-GEMM

The Littoral Acoustic Demonstration Center (LADC) was formed in early 2001 to utilize autonomous Environmental Acoustic Recording System (EARS) buoys developed by the Naval Oceanographic Office (NAVOCEANO) at Stennis Space Center which has provided technical guidance and support to LADC. The purpose of LADC is to make environmental and marine mammal measurements and perform related data analysis. LADC is a consortium of scientists from universities, including the University of New Orleans (UNO), and the U.S. Navy, which has also used the buoys to characterize the three-dimensional acoustic field of a seismic airgun array and to analyze the noise due to nearby storms. LADC has conducted marine mammal experiments in the Gulf of Mexico (GoM) in 2001, 2002, 2007, 2010, and 2015.

EARS buoys were developed as autonomous moored underwater recording systems by NAVOCEANO to make long-term ocean ambient noise measurements. When LADC was formed, the buoys were capable of measuring up to 1000 Hz for 1 year. LADC later added listening to sperm whales to its noise and propagation measurement missions, and so NAVOCEANO quickly modified the buoys to measure up to 5859 Hz for 36 days. The buoys, moored at depths from 550 to 950 meters in the Gulf of Mexico, produced exceptionally clear recordings of sperm whale echolocation and coda clicks and recordings of other whales.

Because of increasing U.S. Navy interest in beaked whales, the technology was further developed. EARS Generation 2 buoys are now capable of recording one channel to 96 kHz, or 4 channels to 25 kHz. All buoy designs include high quality omnidirectional hydrophones.

In 2007, LADC conducted an exercise, to record primarily not only beaked whales but also sperm whales and dolphins in the Northern GoM. The Generation 2 EARS buoys were used in the single-channel mode. It made the first recordings of beaked whales in the GoM. Coincidentally, the two sites selected for deployment, based on a high density of beaked whale sightings, are 9 and 23 miles from the location of the BP oil spill in 2010. Therefore, LADC received funding in 2010 from the National Science Foundation and Greenpeace to go back to the sites.

LADC, along with several new partners, under the name LADC-GEMM (Gulf Ecological Monitoring and Modeling) received 3 years of funding (2015 through 2017) from the Gulf of Mexico Research Initiative (GoMRI) to conduct two exercises. All of the existing marine mammal recordings are still being analyzed, however, the purpose of these additional exercises are to compare sperm whale, beaked whale, and dolphin measurements using ten single-channel Generation 2 moored EARS buoys, autonomous surface vehicles (ASV's), sea gliders, and a test cruise in the off year between the two exercises. All measurements were made at the three sites of the LADC 2010 experiments.

The first exercise began with a cruise to deploy the hydrophones and sea glider, as well as, operate the ASV's in June of 2015. A retrieval cruise took place in October 2015. The second field exercise began with deployment in April of 2017 and retrieval took place in August of that year.

1.2 The Data

The EARS moorings are passive acoustic monitoring (PAM) system used to record marine mammal phonations. Each mooring has two paired single channel EARS buoys with one-meter separation. Glass floats hold the mooring upright with the goal of keeping the receivers at an approximate depth of 1000 meters; where marine mammals are known to feed.

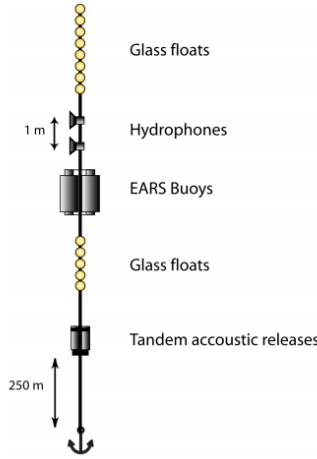


Figure 1: EARS configuration schematic

The data are stored in the binary files as voltage amplitudes that were sampled at a given frequency, in the case of the 2015 experiment, 192 kHz. The data were continuously recorded for one hundred days before being retrieved for processing. This paper will not go into deep detail as to how this process of recording acoustic data takes place, however a very general explanation of transducer theory can be useful, or even important for context.

The hydrophones installed in the EARS platform employ omnidirectional input transducers that convert the mechanical energy of an underwater acoustic wave into electrical energy measured in voltage. Transducers use a diaphragm, which oscillates when struck by an acoustic pressure wave, to produce a measurable potential difference, or voltage. This source of voltage is usually a conductive coil which is pushed through a magnet or vice-versa. The voltage (V) (Kinsler, et. al., 2000) is defined by

$$V = Z_{EB}I + T_{em}u$$

where T_{em} is the transduction coefficient (manufacturer provided in most cases), Z_{EB} is the blocked electrical impedance, I is the current at the electrical inputs, and u is the speed of the radiating surface or diaphragm. The oscillation of the diaphragm, and thus the coil, creates electrical pulses in a circuit, which are directly proportional in amplitude, phase, and frequency to the pressure wave that produced the mechanical movement of the diaphragm. The image below is an example of a basic microphone circuit diagram. This is merely for reference and is not specific to the actual hydrophones utilized by the EARS buoys.

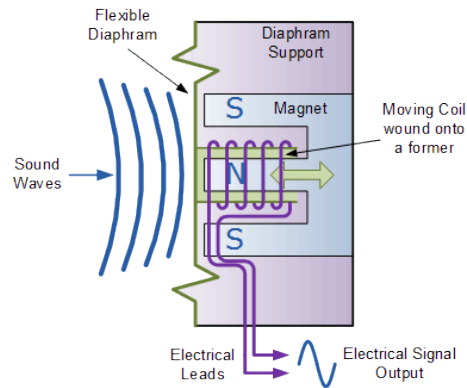


Figure 2: Moving-Coil Transducer. Image thanks to AspenCore, Inc.

The hydrophones employed by the EARS buoys are of a much different design, though the concepts are similar. To achieve an omnidirectional underwater transducer, a cylindrical ceramic piezoelectric design is used. Unlike the diaphragm in the microphone diagram pictured above, which relies on a secondary mechanism to produce an electrical pulse, piezoelectric materials have the ability to generate an electrical charge directly in response to certain forms of mechanical stress. The phenomenon is known as the piezoelectric effect. The ceramic elements are radially poled and are connected electrically in parallel and generate an electric charge that is relative in magnitude to any acoustic pressure wave they may come in contact with.

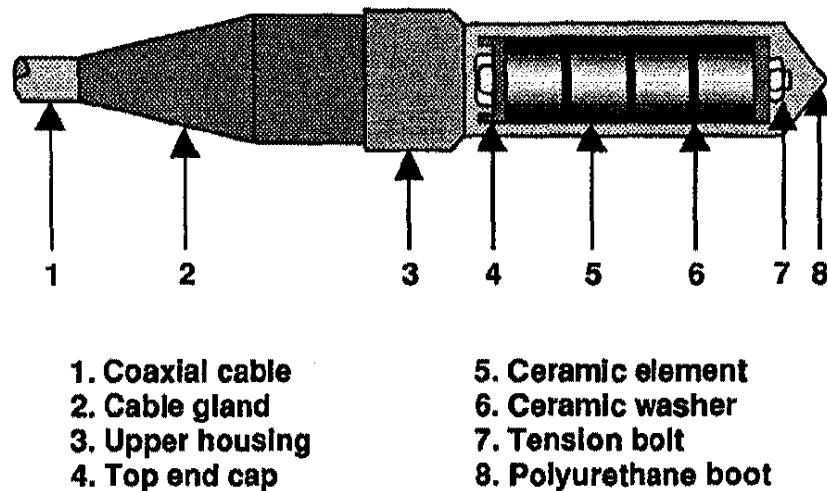


Figure 3: Schematic of a BM024 cylindrical hydrophone (Jones et. al., 1992)

The pulse amplitudes of these waves are then stored in mechanical hard drives in the EARS housing for later retrieval. As discussed above, these hydrophone arrays are left on site for several months to collect large amounts of data in the northern GoM. In 2015 alone, there were 32 terabytes of acoustic data collected from all LADC-GEMM sites.

