Fall 12-20-2017

Automated Species Classification Methods for Passive Acoustic Monitoring of Beaked Whales

John LeBien
University of New Orleans, New Orleans, jlebien@uno.edu

Automated Species Classification Methods
for Passive Acoustic Monitoring of Beaked Whales

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

John LeBien

B.S. University of New Orleans, 2015

December 2017

Acknowledgements

I would like to express sincere gratitude to my advisor, Dr. Juliette Ioup, for her continual support and encouragement. I have developed passionate interests in a number of topics she introduced me to, and her teaching has allowed me the freedom to pursue a wide array of opportunities. She and Dr. George Ioup created the courses which constituted my focus of study. I am grateful for these developments. I would like to thank Dr. Natalia Sidorovskaia for her support as principal investigator of the LADC-GEMM consortium. Also, my LADC-GEMM colleagues at The University of New Orleans, and at The University of Louisiana at Lafayette are thanked for their helpful comments and advice.

I would like to express my appreciation to Dr. Ioannis Georgiou for his encouragement and knowledge. He gave me the opportunity to participate in engaging research during my undergraduate and graduate studies, for which I am very grateful.

I also thank Dr. Ashok Puri, Dr. Leszek Malkinski, Dr. Kevin Stokes, and Dr. Greg Seab for their knowledge, and for aid in developing my plan of study.

Lastly, I wish to express great gratitude to my family and friends for their support and encouragement.

Table of Contents

List of Figures

List of Tables

Abstract

The Littoral Acoustic Demonstration Center has collected passive acoustic monitoring data in the northern Gulf of Mexico since 2001. Recordings were made in 2007 near the Deepwater Horizon oil spill that provide a baseline for an extensive study of regional marine mammal populations in response to the disaster. Animal density estimates can be derived from detections of echolocation signals in the acoustic data. Beaked whales are of particular interest as they remain one of the least understood groups of marine mammals, and relatively few abundance estimates exist. Efficient methods for classifying detected echolocation transients are essential for mining long-term passive acoustic data. In this study, three data clustering routines using k-means, self-organizing maps, and spectral clustering were tested with various features of detected echolocation transients. Several methods effectively isolated the echolocation signals of regional beaked whales at the species level. Feedforward neural network classifiers were also evaluated, and performed with high accuracy under various noise conditions. The waveform fractal dimension was tested as a feature for marine biosonar classification and improved the accuracy of the classifiers. [This research was made possible by a grant from The Gulf of Mexico Research Initiative. Data are publicly available through the Gulf of Mexico Research Initiative Information & Data Cooperative (GRIIDC) at https://data.gulfresearchinitiative.org.] [DOIs: 10.7266/N7W094CG, 10.7266/N7QF8R9K]

Key terms: beaked whale, biosonar, classification, clustering, feedforward neural network, fractal dimension, passive acoustic monitoring

1. Introduction

The Littoral Acoustic Demonstration Center (LADC) consortium has collected passive acoustic monitoring data in the northern Gulf of Mexico (GOM) since 2001. Recordings were made in 2007 near the Deepwater Horizon oil spill that provide a baseline for an extensive study of regional marine mammal populations in response to the disaster. Detections of bioacoustic signals in the ambient recordings allow for the derivation of probabilistic estimates of regional species density. Understanding the dynamics of species abundance and distribution can offer insight into the function of the deep-sea ecosystem. Passive acoustic monitoring (PAM) has been commonly used for stock assessments and behavioral studies of echolocating marine mammals. However, the LADC Gulf Ecological Monitoring and Modeling (LADC-GEMM) effort marks the first step toward long-term monitoring of endangered species abundance based on acoustic recordings.

A number of marine mammals are known to use echolocation to perceive their environment acoustically. These include the beaked whale family (*Ziphiidae*), which are of particular interest to LADC. They have remained one of the least understood groups of marine mammals due to a deep-sea habitat, long dive durations, and an apparent low abundance (Bianucci et al., 2008). Many approaches to automated detection and classification for commonly assessed cetaceans such as sperm whales and dolphins have been developed (Kandia & Stylianou, 2006; Parada & Cardenal-López, 2014; Bittle & Duncan, 2013). However, this is less true for beaked whales. Although several methods of family-level detection exist (Yack, 2010; Parnum et al., 2011), high accuracy species classification methods have remained undeveloped. In this work, several clustering algorithms are shown to distinguish regional beaked whale species with high accuracy based on

selected biosonar features. Furthermore, trained neural network classifiers are evaluated using these features and perform with high accuracy under various noise conditions.

Automated classification of acoustic transients is a priority in passive animal bioacoustics research. Various spectral, temporal and geometric parameters of echolocation clicks are presently considered in determining species-level acoustic signatures. As a classification feature, the fractal dimension (FD) has remained notably absent in marine mammal studies. Many useful applications of fractal analysis have been developed for signal and image processing problems in fields such as biomedical engineering and acoustics (Gómez et al., 2009; Maragos & Potamianos, 1999; Al-Kadi & Watson, 2008). Dimension estimates serve as a complexity measure and have been adopted as a useful tool for detection and classification problems. In this study, we consider algorithms for waveform dimension measure developed by Higuchi (1988), Katz (1988), Henderson (1998), and Castiglioni (2010). The fractal measures are tested against the Shannon and Rényi entropies (Shannon, 1948; Rényi, 1961), which were introduced as bioacoustic features for classification by Han et al. (2011). We also find several spectral and temporal parameters to be useful in distinguishing the echolocation clicks. In the interest of scaling acoustic monitoring studies to large quantities of long-term data, the cost of classification should be minimal. The developed methods use low-dimensional feature vectors. They also perform with higher accuracy and require less computation time compared to a basic wavelet decomposition approach to classification that is tested.

Multi-stage classification routines are currently used in processing long-term data by LADC. Iteratively lower levels of discrimination (with respect to the taxonomical hierarchy) are achieved on resulting subsets of the data. For example, transients from marine mammals may first be classified at the family level (beaked whale vs. dolphin); then the family subset is again

2

considered for species-level classification. The species studied here are Gervais' and Cuvier's beaked whales, as well as a yet unknown species, presumed to be in the beaked whale family, that is regularly recorded in the GOM and referred to as BWG (Beaked Whale Gulf) (Baumann-Pickering et al., 2013). These subsets can then be considered for identification of individual whales.

The range of acoustic encounters in the deep-sea environment has not been thoroughly described. Thus, unsupervised classification, or clustering, is of interest for revealing clusters in detection results corresponding to distinct species, families, or individuals. Self-organizing maps (SOMs) have previously been applied to the problem of unsupervised separation of overlapping click trains by clustering (J. Ioup & G. Ioup, 2004). Three algorithms are presently assessed which utilize SOMs, k-means, and spectral clustering. On the other hand, neural network classifiers trained in a supervised fashion are expected to be more robust to noise, and are capable of classification tasks which are not linearly separable. We show that high accuracy classification results are given by a feedforward neural network trained under various noise conditions, and that the inclusion of the fractal dimension improves performance.

2. Background

2.1 Passive Acoustic Monitoring

The clade of Cetaceans, consisting of whales, dolphins, and porpoises, use acoustic signals to perceive and interact with their environment, as well as for communication. While significant light does not typically penetrate beyond 200 meters (the euphotic zone) in seawater, acoustic signals from whales may travel several kilometers before experiencing significant transmission loss. Sound speed in the ocean is dominated by the thermocline (decreasing with depth) up to a

certain depth, at which point on it is determined by the pressure (increasing with depth). At the depth of minimal sound speed, the so-called deep sound channel (DSC) exists due to the downward and upward angles of refraction above and below the channel, respectively. Surface and bottom reflectivity in the ocean also create a waveguide, preserving much of the propagating acoustic energy within the water layer. Considering these characteristics of underwater sound, approaches based on acoustic recordings have been widely adopted in deep sea ecology research. Passive acoustic monitoring (PAM) concerns the collection of ambient acoustic data for monitoring applications, such as logging noise conditions, or relevant acoustic events. PAM methods are being increasingly used in ecological studies of regularly vocalizing or echolocating marine species. Methods include towed hydrophone arrays (Yack et al., 2013, Holmes et al., 2005), autonomous gliders (Baumgartner et al., 2013), and bottom-moored monitoring instruments (Sousa-Lima et al., 2013), such as the EARS (Environmental Acoustic Recording System) buoys used by LADC. In the interest of estimating species abundance, models of acoustic detection probability may be developed based on hydrophone sensitivity, oceanographic conditions, and known source parameters of target species (such as amplitude and power spectrum) (Frasier et al., 2016). Taking into account the geographic area of probable acoustic detection from a single or set of hydrophones, detection counts may be used to develop regional density estimates (Thomas et al., 2012).

Traditionally during LADC's research, fixed EARS moorings are deployed in water depths between 1 – 2 km. Two hydrophones separated by 1 meter are placed between an anchor and 10 – 12 glass floats, which allow for positioning in the expected depth range of foraging marine mammals. Continuous recordings are made at a 192 kHz sampling rate for approximately 100 days (limited by 4 Tb of hard disk storage per channel). The pre-amplifier output of the two

hydrophones is passed through a high-pass filter (20 Hz) and anti-aliasing low-pass filter (160 kHz). Incorporated hydrophones have a typical sensitivity of approximately -170 dB re 1V per µPa. Data are converted by ADC and stored as 16-bit integers in a proprietary binary format. For recovery, acoustic releases are utilized which detach the recording package from the anchor upon receiving a special acoustic message.