When working with such large data sets, it is important to have efficiency in mind. One goal of the UNO branch of LADC-GEMM is to create a catalog of mammal clicks for future reference. Mining the data can be an arduous process, therefore development of productivity software is vital. Acoustic data from any single LADC experiment can support a variety of marine mammal related studies, often in parallel and complementary to each other. For example, several techniques were explored for individually identifying sperm whales by their clicks for future comparison. Many of these techniques require extensive knowledge of frequency analysis. Chapter 2 will review some of the more important aspects of this field of study.

Chapter 2: Applications of Transform in Signal Analysis

2.1 Bioacoustics

In the underwater-acoustic applications of signal analysis, the time series allows the analyst to observe the instantaneous relative magnitudes of the sound pressure waves which propagate through the medium. Over time, analysts can observe these fluctuations in magnitude as signals. The Fourier transforms, discussed at length in the following sections, of these time signals are their spectra. Aptly referred to as the “frequency domain”, the results of this mathematical operation are literally the amplitudes at the frequencies present in the signal. This is extremely useful to acousticians and marine biologist, as it can be difficult to gather visual data on underwater mammals. Many of these species of use acoustic echolocation for sensory processing of their surroundings. They rely on it for hunting and navigation, since electromagnetic radiation (light) does not penetrate water for more than about 1000 meters. Figure 5 below was taken from the National Oceanic and Atmospheric Administration (NOAA) website and explains a few of the zones of interest to researchers.

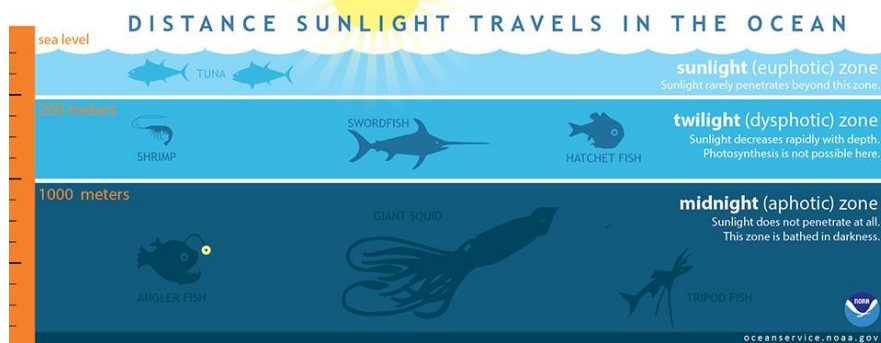


Figure 4: NOAA illustration of oceanographic zones by depth

As many whales rely on species that reside in the aphotic zone for sustenance, they must dive to these depths to hunt their prey. At these depths, their eyes are insufficient for conveying perception. As an example of such species, consider the sperm whale. The sperm whale uses a biological mechanism known as the monkey lips, or museau de singe, to generate a broadband acoustic pressure wave known as a click. This is annotated as “Mo” in the figure below. This signal bounces off of, and is amplified by, the frontal air sac (Fr) and directed through the junk (Ju) where it exits and travels through water. This creates an extremely high sound pressure level that is ideal for the echolocation method utilized by the sperm whale.

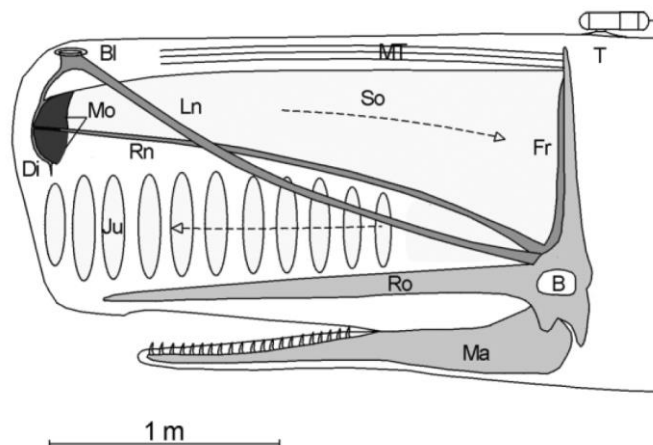


Figure 5: Sperm whale cranial schematic (Madsen, et. al., 2002)

Scientists can exploit these acoustic echolocation data to determine population densities and behavior. Furthermore, the data can be used for localization and classification since each species operates in slightly different frequency bands. Much of the efforts of LADC-GEMM are centralized around the classification of species given their calls. The intent is to build on the works of noted researchers such as Dr. George Ioup, Dr. Juliette Ioup, and Dr. Simone Baumann-Pickering to name a few. As an example of such works, the figure below was taken from Baumann-Pickering, et. al (2013).

	Peak frequency (kHz)	Center frequency (kHz)	−10 dB bandwidth (kHz)	Duration (μs)	IPI (ms)	Click
Baird's beaked whale	16.4 (9.0, 27.0)	20.4 (14.7, 31.3)	8.6 (4.7, 19.1)	504 (275, 875)	204 (80, 369)	yes
Longman's beaked whale	22.0 (14.2, 29.8)	22.7 (16.9, 28.1)	20.8 (11.1, 39.1)	182 (101, 375)	–	yes
Blainville's beaked whale	34.4 (31.3, 44.1)	37.3 (32.3, 44.0)	11.7 (5.5, 23.0)	581 (299, 950)	280 (111, 427)	no
Cuvier's beaked whale	40.2 (20.3, 49.2)	35.9 (28.7, 42.5)	10.9 (5.1, 21.9)	585 (306, 976)	337 (94, 491)	no
BW40	42.6 (33.6, 52.3)	40.1 (33.1, 46.5)	10.5 (4.7, 20.7)	575 (250, 1031)	435 (314, 538)	yes
Northern bottlenose whale	–	43 ± 7	–	276 ± 58 ^a	306 ± 118	no
Gervais' beaked whale	43.8 (35.9, 55.9)	45.2 (37.5, 55.0)	18.8 (7.8, 34.8)	450 (260, 765)	275 (114, 353)	no
BW43	43.4 (37.7, 55.1)	45.2 (37.5, 53.8)	19.9 (7.8, 38.7)	395 (270, 621)	217 (132, 441)	no
Deraniyagala's beaked whale	47.3 (28.9, 69.1)	46.8 (33.5, 57.1)	19.5 (9.0, 36.7)	475 (305, 720)	194 (70, 429)	no
BWG	46.9 (33.6, 56.1)	43.0 (33.8, 52.6)	30.9 (16.8, 43.8)	535 (155, 1023)	133 (72, 355)	no
BWC	46.9 (28.9, 73.8)	47.4 (30.3, 65.4)	26.2 (5.9, 48.4)	779 (270, 1210)	127 (66, 338)	yes
Stejneger's beaked whale	50.4 (45.7, 73.8)	56.1 (46.1, 67.8)	21.1 (8.6, 39.1)	420 (245, 746)	90 (65, 224)	no
BW70	66.4 (60.9, 78.9)	66.9 (60.9, 75.1)	23.4 (12.4, 31.6)	435 (291, 655)	119 (92, 217)	no

^aDurations derived from 95% energy, in comparison to Teager-energy as used in this manuscript, may be slightly shorter and not fully comparable.

Figure 6: Frequency Content of Beaked Whale Signals

Findings, such as these, give a baseline to researchers attempting to classify future signals recorded from other underwater mammals. Many of the values listed in the table above rely on access to the frequency domain of the signal generated by the animal.

2.2 The Fourier Transform

The Fourier transform has been accepted as a staple technique in signal analysis and many other fields of data science. Though other transforms may be used from time to time, it would be difficult to find any research group in the signal analysis field that does not regularly and routinely utilize Fourier analysis in some form. Other transforms, such as the Laplace and Z-transforms, may be treated as strictly mathematical models, and are often used to simplify equations and provide a more easily solvable system prior to returning to the function domain. The Fourier transform and the resulting frequency spectra, on the other hand, have as real an interpretation as the function domains themselves (i.e. time, space, etc.). As discussed above, the time series data must be passed through the Fourier transform to obtain the frequency spectrum. For the purposes of this paper, the Fourier transform (Bracewell, 2000) is given by

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-i2\pi ft} dt \quad (1)$$

and the inverse Fourier transform is given by

$$x(t) = \int_{-\infty}^{\infty} X(f)e^{i2\pi ft} df \quad (2)$$

to return to the function domain, where $x(t)$ is the signal, $X(f)$ is the Fourier transform, dt is an infinitesimally small increment of time, and df is an infinitesimally small increment of frequency. These are defined to be the continuous Fourier transform, and are not indicative of real, sampled data. To rectify

this, the terms dt and df can no longer be considered infinitesimally small. For a signal of integer length N (starting at 1) and sample frequency f_s , the time array T is

$$\Delta t = \frac{1}{f_s} \quad (3)$$

$$T = [0: \Delta t: (N - 1)\Delta t] \quad (4)$$

Note: The notation used here is $[first\ value: period\ or\ step\ interval: last\ value]$. This is shorthand for a column vector valued object with N number rows.

As implied above, equations (1) and (2) take Δt from equation (3) to be infinitesimally small, but of course this is only theoretical. In applied signal analysis, data is sampled and is represented by discrete measurements in time. An adaptation of the Fourier transform for real sampled data is known as the discrete Fourier transform (DFT). If N is the number of data points in a given signal, and the independent variable is n , the DFT $X(k)$ of the data $x(n)$ is given by

$$X(k + 1) = \sum_{n=0}^{N-1} x(n + 1)W_N^{kn} \quad (5)$$

and its inverse is given by

$$x(n + 1) = \frac{1}{N} \sum_{k=0}^{N-1} X(k + 1)W_N^{-kn} \quad (6)$$

where k is the independent variable in the frequency domain and

$$W_N = e^{-j2\pi/N} \quad (7)$$

MATLAB uses the engineering convention of $-j = -i = -\sqrt{-1}$ in the forward DFT and normalizes in the inverse transform operation as is required by the sum of sequence and first value theorem (Bracewell, 2000). This notation was adopted by the author for continuity. Using the results of what is known as the Short Time Fourier transform (STFT), a spectrogram can be built to visualize and analyze the frequency bands. See section 2.5 for discussion and section 3.3 for example images.

2.3 Power Spectral Density

Other aspects of the signal that are of interest include the peak frequency and the various bandwidths that can be obtained from a power spectrum (PS). The PS is the squared, normalized, absolute value of the Fourier transform. This operation allows the analyst to emphasize the changes in amplitude across many frequencies. The PS (Bracewell, 2000, pgs. 285-288) is defined by:

$$PS = \left| \frac{X(k)}{\max[X(k)]} \right|^2 \quad (8)$$

Often when visualizing the power spectrum in acoustics, the amplitudes are displayed in logarithmic scale using the loudness unit, decibels (Kinsler et al, 2000, pg. 130). In this form it is known as the power spectral density (PSD):

$$PSD = 10\log_{10} PS \quad (9)$$

2.4 Bandpass Filtering

Bandpass filtering is a mathematical model of cutting out unwanted frequencies in order to denoise a signal. In this context, windows are utilized in the frequency domain to allow only the frequencies of interest to remain. The most basic of these windows is known to students of Ronald Newbold Bracewell as the Rect function:

$$\Pi(t) = \begin{cases} 1 & -\frac{1}{2} < t < \frac{1}{2} \\ 0 & \text{elsewhere} \end{cases} \quad (10)$$

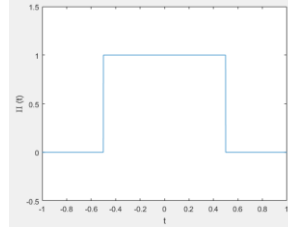


Figure 7: Rect function plot

As the saying goes, “there is no free lunch,” and as expected, there can be drawbacks to this in some situations. A maximally flat magnitude bandpass filter is utilized as to obtain as flat of a frequency response as possible. If this is not done, distortions can occur when returning to the time domain. A phenomenon known as Gibbs’s oscillations can present itself as no filter can be perfectly flat at either end.

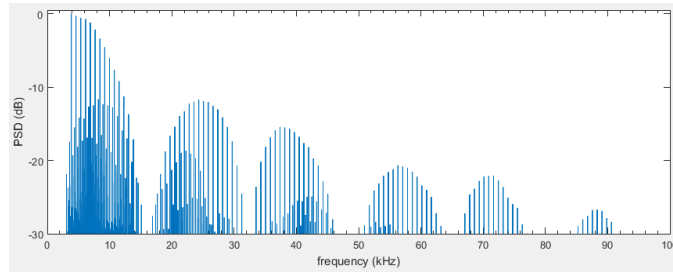


Figure 8: Gibbs oscillations in the PSD as a result of windowing the time series data

This occurs due to the result of carrying the window with the signal in the inverse transform. This can be illustrated mathematically with the convolution theorem (Bracewell, 2000). Give the signal functions $f(t)$ & $h(t)$ with transform $F(s)$ & $H(s)$ respectively, the convolution of such functions results in the following relationships:

$$\begin{aligned} f(t) * h(t) &\supset F(s)H(s) \\ f(t)h(t) &\supset F(s) * H(s) \end{aligned} \quad (11)$$

Since applying a filter in the frequency domain literally means taking the product of the filter and the transform coefficients, the transform of the filter effects the time series data as well. The oscillatory behavior is a direct result of the Fourier transform of window like function. Consider again the Rect function; its Fourier transform is the sinc:

$$\Pi(t) \supset \text{sinc}(s) = \frac{\sin(\pi s)}{\pi s} \quad (12)$$

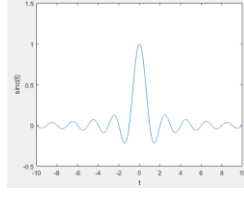


Figure 9: Sinc function plot

From his original paper, S. Butterworth (1930) is quoted as saying, "An ideal electrical filter should not only completely reject the unwanted frequencies but should also have uniform sensitivity for the wanted frequencies." Obviously, this is impossible, but the 4th order Butterworth Filter does well enough for the purposes of this software. This filter uses the frequency responses (Oppenheim & Shafer, 1975, pg. 211):

$$\text{lowpass: } |H_{lp}(jk)|^2 = \frac{1}{\sqrt{1 + \left(\frac{jk}{jk_c}\right)^{2n}}} \quad (13)$$

$$\text{highpass: } |H_{hp}(jk)|^2 = \frac{\left|\frac{jk}{jk_c}\right|^{2n}}{\sqrt{1 + \left(\frac{jk}{jk_c}\right)^{2n}}} \quad (14)$$

Where k_c is the cutoff frequency and n is the order. These filters are difficult to visualize due to their imaginary part and the variability of their scale and order, however, the following figure can give an idea of its shape by considering only the real parts:

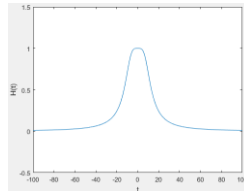


Figure 10: Example of $H(t)$ with values $k_c = 10$ & $2n = 4$

Note that the shape attempts to become as flat as possible as it approaches zero on either side. These types of filters are known as "max flat".