2.2 Beaked Whales

The beaked whale family, Ziphiidae, currently includes 22 known species, constituting nearly one fourth of Cetacean species. They are widespread, ranging from the ice boundaries at the poles, to the equator in all oceans (MacLeod et al., 2006). They are difficult to study visually due to a deep-sea habitat, and extreme dive profiles. For this reason, only a small subset of these species is reasonably well-known thus far, including Baird's, and Cuvier's beaked whales.

The length of beaked whales ranges from 4 to 13 meters, and their weight from 1 to 15 tons. Distinctive physical features are a protruding snout similar to dolphins, which inspires the family's common name, and the absence of a notch in the tail fluke. Males exhibit a pair of tusk-like teeth. Dives beyond 500 meters are typically performed by beaked whales to forage using echolocation (Zimmer, 2007). Cuvier's beaked whales have been recorded breaking dive records for mammals, surpassing depths of at least 2992 meters, and durations of at least 137.5 minutes (Shorr et al., 2014).

Their biosonar is characterized by high frequency clicks in the approximate range 5-95 kHz, with a chirp-like frequency upsweep. Inter-pulse-intervals have been recorded to range from 90 to 435 ms, and click durations from 182 to 779 microseconds (Baumann-Pickering et al., 2013).

2.3 Classification

For long-term studies, automated detection and classification are essential processes, since visual surveying of acoustic data becomes inviable. The problems of detecting and classifying time-series events are closely related, since detection algorithms generally use temporally local features to partition data into two classes: periods of detection and of no detection. In this work, the problem of classification is that of automatically grouping echolocation signals which have been previously extracted as segments from the hydrophone data.

Methods of classification can be generalized into two categories: supervised, and unsupervised. In the former approach, an algorithmic model is trained to recognize known classes. In unsupervised classification, also referred to as cluster analysis or clustering, the algorithm is designed to partition the data such as to reveal inherent groups of similar elements.

2.3.1 Feedforward Neural Network

Feedforward neural networks (Svozil et al., 1997) were tested as supervised classification models in this study. The feedforward neural network (FNN) is a biologically inspired computing system in the form of a directed graph, as the example in Fig. 1.
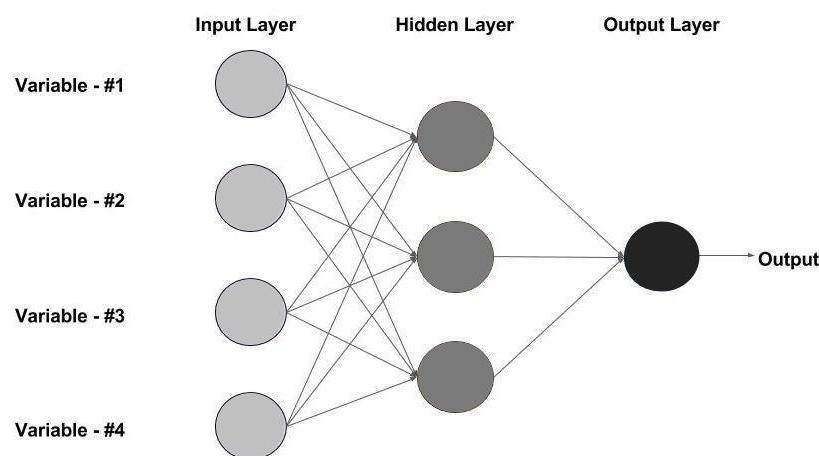
Fig. 1. A basic feedforward neural network with three nodes in the hidden layer, and a single output node.

The nodes are arranged in layers as shown above; each layer feeds data forward to the next layer in sequence. There are no recurrent connections. Variables representing various features of a given class are fed to the nodes of the input layer, which forms a complete bipartite graph with the next layer, named the 'hidden layer'. In general, more than one hidden layer may exist, the arrangement of connections may differ, and multiple outputs may exist. A single layer was used in this study, and a single output node which provided the class estimate.

In the FNN model, each network link is assigned a weight, which is adjusted during the learning process. Each node of the hidden layer models a biological neuron, and is a basic processing unit of the network.
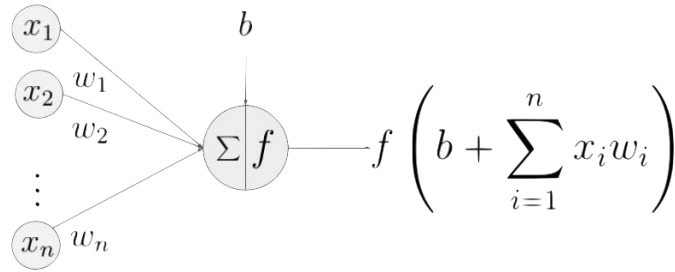
Fig. 2. An artificial neuron, with the array of inputs $(x_1, \ldots x_n)$, their respective weights $(w_1, \ldots w_n)$, a bias b, and the activation function f.

A neuron of the hidden layer first computes the weighted sum of its inputs and then applies an activation function which normalizes the sum. A nonlinear activation function qualifies a feedforward neural network with a single hidden layer as a universal function approximator (Cybenko, 1989). It also prevents the growth of output values for deep networks, with multiple hidden layers. For the model evaluated presently, the hidden layer nodes used a computationally efficient approximation of the hyperbolic tangent function:

$$S(n) = \frac{2}{1 + e^{-2n}} - 1$$

which satisfies the demand for nonlinearity. As depicted in Fig. 2, each neuron in the hidden layer is shifted by a bias, which provides a constant value of 1. By adjusting the bias weight, the entire output of a neuron may be shifted, which increases the range of possible learned functions for the network. It also allows for classification when all inputs are equal to 0.

For multiclass classification problems such as that faced in this study, a softmax function is used by the output node of the network. The softmax function can be defined, for an N-dimensional vector $x(n)$ of real values,

$$\sigma(x)_m = \frac{e^{x_m}}{\sum_{n=1}^{N} e^{x_n}} \quad \text{for } m = 1, \dots, N$$

It is a generalization of the sigmoid transfer function, and normalizes the N-length array of predictions to an equal length array of real values in the range [0, 1] that sum to 1. The class corresponding to the largest element of the array is chosen as the predicted class.

A dataset with known corresponding outputs is prepared for training. The weights associated with each link are initially randomized, then inputs are iteratively fed through the network to the output layer. In each iteration, the output is compared with the ground-truth output by an error function such as the mean squared error (MSE), or cross-entropy. Cross entropy between two discrete probability distributions p, and q may be defined

$$H(p, q) = -\sum_{x} p(x) \log q(x)$$

For neural network training, the discrete distributions are the class probability distributions that are output by the network. A gradient descent algorithm is then applied in the error function space. As mentioned, the weights are iteratively adjusted during the learning process. The total error E has a nested dependency on the weights of the network. The final outputs are dependent on the

hidden neuron outputs and weights, which are further dependent on the inputs and input weights. Applying the chain rule

$$(f \circ g)' = (f' \circ g) \cdot g'$$

the gradient of the error function with respect to each network weight can be computed. Then the weights may be adjusted by

$$\Delta w_{ij} = -\frac{\eta \partial E}{\partial w_{ij}}$$

where w_{ij} is the weight of link i, j and E is the error. The weight will be adjusted by the product of the learning rate η and the gradient, multiplied by -1. The parameter η, the learning rate, may be adjusted throughout the learning process for improved optimization. In this way, the gradient descent algorithm seeks to find a network weight configuration which minimizes the output of the chosen error function.

2.3.2 Clustering

Supervised classifiers can only make predictions based on a set of known classes. Cluster analysis concerns the general problem of automatically partitioning data in a meaningful way. It has been used previously in acoustic monitoring research (Bittle et al., 2013; Stimpert et al., 2011; J. Ioup & G. Ioup, 2004). Clustering models are not trained with data of known classes; they are designed to partition the data such that members of each class are more similar (by a chosen measure) to each other than to members of other classes. This is a useful technique in passive acoustic monitoring, since a wide range of acoustic events are expected. Clusters may be automatically revealed which correspond to distinct marine mammal families, species, or individuals, based on the parameters of the algorithm. Here, we compare three different approaches based on their ability to distinguish beaked whale species over a range of ambient noise levels. As

a baseline, the most widely used algorithm in data clustering, k-means, is implemented. This is compared to approaches based on the self-organizing map, and spectral clustering.

k-means Algorithm

The k-means algorithm (MacQueen, 1967) partitions a set of n observations into k classes, and requires the parameter k as input. It aims to find

$$\min_{S} \sum_{i=1}^{k} \sum_{x \in s_i} \|x - \mu_i\|^2$$

for $X = \{x_1, x_2, \dots, x_n\}$ a set of d-dimensional vectors; $S = \{s_1, s_2, \dots, s_k\}$ a set of clusters, each a list of which vectors it contains; μ_i, the mean of the members of s_i; and $\|x - \mu_i\|^2$, the squared Euclidean distance between a vector x and μ_i. In words, it aims to find the set of clusters S such

---

begin
- Initialize k cluster centers $c_1, c_2, \dots, c_k$ at randomly selected points in the d-dimensional space
- Compute pairwise distances between the data vectors and the cluster centers
- Assign each vector to the nearest cluster center

while (1) {
- Adjust the positions of the cluster centers using

$$c_i = \mu_i = \frac{1}{n_i} \sum_{x \in S_i} x_i$$

      where n_i is the number of vectors in cluster s_i
- Compute pairwise distances between the data vectors and the cluster centers
- Assign each vector to the nearest cluster center

if (number of re-assigned vectors equals 0) {
- **end**

    }
}

Fig. 3. Pseudo-code for the k-means algorithm

that the sum of squared Euclidean distances between the observations of each cluster is minimized. The algorithm cannot reveal overlapping clusters; each data point belongs to a single class. Pseudo-code for the algorithm is provided in Fig. 3.

Self-Organizing Maps

The self-organizing map (Kohonen, 1982) is a network model which uses unsupervised learning to produce a low-dimensional representation of the input data. Graph topology differs amongst applications, however a two-dimensional network with a square or hexagonal grid topology is common, as shown in Fig. 4. Each node is assigned a weight vector, determining its position in the data space. The training algorithm typically iterates over the data points, or, for larger datasets, randomly samples them. At each iteration, the weight vector of the nearest node, termed the Best Matching Unit (BMU), is updated in a similar way as the cluster centers of the k-means algorithm. However, as each BMU is adjusted, its neighborhood will also be adjusted. A neighborhood function Θ dictates the relationship between the magnitude of adjustment and distance from the BMU. The magnitude of adjustment is also typically dependent on a learning rate function α. The adjustment formula for a node n with weight vector $\boldsymbol{W}_n(i)$ at iteration i then has the form

$$\boldsymbol{W}_n(i + 1) = \boldsymbol{W}_n(i) + \Theta(u, v, i)\alpha(i)\big(\boldsymbol{D}(i) - \boldsymbol{W}_n(i)\big)$$

Here, $\boldsymbol{D}(i)$ is the input vector of the currently sampled data point, and u is the index of the current BMU.

The primary interest in using self-organizing maps for this study is in their ability to generate a smaller representation of the input data containing fewer outliers. In the method presently evaluated, each input vector is assigned to the nearest map node after mapping, then k-means is applied to the map nodes. In this way, a two-level hierarchical clustering is formed. The
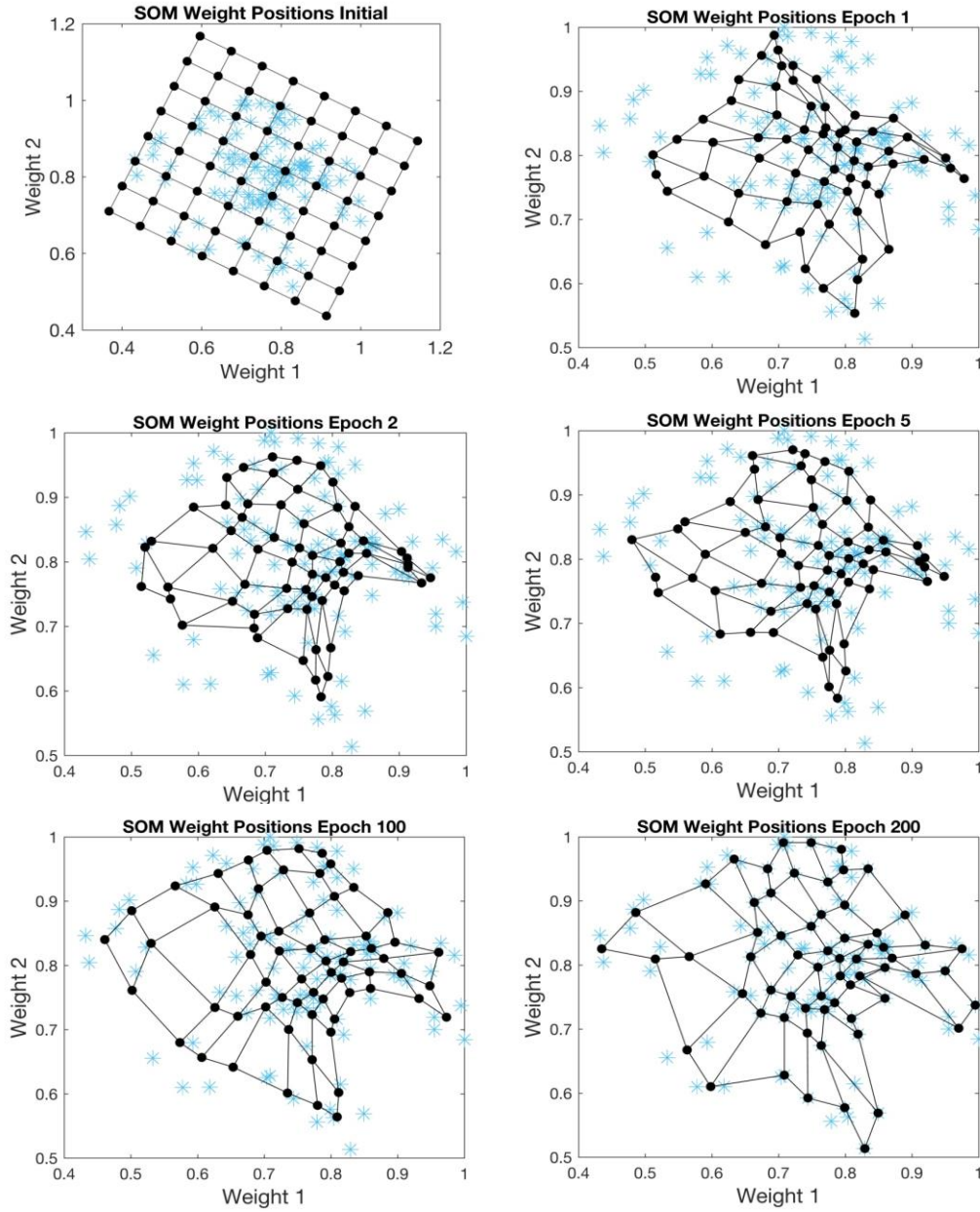
11

Fig. 4. An example of self-organizing map node positions at various iterations of training. Blue asterisks represent the example data points. Over an epoch of training, the entire set of data points is sampled by the algorithm. The effect of the learning rate function can be seen in that the adjustments become small after many epochs.

set of input vectors assigned to nodes within a given k-means cluster are then assigned to that cluster. This approach reduces the influence of outlying input vectors on the partitioning compared to k-means alone, which we speculate will lead to more meaningful clusters.

Spectral Clustering

Spectral graph theory addresses the relationship between graph structure and the eigenvectors and eigenvalues of the associated adjacency and Laplacian matrices. For an n-length set of data vectors, the $n \times n$ adjacency matrix A contains pair-wise distances between the vectors, such that $A_{ij} \geq 0$. The diagonal degree matrix is then defined

$$D_{ii} = \sum_j A_{ij}$$

The Laplacian matrix of an undirected graph $G = (V, E)$ can be defined

$$L_G(u, v) = \begin{cases} d(u) & \text{if } u = v \\ -1 & \text{if } u \neq v \text{ and } (u, v) \in E \\ 0 & \text{otherwise} \end{cases}$$

$$= D_G - A_G$$

where D_G and A_G are the degree and adjacency matrices of the graph, respectively. For an undirected graph, the Laplacian matrix will be symmetric and singular. Derived from Cheeger's inequality (1970), the sparsest cut of a graph may be approximated by the second smallest eigenvector of the graph Laplacian. A transformed space can be created by the Laplacian eigenvectors in which previously non-linearly separable classes may become linearly separable. In traditional spectral clustering methods, the first several (d) singular vectors of the graph Laplacian are found and used to create an $n \times d$ matrix as a transformed space (for G with n vertices) in which rows are clustered using a partitioning method such as k-means.

3. Objective

The current goal is to develop and evaluate a set of computational methods for distinguishing beaked whale echolocation signals at the species level. These methods may aid future acoustic monitoring research of beaked whale populations. Specifically, we wish to develop efficient methods of classifying three beaked whale species regional to the Gulf of Mexico, which have been detected in passive acoustic monitoring data. An essential aspect of the development is the evaluation of various signal features for their potential in quantifying unique acoustic signatures. The most relevant feature sets will be used to assess the performance of feedforward neural network classifiers, as well as several unsupervised classification approaches.