2.5 The Wavelet Transform

As has been illustrated above, the Fourier transform is extremely useful in many forms of data analytics, but it does have its drawbacks. Its use of then analyzing function, shown here as Euler's equation,

$$e^{-j2\pi/N} = \cos\left(\frac{2\pi}{N}\right) - j \sin\left(\frac{2\pi}{N}\right) \quad (15)$$

employs the method of adding a series of products of infinitely long wave functions with data amplitudes to obtain frequency data. This can often be useful; however, the use of such infinite analyzing functions fails to localize well with respect to time.

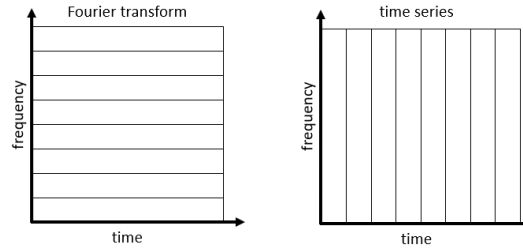


Figure 11: The Fourier transform does not localize frequencies with the time of occurrence

The STFT attempts to amend this by taking small chunks of the data to perform transform in windows. These are then assembled later into spectrograms creating a pixelized pattern that can do a decent job of providing time data with the frequencies obtained. Unfortunately, due to the nature of the operation, aspects of the uncertainty principle being to apply, as is demonstrated by the following figure.

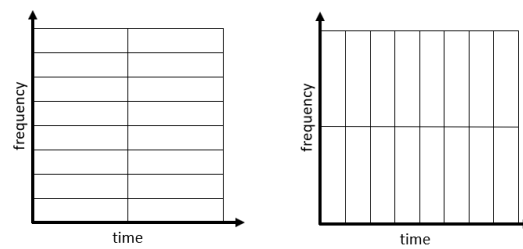


Figure 12: Localizing better in time sacrifices resolution in frequency, and vice-versa

In order to address these drawbacks, a mathematical operation known as the wavelet transform is considered. Wavelets are finite wave functions that have compact support and other special properties that allow for some very interesting analytics.

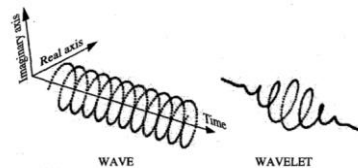


Figure 13: Wave vs. Wavelet (Mann and Haykin, 1995)

In contrast to the analysis function utilized by the Fourier transform, where pixels are uniform, wavelets are able to localize frequencies much better in time through the concept of multiresolution. At lower frequencies, this may not be necessary, as the analyzing function is wide enough to capture them. As the frequencies get higher, a scaling factor adjusts the wavelet by compressing it. Some resolution with respect to frequencies captured is lost, but it allows analysts to know when these higher frequencies occurred with relative accuracy, whereas the STFT could have missed them all together.

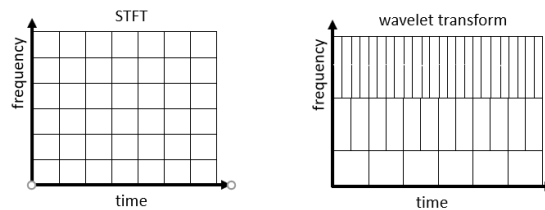


Figure 14: Uniform resolution STFT vs. multiresolution wavelet transform

The wavelet transform is given by

$$F(a, b) = \frac{1}{|a|^{1/2}} \int_{-\infty}^{\infty} f(t) \psi\left(\frac{t-b}{a}\right) dt \quad (16)$$

where ψ , a wavelet, has taken the place of the exponential function in the Fourier transform. The variable a is known as the scaling factor which either dilates or compresses the wavelet and is analogous to the frequency content of the Fourier transform, and b is known as the translation factor which acts as a phase shifting mechanism that localizes the frequencies captured with the time they were present. The equivalent frequency (F_{eq}) can be obtained by calculating the center frequency of the wavelet (C_f), and is inversely proportional to the product of the scaling factor and the sampling interval.

$$F_{eq} = \frac{C_f}{a} \cdot \frac{1}{\Delta t} \quad (17)$$

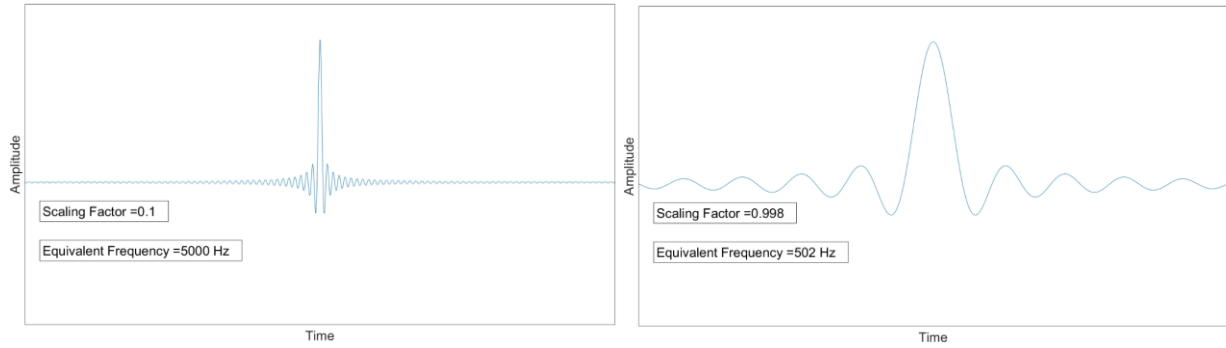


Figure 15: Compression and Dilation of a sinc wavelet

Scaling any wavelet by a factor of 2, for example, results in reducing F_{eq} by an octave. Since this transform results in a function of two variables, the inverse transform is not symmetric like its cousin the Fourier transform. The inverse can be obtained by the equation

$$f(t) = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(a, b) \frac{1}{|a|^{1/2}} \tilde{\psi}\left(\frac{t-b}{a}\right) db \frac{da}{a^2} \quad (18)$$

where $\tilde{\psi}$ is the dual function of ψ . The explanation of this relationship is outside the scope of this paper, but it is important to note that, in all wavelets used by the author, ψ is biorthogonal and square integrable, and thus $\psi = \tilde{\psi}$ (Unser, 1996). Further, C_ψ is known as the wavelet admissibility constant and is given by

$$C_\psi = \int_{-\infty}^{\infty} \frac{|\psi(\omega)|^2}{|\omega|} d\omega \quad (19)$$

The continuous wavelet transform defined above can also be translated to their discrete versions as well, however this is very laborious process due to its lack of symmetry and not necessary in the scope of this paper.