4. Methods

4.1 Data Collection

More than 32 TB of data were collected from June to October of 2015 at ten locations in the northern GOM. Data used for this evaluation were taken from two deployments (Fig. 5). Processing routines were developed in MATLAB and Bash. Recordings were made at a 192 kHz sampling rate approximately 530 meters above the seafloor. A high-level survey of the entire 2015 monitoring dataset for transient acoustic events with frequencies within 3-90 kHz has been performed by LADC members using a primarily automated routine. Metadata is available for each buoy listing files determined to contain clicks spanning any set of three frequency bands. To collect signals from each species, server files with bioacoustic events spanning the same bands as a target species' known dominant frequency range were randomly sampled. From these files, events of a desired species were automatically detected by matched filtering with a template echolocation signal. Specifically, the cross-correlation of the template and the file waveform was taken and

segments of the waveform were extracted centered on peaks of the correlation which surpassed a visually chosen threshold. The cross-correlation for two discrete functions f, and g is defined

$$(f \star g)[n] = \sum_{m=-\infty}^{\infty} f^*[m]g[m+n]$$

where f^* denotes the complex conjugate of f.
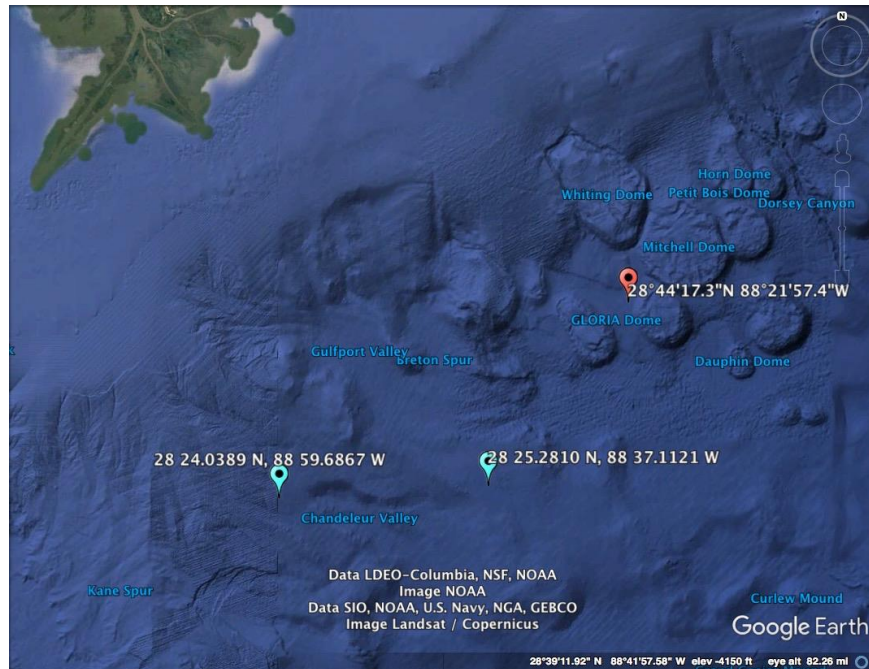
Fig. 5. Locations of EARS buoys from which data were collected (blue tacks), and of Deepwater Horizon (red tack). The eastern and western buoys are approximately 27 and 44 miles from the Deepwater Horizon site, respectively.

These templates were processed using the routine explained below. Prior to running the detection algorithm, raw files were digitally filtered with a fifth-order Butterworth bandpass filter with cutoff frequencies 15 and 95 kHz. This is the expected band of the regional beaked-whale echolocation signals as determined by Baumann-Pickering et al. (2013). Using this matched filtering method, echolocation patterns could be detected at very low signal-to-noise ratio (SNR).

Frames were extracted at a constant buffer length of 2.1 milliseconds, which was expected to contain regional biosonar clicks (Baumann-Pickering et al., 2013). Upon the detection of a desired species in the sampled files, files with temporally preceding and following data were scanned. Programs for efficient visual surveying of time series and spectrogram data (Hann window, 1024-point FFT, 98% overlap) were developed for validation. The extracted frames were stored in a MATLAB structure array with fields containing time, amplitude, and source file information. A noise estimate was also recorded for each extraction, which was taken as the first 3-13 milliseconds of the respective source file which spans 21.33 seconds. In this format, the size of relevant echolocation data for long-term studies at a typical sampling rate (190 – 200 kHz) is easily manageable, with files containing hundreds of events being only a few megabytes in size, depending on the frame size.

4.2 Preprocessing

Data from acoustic encounters of three beaked whale species (Cuvier's, Gervais', BWG) were used in the evaluation. Quantification of some biosonar source properties requires the range and bearing of the source signal from the receiver to be known. These parameters are not assumed to be known and are not presently considered in classification. Methods of click localization from multiple hydrophone array data are in development and could allow for modeling of source properties.

The extracted frames were Hann windowed, then noise removal was performed by wavelet thresholding. Five-level wavelet decomposition coefficients were calculated with the Fejer-Korovkin length-8 (fk8) wavelet (Nielsen, 2001), and hard minimax thresholding applied (Donoho & Johnstone, 1994). Framed signals were normalized by the maximum of their absolute value. Further analysis was done on these processed frames.

4.3 Signal Features

4.3.1 Spectral Analysis

Spectral features were derived from a 1024-point Fast Fourier Transform (FFT) periodogram. The FFT is a computationally efficient algorithm for computing the Discrete Fourier Transform (DFT) of a complex sequence $x_0, \ldots, x_{N-1}$. The DFT is defined

$$X_k = \sum_{n=0}^{N-1} x_n e^{-i2\pi kn/N} \quad k = 0, \ldots, N-1$$

Features considered in the evaluation include peak frequency (f_p); -10 dB bandwidth (BW_{-10dB}); and spectral centroid (f_c), calculated as the weighted mean of the FFT spectrum

$$f_c\big(x(n)\big) = \frac{\sum_{n=0}^{N-1} f(n)x(n)}{\sum_{n=0}^{N-1} x(n)}$$

where $f(n)$ is the center frequency of bin n, and $x(n)$ is its magnitude. Also tested was the -20 dB center frequency (f_{ctr}), taken as the mean of frequencies with magnitude surpassing 20 dB below f_p.

4.3.2 Click Duration

Measures of transient duration have been developed specifically for biosonar research. The 95% energy duration (τ_E) was calculated as presented by Madsen & Wahlberg (2007). Prior to the calculation of τ_E, the signal was interpolated by a factor of 5 using the MATLAB 9.1 low-pass interpolation routine with a symmetric filter. The 95% squared-amplitude duration (τ_A) was calculated by taking the end points at 5% of the peak squared amplitude as the onset and offset times. The same method was also tested using the absolute value of the Teager-Kaiser energy (Kaiser, 1990) in place of the squared-magnitude signal (95% TK amplitude duration, τ_{TA}). The Teager-Kaiser energy operator $\Psi\big[x[n]\big]$ is a non-linear operator which provides nearly

instantaneous tracking of high-amplitude, high-frequency events. The operator is defined, for discrete signals

$$\Psi[x[n]] = x^2[n] - x[n-1]x[n+1]$$

The -10 dB duration ($\tau_{-10\text{dB}}$) was found to be the most useful for this study. The -10 dB end points relative to the envelope peak were taken as the onset and offset times (Moehl et al., 1990). The envelope of the framed click was taken as the absolute value of its analytical signal – consisting of the raw signal as its real part and the Hilbert-transformed signal as its imaginary part (Fig. 6).
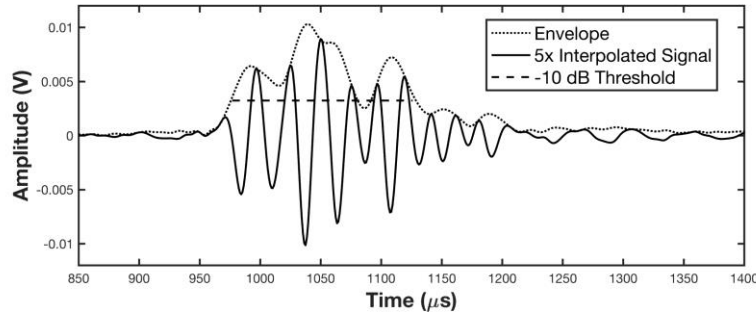
Fig. 6. An example Cuvier's beaked whale click (interpolated by a factor of 5) with the signal envelope (dotted line) and -10 dB threshold (dashed line).

4.3.3 Fractal Dimension and Entropy

Dimension estimates tested included that of Higuchi, Katz, Castiglioni, and Henderson. Applications of the fractal dimension include approaches in the time and phase-space domains. The former regards the time series itself as a geometric object. In the phase-space approach, the time series is considered as a set of observations of a dynamical system, and is embedded in an approximated phase-space domain. This constructed trajectory is then the geometric object considered. Estimates made directly in the time-domain are generally less computationally costly. In either case the FD characterizes the complexity of the set as a ratio of the change in detail to the

change in measuring resolution, or scale. The methods tested here used the waveform directly. FD estimates were computed for the processed frames after interpolation by a factor of 5. The accuracy of the FD algorithms in approximating theoretical values such as for Weierstrass functions is dependent on the number of samples of the curve. The 2.1 millisecond frames consisted of 401 samples without interpolation, which resulted in higher variance of the FD measurements than for the interpolated signals, and in more cases values slightly exceeding the theoretical limit for curves of topological dimension 1 (a fractal dimension of 2).

Higuchi's Algorithm

Consider the time series to be analyzed

$$x(1), x(2), \ldots, x(N)$$

Create k new time series x_m^k as

$$x_m^k = \left\{ x(m), x(m+k), x(m+2k), \ldots, x\left(m + \left\lfloor \frac{N-m}{k} \right\rfloor k\right) \right\}, \quad \text{for } m = 1, 2, \ldots, k$$

Here, m is the initial time value, and k is the discrete time interval, or scale. The effective measurement resolution is determined by k. For each x_m^k constructed, the average length $L_m(k)$ is computed as

$$L_m(k) = \frac{\beta}{k} \sum_{i=1}^{\left\lfloor \frac{N-m}{k} \right\rfloor} |x(m+ik) - x(m+(i-1)k)|$$

where $\beta = (N-1)/\lfloor (N-m)/k \rfloor k$ is a normalization factor. The average value $\langle L(k) \rangle$ of the k lengths $L_m(k)$ is taken. This is repeated for a set of varying time intervals. If the average value follows a power law:

$$\langle L(k) \rangle \propto k^{-D}$$

then the curve is said to be fractal with dimension D. To observe the relationship between time interval k and the curve length $L(k)$, k was varied from 1 to a chosen value k_{max}. The slope of the least squares linear best fit of the $\ln(L(k))$ vs. $\ln(1/k)$ curve is taken as the dimension estimate.

To choose k_{max} for this study, the value was varied from 10 to 100 in intervals of 5, and k-means clustering purity (Zhao & Karypis, 2002) was calculated for each. Peaks in clustering purity were observed at values of 40 and 60. Between these values there was an approximate 2% increase in purity and 36% increase in computation time, thus a value of $k_{max} = 40$ was chosen for this study.

Katz's Algorithm

The fractal dimension as proposed by Katz (1988) can be expressed

$$D = \frac{\log_{10}(n)}{\log_{10}(n) + \log_{10}(d/L)}$$

where the number of points in the sequence is $N = n + 1$, and

$$L = \sum_{i=1}^{n} l_{i,i+1}$$

$$d = \max\{l_{1,j}\}$$

for

$$l_{i,j} = \sqrt{(y_i - y_j)^2 + (x_i - x_j)^2}$$

where L is the sum of distances between successive points in the waveform and d is called the diameter, defined as the maximum of distances between the first point of the sequence and all other points.

Castiglioni's Algorithm

Castiglioni (2010) claimed that Katz's algorithm is flawed in that it is influenced by the unit of measurement. He proposed a variation in which d and L are redefined:

$$L = \sum_{i=1}^{n} |y_{i+1} - y_i|$$

$$d = \max\{y_k\} - \min\{y_k\}$$

Notice that d is now simply the range of the set, and the sampling interval is not involved in the calculation of L.

Adapted Box Dimension

The adapted box dimension (ABD) developed by Henderson et al. (1998) is found by dividing a time series of N samples into sets of Δt samples. The range ("extent") of each segment is calculated. The mean extent $E(\Delta t)$ is then taken over a range of Δt. The dimension is then computed by finding the best fit to the following equation:

$$A(\Delta t) = NE(\Delta t) \approx A_0 \Delta t^{2-D}$$

Shannon and Rényi Entropy

Han et al. (2011) introduced the Shannon and Rényi entropy information as bioacoustic features for frog classification. The Shannon entropy can be expressed as

$$H = -\sum_{i} p_i \log_2 p_i$$

where p_i is the probability of occurrence of element i in the set of unique values in the sequence or signal (unique voltage readings for EARS signals). It is the expected information content of a sequence or signal. Similarly, the Rényi entropy of order $\alpha \geq 0$ is defined as

$$H_\alpha = \frac{1}{1-\alpha} \log_2 \left(\sum_{i} p_i^\alpha \right)$$

Following the study of Han et al., an order of $\alpha = 3$ is chosen for the Rényi entropy estimation. The Rényi entropy generalizes the Shannon entropy and as the order is increased, the value is increasingly dependent on the highest probability events.

4.3.4 Wavelet Decomposition

A common representation of data for feature selection is given by the discrete wavelet transform, which decomposes a signal into an approximation (low) and detail (high) frequency band representation by a quadrature mirror filter pair. Wavelet packet decomposition (WPD) refers to the iterative application of a wavelet transform to each resulting sub-band representation, creating a binary tree. Energy maps of the WPD tree were tested against other feature vectors as input for feedforward neural networks. Here, the energy map refers to the percentages of energy contained in the terminal nodes of the decomposition tree relative to the sum of their energies. Six-level trees were computed using a length-4 Daubechies, or D4 wavelet.

4.4 Feature Subset Selection

In the interest of minimizing the computational cost of feature extraction necessary for high accuracy classification, a subset of the investigated features with relative predictive power was chosen. The strengths of individual features in their ability to isolate pairs of species were investigated using unpaired two-tail t-tests assuming unequal variances. Data visualization tools were also developed for this purpose. To quantitatively validate the choice of features, a sequential forward selection (SFS) algorithm was applied, which maximized the classification accuracy determined by fitting multivariate normal distributions to each class. The algorithm iteratively adds to an empty set of features those which increase the accuracy by the maximum amount at each step. Termination occurs when no further features increase the accuracy. To quantify the

response of individual features to noise, feedforward neural networks as described in the above section were trained using individual features and tested on datasets of various mean SNR.

4.5 Clustering Algorithms

The three methods evaluated made use of k-means, spectral clustering, and self-organizing maps. Cluster purity was used to quantify the accuracy of the clustering. In computing cluster purity, the value for each cluster is often weighted by its size: the number of data points in the cluster. However, assuming that each cluster has a unique modal class, the unweighted cluster purity will give the fraction of observations that are correctly assigned. Here, the correct class assignment of an observation is defined as the cluster's mode class. Unweighted cluster purity can be expressed

$$\text{purity}(\mathbb{Q}, \mathbb{C}) = \frac{1}{n} \sum_j \max_i |\omega_j \cap c_i|$$

where $\mathbb{Q} = \{\omega_1, \omega_2, \ldots, \omega_m\}$ is the set of clusters, and $\mathbb{C} = \{c_1, c_2, \ldots, c_I\}$ is the set of classes. To compute purity, each cluster is assigned to the most frequent class of its objects, then the number of correctly assigned objects is divided by n. The implementations developed for this study check for unique modal classes of the clusters. Non-unique modal classes indicate a poor clustering job, yet can result in relatively good purity and entropy values. Validity measures are only reported here for results in which the clusters have unique modal classes.

Although the appropriate number of clusters k was known to be three for this study, automated selection of this value is of interest for mining passive acoustic data. The appropriate number was chosen by sweeping k from 2 to 10, and selecting the value which yields the minimum Davies-Bouldin index (Davies & Bouldin, 1979).

For spectral clustering, 15 neighbors were used in the generation of a k-nearest neighbors similarity graph, and the normalized graph Laplacian was calculated as

$$L = D^{-1}(D - W)$$

where D is the degree matrix, and W the k-NN similarity graph. Eigenvectors corresponding to the k smallest eigenvalues of the graph Laplacian were concatenated to form the space $U \in \mathbb{R}^{n \times k}$. Here, k is the number of clusters. Rows of U were then considered as transformed data vectors and clustered using *k-means*. Initialization of centroid positions for k-means was done by the k-means++ algorithm (Arthur & Vassilvitskii, 2007), and a squared Euclidean distance measure was used.

Self-organizing maps of dimensions 10×10 were applied to the data. The networks had grid topology, and a Manhattan distance metric with MSE performance was used for the mapping. The nodes come to represent local averages of the data, and are thus less sensitive to random variations than the original data (Vesanto & Alhoniemi, 2000). Resulting node positions were then considered for clustering by k-means. Map nodes with no assigned data points were removed.

4.6 Feedforward Neural Network

Feedforward neural networks with a single hidden layer were trained on randomized subsets of the data using scaled conjugate gradient backpropagation (Møller, 1993). Cross-entropy was used to measure performance. The training data spanned different levels of simulated noise. Several hidden layer sizes were tested with the feature vectors. A size of twenty nodes was chosen. Input and hidden layers each used a hyperbolic tangent sigmoid transfer function with biasing. A softmax transfer function was used by the output node for classification purposes. The accuracy was taken as the percent of correctly classified observations in the test data.

Training data were feature vectors computed from 250 processed clicks. Fifty were taken randomly from the original data and from data at each simulated noise level. Training data were

excluded from subsequent testing, which consisted of classifying the remaining 290 clicks for each dataset and computing accuracy as the percentage of correctly classified data vectors.

4.7 Noise Simulation

There are inherent signal-to-noise ratio differences in the observations dependent on the ambient noise, source bearing and range, and amplitude. To measure performance under a controlled range of noise conditions, simulated noise was added to the extractions (Fig. 7). This was done by filtering the source file noise estimate through a band-pass filter of cutoff frequencies
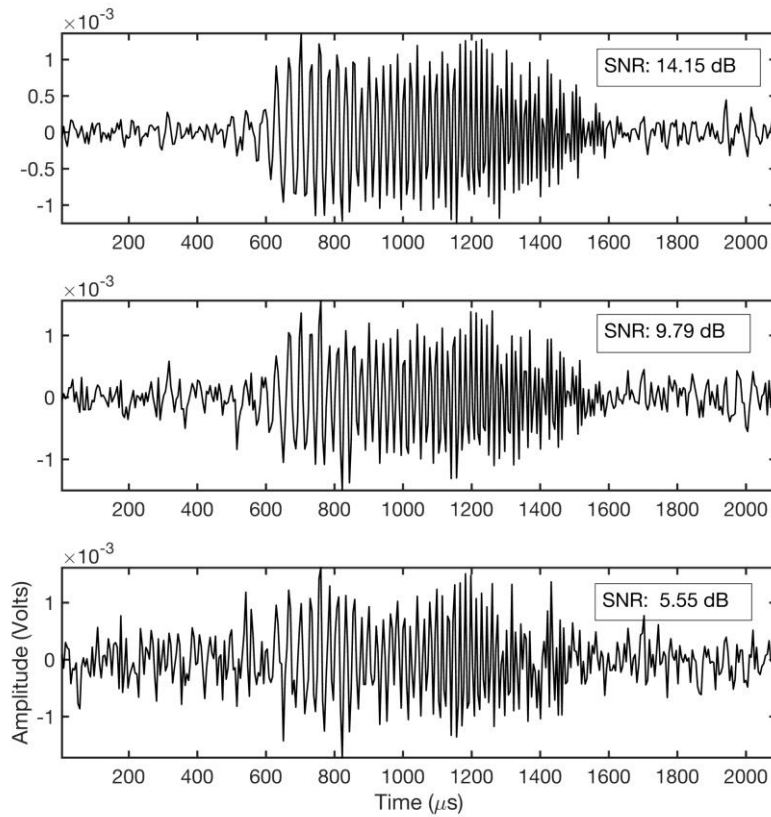

Fig. 7. An example BWG click with two levels of simulated noise corruption and computed SNR.