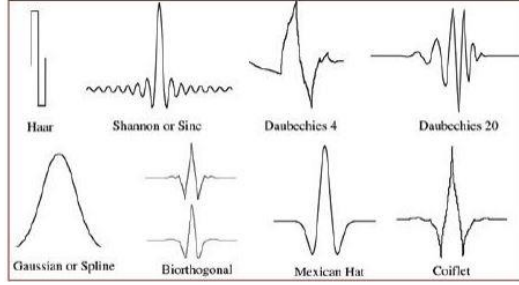


Figure 16: Examples of wavelets $\psi(\omega)$

Once the results of the transform have been obtained, a process known as thresholding can be utilized to denoise the signal. In signal analysis, the most widely used process is that of soft thresholding, where a data dependent threshold δ is chosen. The standard method is to choose the threshold value statistically (Kolaczyk, 1996).

$$\delta = \sigma \sqrt{2 \log_{10}(N^2)} \quad (20)$$

Where σ is the standard deviation of the noise and N is the number of data points. This is a basic equation for calculating entropy. Values can also be chosen manually for testing purposes. All values present in the signal that fall below the absolute value of δ are set to zero. For the practice of hard thresholding the process would end here. For soft thresholding, δ is then subtracted from the remaining values. The filter is applied as follows (Strang, 1996).

$$y_{soft}(t) = \begin{cases} 0 & \text{for } |x(t)| < \delta \\ \frac{x(t)}{|x(t)|} (|x(t)| - \delta) & \text{elsewhere} \end{cases} \quad (21)$$

This ensures a reconstructed thresholded function that is at least as smooth as the original function. Hard thresholding tends to exhibit distortions due to discontinuities. As an example of thresholding, consider the function $x(t) = \cos t$, $0 < t < 4\pi$

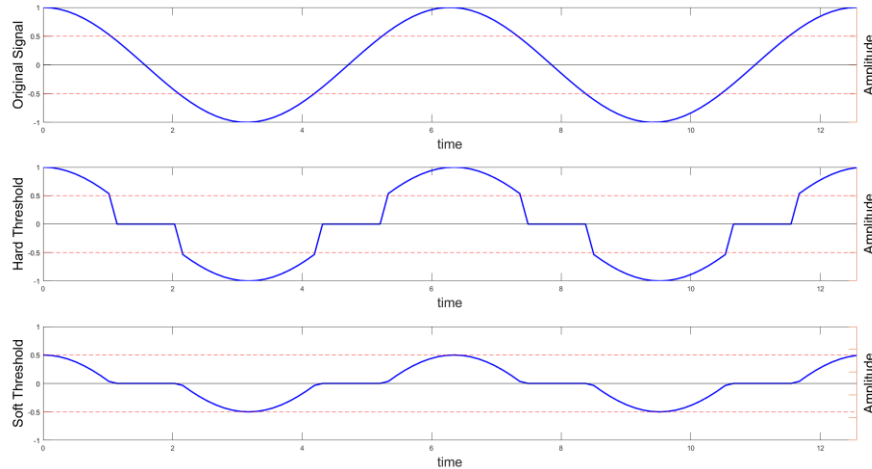


Figure 17: Thresholding example with $\delta = 0.5$

The techniques discussed in this chapter have all been utilized in the development of the Banshee software. Example images, explanations of the capabilities, features, and processes are shown and discussed in the following chapter.

Chapter 3: The Analysis Toolkit

The original toolset, known henceforth as the Analysis Tab, includes a GUI that accepts a file path, aided by a secondary interface of the computer's file structure. It is compatible with both EARS binary files and text data files. This interface can load the data into the computer's memory and perform a series of operations to visualize the data in many useful ways, including plots of a time series, its time-frequency spectrogram, and its PSD.

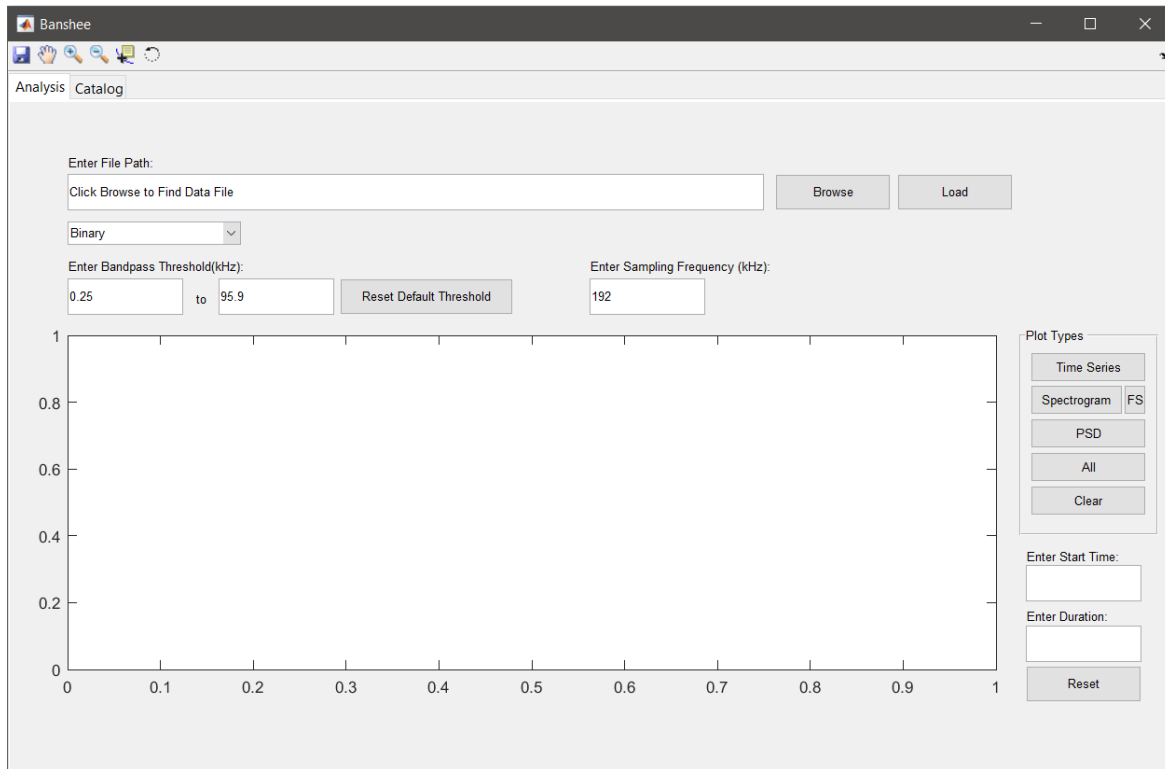


Figure 18: Image of the analysis tab of the Banshee program

3.1 Load Sequence

In order to read these data, the GUI was constructed in the MATLAB environment, which has built-in read functions for binary files. This makes the process less arduous than using open source development environments, however some tailoring was still necessary to organize the data into a usable format. The "Load" button has two possible operations. By default, it will run a binary read algorithm that decodes and loads the data of a single EARS binary file into the computer memory for later access. This includes not only the voltage amplitudes, but also time stamp information for each file of slightly more than 21 seconds. These data are then stored as an array or "vector" object inside a class along with all other GUI objects (i.e. buttons, edit boxes, toggles, etc.). The secondary operation is dictated by a drop-down menu located below the file path edit box. The "Text" option will allow the user to load a .dat file containing an array of amplitude values. These files can be created with the GUI through the save icon at the top left-hand corner or by any other method. The format of the .dat files is extremely important, and caution should be taken when creating said files outside of the GUI program.

3.2 Time Series:

The simplest of the operations, but every bit as important, is the time series. It allows the analyst to navigate the data file and decide on which portions of the signal to focus. The “Times Series” button accesses the data loaded into memory and calculates a time array (based on the given sample rate and number of data points), as the times of the samples are not stored in the data for every sample point.

This array is then coupled with the signal array for a two-dimensional time series plot of signal amplitude on the vertical axis versus time on the horizontal axis as seen in the figure below. The time array is the relative time from the beginning of the file or the first plotted point. The actual time can be calculated by taking the relative time and adding it to the timestamp given under the “All” button or by accessing the “Catalog” tab. This value, in future releases will be displayed for the user on the GUI for more efficient access to the real time of the event being analyzed. This could save users’ time when generating reports for Digital Object Identifier (DOI) numbers. The following figure is an example of a time series plot of several sperm whale clicks, recorded in the Northern GoM. As discussed in Chapter 1, the amplitudes are recorded voltages relative to the ambient pressure at the hydrophone depth.

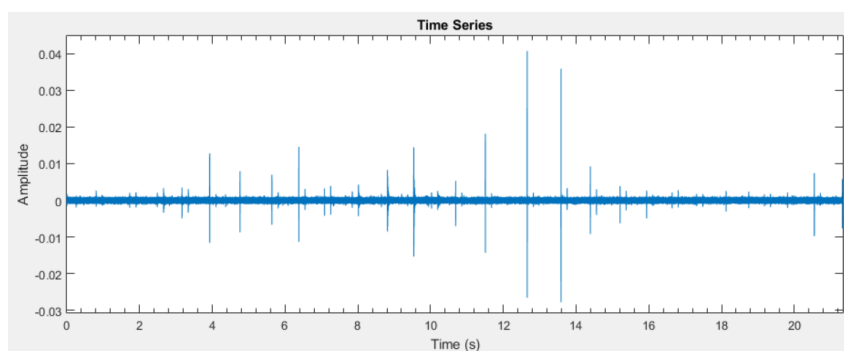


Figure 19: Time series of an EARS file captured in the Northern Gulf of Mexico generated by Banshee

Additional tools have been included to further manipulate the plot, such as zoom, pan, and a data cursor for closer inspection. Once a region of interest is found, the start time and duration edit boxes can be utilized to select a much smaller array for further analysis. The axes are automatically reset so the signal extrema span the entire viewing window to allow easier inspection.

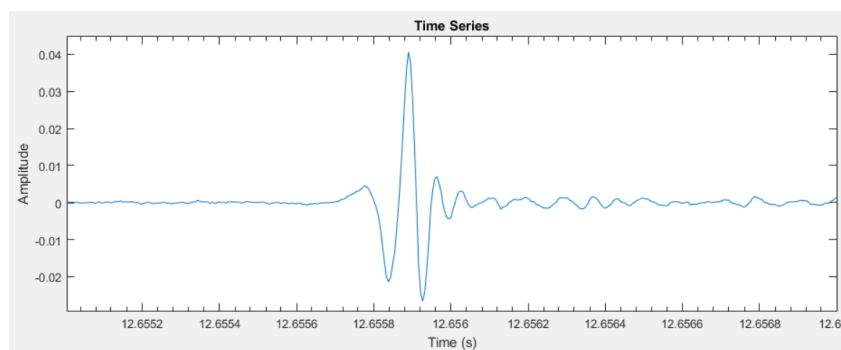


Figure 20: Time series of blown up click using start time and duration edit boxes

This not only allows the visualization of the smaller set of data, as the zoom would provide, but also permits the user to apply more complex operations on only the shorter data set. In other words, all operations requested will then be performed on the set specified by the edit boxes rather than the full data set loaded into memory. This object-based system uses far less random-access memory (RAM) and processing power than previous methods utilized (i.e. command line called programs).

3.3 Spectrogram

In order to visualize the frequency bands, the spectrogram operation of the GUI takes the portion of the data dictated by the edit boxes discussed above and utilizes the STFT and performs several analysis operations. For large data sets, an algorithm has been included to down sample the signal for faster processing. For small data sets, however, we can achieve much higher resolution pictures. As an example, the figures below are over 4 million data points and only 380 data points, respectively.

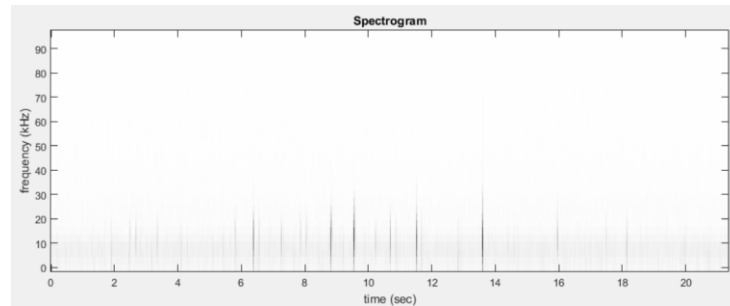


Figure 21: Spectrogram of entire EARS file

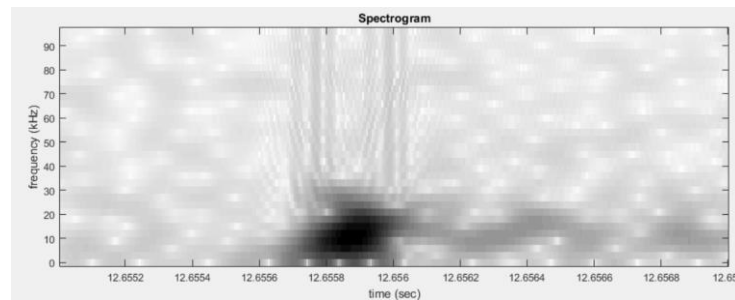


Figure 22: Spectrogram of smaller selection of data from EARS file

The spectrogram images displayed here use the three-dimensional method of time on the horizontal axis, frequency on the vertical axis, and pixels of varying darkness to indicate amplitude (higher amplitudes are darker). The advantage of this type of visualization is the ability to compare behavior in both the frequency domain and the time domain at selected times and frequencies.

3.4 Power Spectral Density

The PSD button will run the calculations as described in Chapter 2. The amplitudes are displayed in logarithmic scale using the loudness unit, decibels. The maximum amplitude therefore occurs at zero dB, known as the peak frequency, with all other levels relative to this max value.

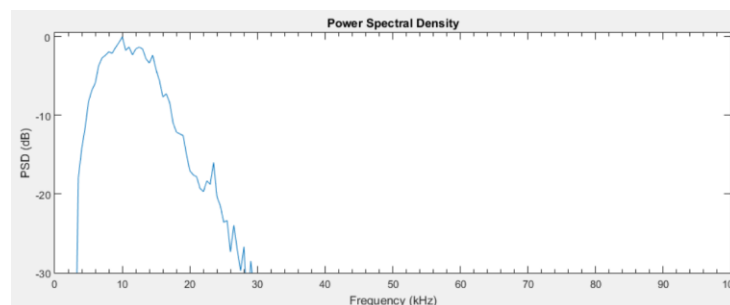


Figure 23: Power Spectral Density (PSD) generated by Banshee

The resulting transformed data allow the analyst to obtain important information characteristic of the signal. The -10dB bandwidth is the difference of the frequencies where the PSD crosses this threshold. The center frequency is half the -10dB bandwidth plus the frequency value of the first crossing. These values are often unique to many species, and as such are important for classification. The Catalog tab of Banshee has been programed, in large part thanks to Mr. Britt Aguda, to calculate all of these values of interest. An algorithm was also developed that can accurately distinguish the -10dB bandwidth, even if the PSD fluctuates across the threshold prematurely. Once all values are obtained, the “Report” button will create a .mat object file that stores this meta data along with a .dat file of the amplitudes for future processing.