15 and 95 kHz, and calculating the standard deviation. Then random Gaussian noise of equal standard deviation was added to the click and to the filtered noise estimate. Signal-to-noise ratio (SNR) was computed as the ratio of the summed squared magnitude of the signal to that of an equal length segment of the noise estimate.

The original dataset of clicks was duplicated, and the clicks of the copied dataset were corrupted with noise using the process described above. A second copy was then made of the noise corrupted clicks, which was further corrupted. This process was iterated four times to create datasets of different average noise level.

5. Results

For the analysis, 111 BWG clicks, 110 Cuvier's BW clicks, and 119 Gervais' BW clicks were extracted from the EARS data. See Table 2 for a summary of the computed biosonar parameter statistics for each species. Except for the Castiglioni fractal dimensions (CFD) calculated for Gervais' and Cuvier's whales, the fractal dimension measures differed significantly for each pair of species ($p < 0.05$) as determined by the unpaired t-tests (Table 1). This supports the potential usefulness of fractal dimension measures as signal classification features.

	Higuchi FD	Castiglioni FD	Adapted Box FD	Katz FD	Shannon Entropy	Rényi Entropy
BWG vs. Cuvier's	5.13E-87	6.71E-93	9.56E-07	3.19E-44	2.95E-63	1.46E-33
BWG vs. Gervais'	2.38E-35	2.51E-99	2.76E-27	2.17E-41	1.31E-83	5.20E-57
Gervais' vs Cuvier's	2.04E-58	9.64E-01	3.03E-54	4.93E-22	2.30E-14	3.76E-09

Table 1. p-values for each fractal dimension measure between species pairs.

Parameter	BWG	Cuvier's BW	Gervais' BW
	Mean (*S.E.*; **Range**; N=111)	Mean (*S.E.*; **Range**; N=110)	Mean (*S.E.*; **Range**; N=119)
-10 dB bandwidth (Hz)	217.49 (*5.01*; **106-322**)	138.82 (*1.53*; **92-174**)	165.50 (*5.42*; **73-305**)
-20 dB center freq. (kHz)	50.33 (*0.270*; **43.7-56.6**)	35.03 (*0.250*; **30.2-43.0**)	53.05 (*0.435*; **41.4-62.0**)
Spectral centroid (kHz)	48.21 (*0.365*; **40.3-58.1**)	36.66 (*0.170*; **32.8-40.7**)	49.27 (*0.307*; **43.9-56.8**)
-10 dB duration (μs)	840 (*11.72*; **505-1093**)	198 (*7.98*; **99-500**)	130 (*4.17*; **57-333**)
95% energy duration (μs)	766 (*9.86*; **473-960**)	303 (*12.2*; **168-731**)	191 (*10.3*; **108-573**)
95% TK amplitude duration (μs)	925.16 (*11.94*; **583-1151**)	246 (*11.48*; **141-823**)	163 (*6.02*; **94-401**)
Higuchi FD	1.82 (*0.0032*; **1.71-1.88**)	1.65 (*0.0037*; **1.57-1.74**)	1.76 (*0.0021*; **1.69-1.82**)
Castiglioni FD	2.05 (*0.011*; **1.79-2.27**)	1.48 (*0.004*; **1.41-1.65**)	1.48 (*0.006*; **1.41-1.69**)
Adapted Box FD	1.60 (*0.005*; **1.48-1.69**)	1.63 (*0.003*; **1.51-1.69**)	1.52 (*0.004*; **1.45-1.64**)
Katz FD	1+5.08e-4 (*1.89e-5*; **1+1.59e-4 – 1+9.57e-4**)	1+7.11e-5 (*1.14e-6*; **1+4.08e-5 – 1+1.15e-4**)	1+1.05e-4 (*2.78e-6*; **1+6.03e-5 – 1+2.12e-4**)
Shannon Entropy (bits)	3.68 (*0.018*; **3.20-4.05**)	2.32 (*0.043*; **1.38-3.62**)	1.82 (*0.043*; **0.99-2.95**)
Rényi Entropy (bits)	1.65 (*0.017*; **1.23-2.07**)	1.09 (*0.032*; **0.41-1.99**)	0.82 (*0.031*; **0.26-1.55**)

Table 2. Summary of click parameter statistics: mean, standard error (*S.E.*) and range. Standard error is of the mean, computed as the sample standard deviation divided by the square root of the sample size.
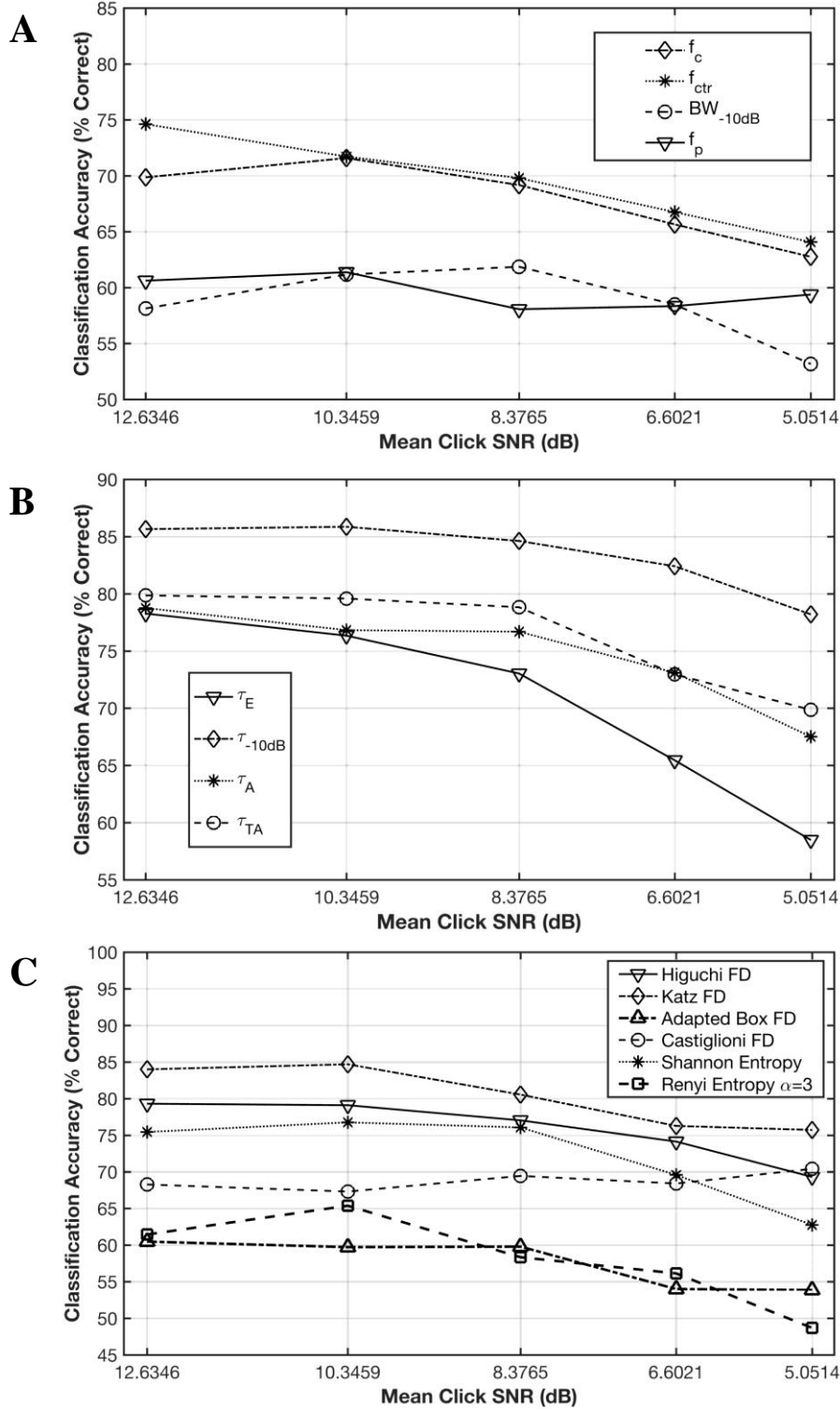
Fig. 8. FNN classification performances using individual features; (A) spectral parameters, (B) duration measures, (C) fractal dimensions, and Shannon and Rényi entropies.