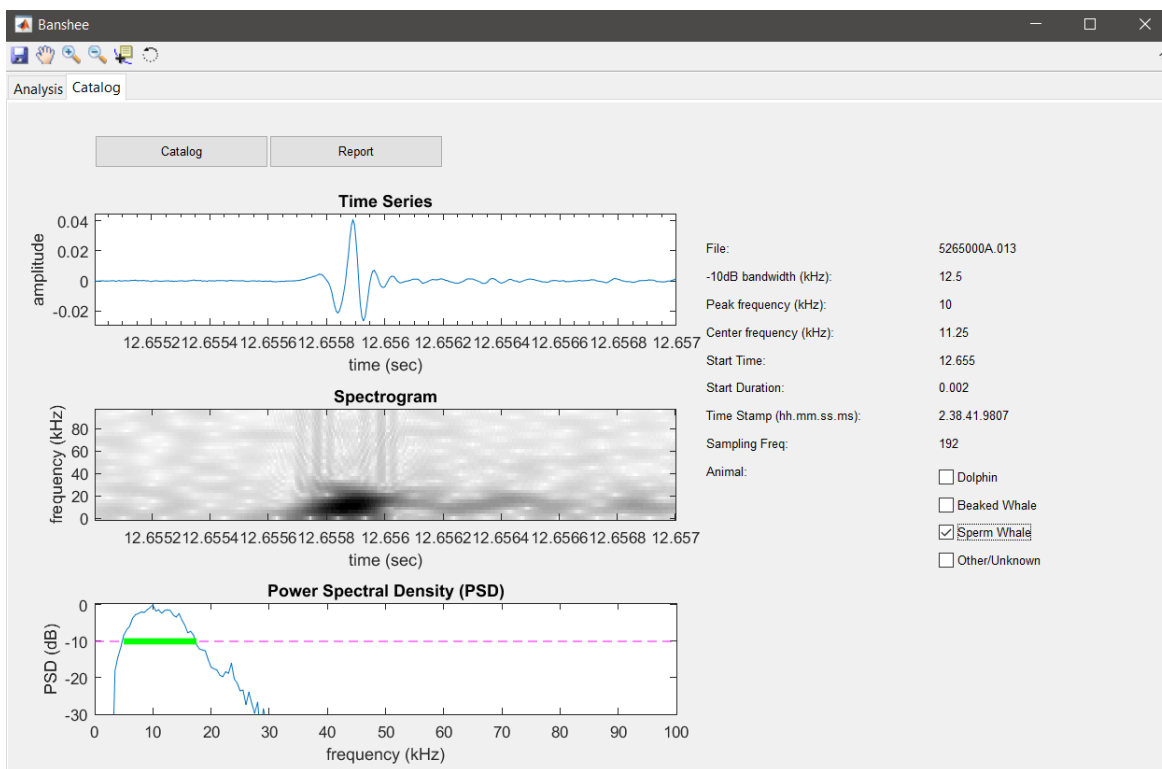


Figure 24: Catalog Tab of Banshee

3.5 Bandpass Filter

In the practical sense, the intensity of sound is nonuniformly distributed over different frequencies across time. The receivers record a range of these frequencies from a multitude of sources. Much of this is of no interest to any particular group, and as such is considered to be noise. It can be said that one individual's noise is another's signal. With this idea in mind, a bandpass filter was included in the toolset. The GUI was designed to always utilize the filter for noise removal. By default, it is set at the receiver extrema, resulting from the sample interval or rate:

Enter Bandpass Threshold(kHz): 0.25 to 95.9 Enter Sampling Frequency (kHz): 192

As LADC-GEMM 2015 data has a sample rate of 192kHz, this was set as the default, but it can be changed if older or newer data has a different sample rate. The bandpass threshold will automatically set itself from a small frequency (approximately zero with respect to these data) to the cutoff frequency, or in other words, half the sample frequency. The figures below demonstrate the capabilities of the

Butterworth filter employed by the GUI (See Chapter 2 for more detail). Note, the signal displayed as (c) has far less jagged changes and the resulting PSD no longer includes the higher frequencies.

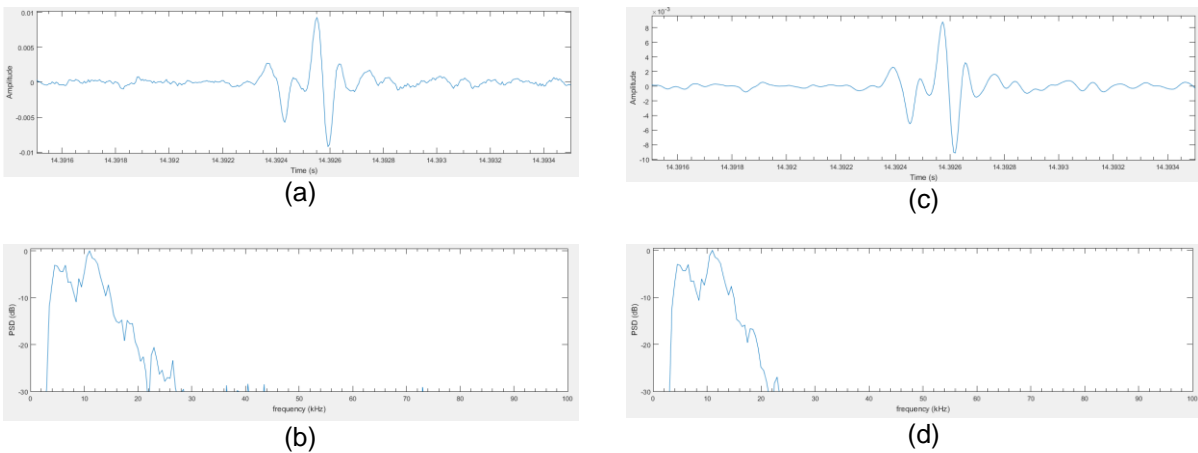


Figure 25: (a) original time series; (b) original PSD; (c) time series with $k_c = 20\text{kHz}$ (d) PSD with $k_c = 20\text{kHz}$

3.6 Fine Scale Frequency Analysis

The FS toggle located next to the Spectrogram button utilizes MATLAB's Wavelet toolbox to produce "fine scale" scaleograms. These functions allow the creation of very high resolution, denoised images for analyst and are analogous to the spectrograms obtained through Fourier analysis. The following figure compares the spectrogram with the same image obtain through wavelet analysis. The image on the right utilizes equation 17 from chapter 2 to obtain the equivalent frequencies.