As described previously, simulated noise was added to the original dataset of 340 clicks, and the resulting dataset was again corrupted with noise. This was done iteratively to create four additional datasets with progressively decreasing SNR (datasets 2-5). Including the original dataset, mean SNR values of the datasets spanned roughly 5.1 – 12.6 dB. See Fig. 8 for a summary of the individual feature performances in classifying observations of decreasing mean SNR.

Click spectra and duration have no practical correlation in this case, and vary amongst species. Waveform fractal dimensions generally have some correlation with central frequency measures. However, they will deviate in their correlation depending on the frequency distribution of the waveform and the FD algorithm. Pearson correlation coefficients amongst FD measures and spectral features range from 0.23 to 0.88. The Shannon and Rényi entropy show less correlation with coefficients spanning 0.04 to 0.33. However, Katz' and Higuchi's measures yield better accuracy as individual features for the trained networks. The sequential forward selection algorithm returned, as the optimal feature set, the parameters spectral centroid f_c, -10 dB duration τ_{-10dB}, and the Higuchi fractal dimension (HFD). The main disadvantage of the SFS algorithm is the inability to remove features from the set that become dispensable after the addition of other features. For this reason, networks were trained and tested using various combinations of these three parameters. The highest accuracy was still provided by the set chosen by SFS. A network was also tested using the Shannon entropy in place of the HFD, since the Shannon entropy showed little correlation with f_c and strong performance as an individual feature.

Clicks recorded at angles off-axis from the whale's forward bearing may increase variance in duration and central frequency measures. Madsen et al. (2004) reported recordings of Risso's dolphins at various angles off the presumed zero-degree axis (aligned with the dolphin's bearing) and observed a positive correlation between the angle and τ_E. Decaying low-frequency oscillations

were observed at the tail of the off-axis click recordings which increased τ_E and reduced f_c. Similar

patterns are noticed in the collected Cuvier's data, yet source range and bearing are not known.

There is an apparent negative correlation between peak received amplitude and this effect. The

biosonar characteristics of Cuvier's beaked whales may be more similar to that of Risso's dolphins

than other beaked whale species. The effect is not seen in low-SNR Gervais or BWG clicks. The

$\tau_{-10\text{dB}}$ measure is more robust to the effect than τ_E (Fig. 9). The -10 dB duration significantly

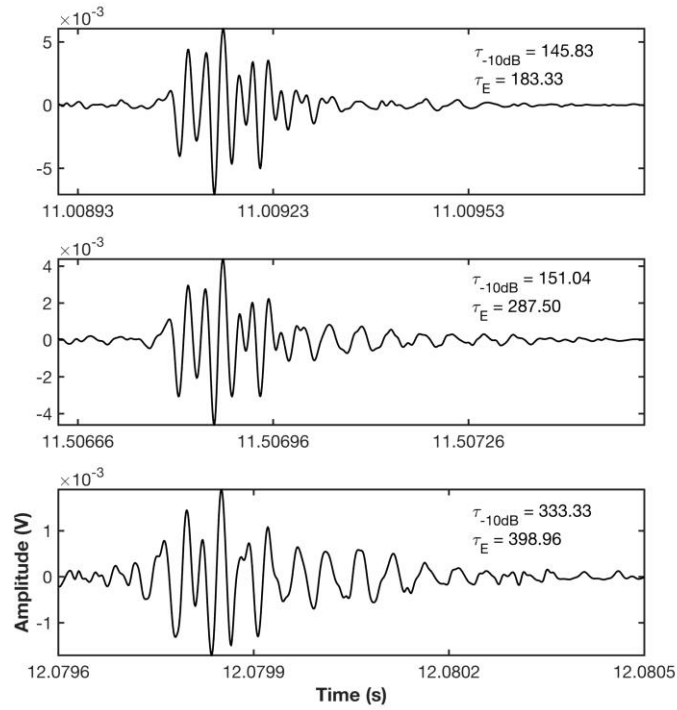outperformed the other duration measures as an individual feature.

Fig. 9. Three clicks from a Cuvier's beaked whale click train, presumably received at varying off-axis angles. Decaying oscillations of lower frequency appear at the tail of the click.

The -20 dB center frequency performed marginally better than the spectral centroid as an

individual feature, but the spectral centroid separated the Gervais' and Cuvier's observations more

strongly, complementing the strong separation of BWG from the other species provided by $\tau_{-10\text{dB}}$.

Katz' method yields slightly higher accuracy than Higuchi's as an individual feature, yet it strongly

separates BWG from the other species, which correlates with the contribution of $\tau_{-10\text{dB}}$ as a feature. The HFD on the other hand separates the species relatively evenly, and is thus more valuable in combination with f_c and $\tau_{-10\text{dB}}$.

For the evaluation of clustering algorithms, the HFD, $\tau_{-10\text{dB}}$, and f_c were used together as features. The results shown in Fig. 10 are average values over 10 trials of clustering the datasets. All algorithms yielded a purity of 1 for the first two datasets, with mean SNR values of 12.63 and 10.34 dB, respectively. That is, the algorithms completely isolated the species based on the three click features. Differences in performance between algorithms were marginal for all noise conditions except the last dataset, with a mean SNR of 5.05 dB. The SOM method outperformed the others in this case with a purity of 0.937, followed by spectral clustering at 0.921, and k-means at 0.909. Thus, for all noise conditions purity values were above 0.9, which corresponds to 90% accurate unsupervised classification. See Fig. 11 for a visualization of the feature space computed for dataset 1.

Results for FNNs trained using various feature sets are shown in Fig. 12. Results are shown for four networks. The first (N1) was trained using f_c, $\tau_{-10\text{dB}}$, and HFD, which was compared to another that substituted the Shannon entropy for the HFD (N2), and another which did not include either (N3). A fourth network (N4) was trained using the WPD energy map. Results shown in Fig. 12 are averaged over five trails of training and testing. Notable improvement of classification accuracy is seen with the inclusion of the HFD as a feature. N1 yielded the highest accuracy in all cases. All feature sets result in a similar trend of accuracy with increasing noise. A small rise in performance between trial 1 and 2 or 3 is seen for all networks, which is a consequence of most of the training data consisting of feature vectors from transients of non-optimal SNR. For all noise

conditions, N1, N2, and N3 show accuracy above 93%. N4 which used the WPD energy map as input was outperformed by the other networks in all cases.

5. Conclusions

High accuracy classification of several beaked whale species regional to the northern Gulf of Mexico is achieved by using spectral, temporal, and geometric features of recorded echolocation waveforms. The spectral centroid, -10 dB duration, and Higuchi's fractal dimension determine a strong acoustic signature. Katz's approach to measuring fractal dimension competes
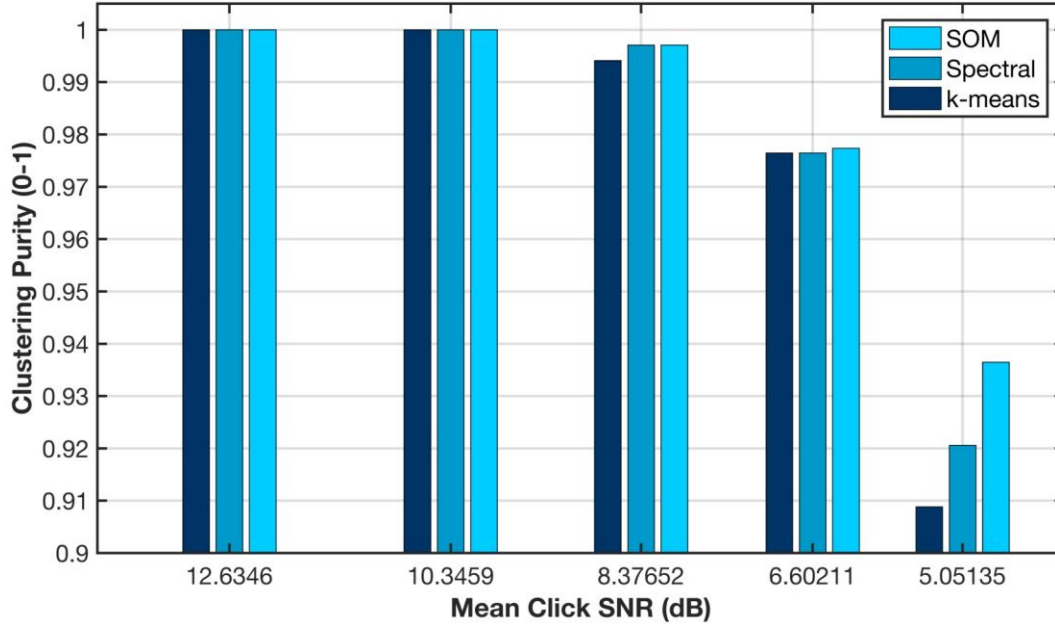