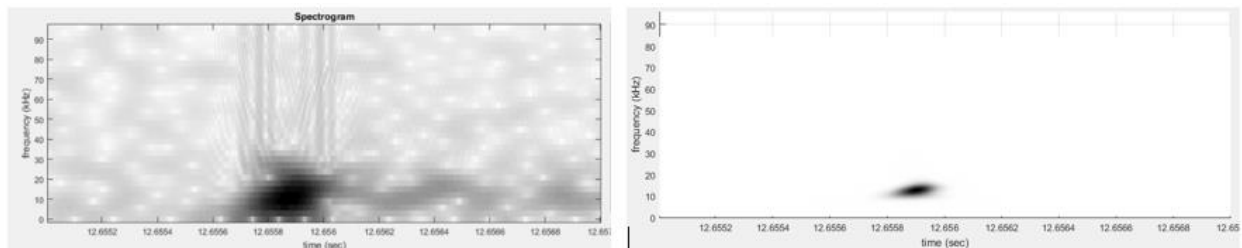


Figure 26: Fourier vs. Wavelet analysis

The toolbox utilizes the wavelet transform and has an algorithm to optimize a δ value for thresholding. This process can achieve remarkable results not possible in traditional Fourier analysis.

Results and Conclusion

The development of this software has been tremendously successful, and toolsets and functionality continue to grow with each month. Manual analysis, cataloging, and data mining of whale clicks now takes a fraction of the time it did with previous methods. The cataloging functionality of Banshee has been expanded to take several files from any given folder; it will find all clicks, calculate the values of interest discussed in this paper, and provide reports on them all, sorting them by their file name. Much of the work that several students produced in the summer of 2016 was replicated in less than 10 minutes by a single user.

Development began in the summer of 2016 as a way to better visualize the data collected and processed by LADC-GEMM. Currently, the software is in the hands of multiple researchers within the consortium, as well as, high school students across the New Orleans metropolitan area. Thus far, Banshee has been received well by users. Feedback tends to be requests for additional features, which have been implemented when completed.

Under the guidance of Mr. Kendal Leftwich outreach programs have been designed to use the software to aid and educate students in data mining and collection procedures. Many of these students are required by the schools to conduct research and UNO LADC-GEMM has a significant amount of data to mine and catalog. As such, this is a mutually beneficial relationship. Thus far the program has been well received, and students are motivated to learn about bioacoustics and signal analysis techniques.

Future plans for the software include expanding the cataloging functionality to include even more data mining capabilities. For instance, the most recent work of Mr. Jack Lebien of UNO uses multiple statistical algorithms to classify underwater mammals based on the criteria extracted above. This includes k-means, self-organizing maps, and several other iterative machine learning techniques. Mr. Matt Firreno, another researcher affiliated with UNO LADC-GEMM, has begun to explore the classification potential of Lorenz attractors, time-embedded manifold creation, fractal dimension, and other applications of chaos theory. These algorithms should be migrated into Banshee to further improve its capabilities. Mr. Britt Aguda has begun development of a database management tool for the reports collected has also just begun, but will eventually operate much like that of other database querying software only specific to LADC-GEMM data. It will generate spreadsheet reports utilizing end user defined parameters to further improve research efficiency.

References

- AspenCore, Inc. (n.d.). Sound Transducers. Retrieved November, 2017, from http://www.electronics-tutorials.ws/io/io_8.html
- Baumann-Pickering, S., McDonald, M. A., Simonis, A. E., Berga, A. S., Merkens, K. P., & Oleson, E. M., ... Hildebrand, H.A. (2013). Species-specific beaked whale echolocation signals. *The Journal of the Acoustical Society of America*, 134, 2293-2301. doi: <http://dx.doi.org/10.1121/1.4817832>
- Bracewell, R. N. (2000). *The Fourier transform and its applications*. Boston, MA: McGraw-Hill.
- Butterworth, S. (1930), "On the Theory of Filter Amplifiers", *The Wireless Engineer*.
- Jones, D. F., Prasad, S. E., & Kavanaugh, S. R. (1992). *An End-Capped Cylindrical Hydrophone for Underwater Sound Detection* (Canada, National Defence, Research and Development Branch).
- Kinsler, L. E., Frey, A.R., Coppens, A.B., & Sanders, J.V. (2000). *Fundamentals of acoustics*. New York: John Wiley & Sons, Inc.
- Kolaczyk, E. D. (1996). *Wavelets in medicine and biology* (A. Aldroubi & M. Unser, Eds.). Boca Raton, FL: CRC.
- Madsen, P.T., Payne, R., Kristiansen, N.U., Wahlberg, M., Kerr, I., & Mohl, B. (2002). Sperm whale sound production studied with ultrasound time/depth-recording tags. *The Journal of Experimental Biology*, 205, 1899-1906.
- Mann, S. & Haykin, S. (1995). "The Chirplet Transform: Physical Considerations", *IEEE Transactions on Signal Processing*, Volume 43, Number 11, pp. 2745–2761.
- Oppenheim, A. V., & Schaffer, R. W. (1975). *Digital signal processing Alan V. Oppenheim Ronald W. Schaffer*. Englewood Cliffs, NJ: Prentice-Hall.
- Strang, G., & Nguyen, T. (1997). *Wavelets and filter banks*. Wellesley, MA: Wellesley - Cambridge Press.
- The MathWork, Inc. (n.d.). MATLAB Documentation. Retrieved November, 2017, from <http://www.mathworks.com/help>.
- Unser, M. (1996). Wavelets, Statistics, and Biomedical Applications. *IEEE*, 244-249.
- US Department of Commerce, National Oceanic and Atmospheric Administration. (2013). How far does light travel in the ocean? Retrieved August 24, 2017, from https://oceanservice.noaa.gov/facts/light_travel.html

Vita

Born May 25, 1985, Kirk D. Bienvenu, Jr. attended school at St. Angela Merici at the age of 6, and continued there from kindergarten to eighth grade. He then attended high school at Archbishop Rummel High School. In 2004, upon graduation, he enlisted in the United States Army (USA) where he was trained as a paralegal specialist in the USA Judge Advocate General (JAG) Corps.

Mr. Bienvenu was subsequently assigned to the 1st Armored Division in Wiesbaden, Germany where he served as legal assistant to the Judge Advocate for 205th Military Intelligence Brigade. He was deployed to Iraq in 2005 where he supported both the military intelligence and detainee operations missions for a year. He returned to Germany where he served the remainder of his contract and was honorably discharged from the USA in 2008, also receiving a commendation and an achievement medal for his service.

In August of 2008 Mr. Bienvenu entered the Jefferson Parish Sheriff's Office (JPSO) academy where he was trained to be a patrol deputy. He worked in corrections for a year in the Jefferson Parish Correctional Facility. At the end of 2009 he was assigned the 3rd District Patrol of JPSO where he provided emergency services and security to West Bank Jefferson Parish neighborhoods. After completion of his two-year commitment to JPSO, Mr. Bienvenu worked a short time for St. John's Sheriff's Office prior to enrolling in classes at Embry-Riddle Aeronautical University. Mr. Bienvenu obtained his Bachelor's degree in Aeronautics in December of 2013.

In the spring of 2014, Mr. Bienvenu enrolled in the engineering school at the University of New Orleans to continue his education. After a year and a half in naval architecture program, he decided that physics was a better fit. He was offered a position as a teaching assistant and admitted into the graduate program for applied physics in the summer of 2015. A year later, in the summer of 2016, Dr. Ioup hired Mr. Bienvenu as her research assistant. Since this time he has been working on acoustic data collected in the northern Gulf of Mexico for the research group LADC-GEMM. This thesis is a portion of the work he has contributed to the group, which also includes work in classification of species and localization of underwater mammals.