Fig. 10. Performance of clustering methods using three features: Higuchi FD, spectral centroid f_c, and -10 dB duration $\tau_{-10\text{dB}}$. Purity is unweighted and corresponds to the fraction of correctly assigned data points, where the correct class of a data point is defined as the mode class of its cluster.
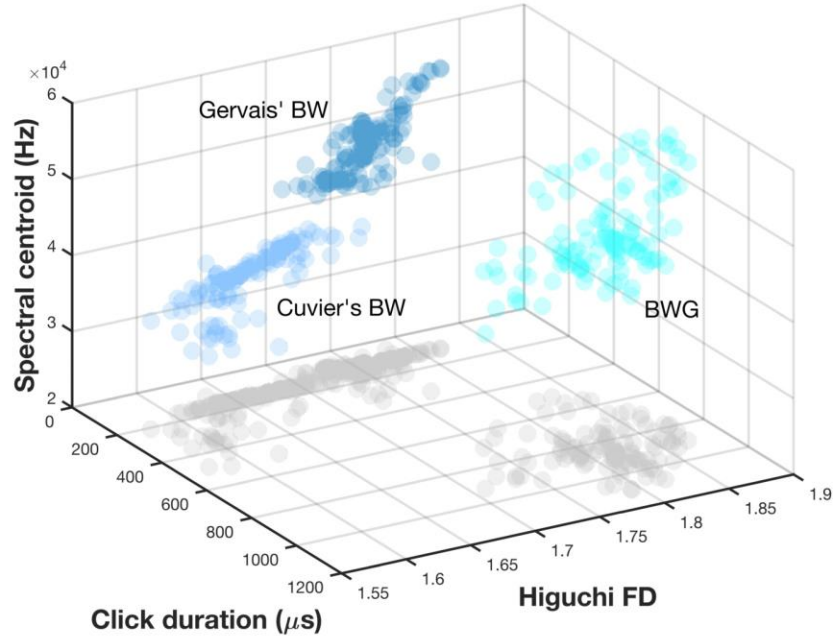
Fig. 11. Feature space visualization of f_c, $\tau_{-10\text{dB}}$, and HFD computed from the original, non-corrupted data. The clusters corresponding to each species are easily discerned.
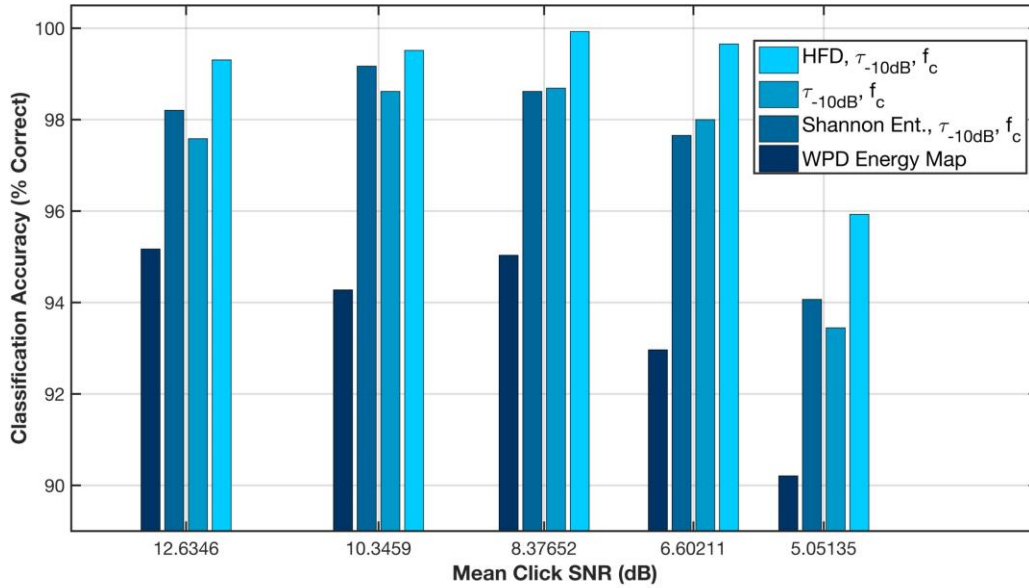

Fig. 12. Classification performance of feedforward neural networks trained using various inputs. The features are Higuchi's fractal dimension (HFD), spectral centroid (f_c), -10 dB duration ($\tau_{-10\text{dB}}$), Shannon entropy, and wavelet packet decomposition (WPD) energy map.

33

with Higuchi's method as an individual feature, yet separates BWG strongly from the other species. The spectral clustering and self-organizing map routines partition observations of the species with slightly higher accuracy than k-means for low SNR conditions. These methods have potential for automatically isolating observations of multiple species or individuals in acoustic data. Measured click duration and central frequency are dependent on source bearing. The dependency was observed most prominently for Cuvier's beaked whale clicks. Although signal source bearings were not known in this study, the wave behavior of Cuvier's beaked whale clicks at off-axis angles is hypothesized to have a stronger similarity to that of Risso's dolphins as compared to Gervais' and Cuvier's whales. FNNs trained under various noise conditions can classify the species effectively using f_c, $\tau_{-10\text{dB}}$, and HFD as features, with average accuracy over five test trials ranging from 96.0 – 99.9 % correct classification of datasets with mean SNR estimates spanning 5.05 – 12.6 dB. The inclusion of the Higuchi fractal dimension with the spectral centroid and -10 dB duration measure improved performance of the neural network classifiers compared to the Shannon entropy. Computing these features is less costly in terms of space usage and time than the WPD energy map, and using them for training resulted in classifiers with greater performance.

References

Al-Kadi O.S, and Watson D. **(2008)**. "Texture analysis of aggressive and non-aggressive lung tumor CE CT images," IEEE Transactions on Biomedical Engineering **55**, 1822–1830.

Arthur, D., and Vassilvitskii, S. **(2007)**. "K-means++: the advantages of careful seeding," in 2007 ACM-SIAM symposium on discrete algorithms (SODA'07).

Baumann-Pickering, S., McDonald, M.A., Simonis, A.E., Solsona Berga, A., Merkins, K.P., Oleson, E.M., Roch, M.A., Wiggins, S.M., Rankin, S., Yack, T.M., and Hildebrand, J.A. **(2013)**. "Species-specific beaked whale echolocation signals," J. Acoust Soc Am. **134**, 2293−2301.

Baumgartner, M.F., Fratantoni D.M., Hurst, T.P., Brown, M.W., Cole, T.V., Van Parijs, S.M., and Johnson, M.P. **(2013)**. "Real-time reporting of baleen whale passive acoustic detections from ocean gliders," J. Acoust. Soc. Am. **134**, 1814–23.

Bianucci, G., Post, K., and Lambert, O. **(2008)**. "Beaked whale mysteries revealed by seafloor fossils trawled off South Africa," South African Journal of Science **104**, 140–142.

Bittle, M., and Duncan, A. **(2013)**. "A review of current marine mammal detection and classification algorithms for use in automated passive acoustic monitoring," Proc. Acoust. Sci. Technol. Amenity. pp. 1–8.

Castiglioni, P. **(2010)**. "What is wrong in Katz's method? Comments on: 'a note on fractal dimensions of biomedical waveforms'," J. Computers in Biology and Medicine **40**, 950–952.

Cheeger, J. **(1970)**. "A lower bound for the smallest eigenvalue of the Laplacian," In: Gunning, R.C. (Ed.), *Problems in Analysis*. Princeton Univ. Press. pp. 195–199.

Cybenko, G.V. **(1989)**. "Approximation by superpositions of a sigmoidal function," Math. Control Signals Systems **2**, 303–314.

Davies, D.L., and Bouldin, D.W. **(1979)**. "A cluster separation measure," IEEE Transactions on Pattern Analysis and Machine Intelligence **PAMI-1**, 224–227.

Donoho, D.L. and Johnstone, I.M. **(1994)**. "Ideal spatial adaptation by wavelet shrinkage," J. Biometrika **81**, 425–455.

Frasier, K.E., Wiggins, S.M., Harris, D., Marques, T.A., Thomas, L., and Hildebrand, J.A. **(2016)**. "Delphinid echolocation click detection probability on near-seafloor sensors," J. Acoust. Soc. Am. **140**, 1918–1930.

Gómez, C., Mediavilla, A., Hornero, R., Abásolo, D., and Fernández, A. **(2009)**. "Use of the Higuchi's fractal dimension for the analysis of MEG recordings from Alzheimer's disease patients," J. Med Eng Phys. **31**, 306–313.

Han, N.C., Muniandy, S.V., and Dayou, J. **(2011)**. "Acoustic classification of Australian anurans based on hybrid spectral-entropy approach," J. Appl. Acoust. **72**, 639–645.

Henderson, G.T., Wu, P., Ifeachor, E.C., and Wimalaratna, H.S.K. (**1998**). Proceedings of the 3rd International Conference on Neural Networks and Expert Systems in Medicine and Healthcare, pp. 322–330.

Higuchi, T. (**1998**). "Approach to an irregular time series on the basis of the fractal theory," Physica D. **31**, 277–283.

Holmes, J.D., Carey, W.M., Lynch J.F., Newhall, A.E., and Kukulya A. (**2005**). "An autonomous underwater vehicle towed array for ocean acoustic measurements and inversions," In: *IEEE OCEANS'05 Europe Conference Proceedings*, **2**, 1058–1061.

Ioup, J.W. and Ioup, G.E. (**2004**). "Sperm whale identification using self organizing maps," J. Acoust. Soc. Am. **115**, 2556.

Kaiser, J.F. (**1990**). "On a simple algorithm to calculate the 'energy' of a signal," Proceedings of the IEEE International Conference of the Acoustical Speech, and Signal Processsing, Albuquerque, NM, pp. 381–384.

Kandia, V., and Stylianou, Y. (**2006**). "Detection of sperm whale clicks based on the Teager-Kaiser energy operator," J. Appl. Acoust. **67**, 1144-1163.

Katz, M. (**1988**). "Fractals and the analysis of waveforms," J. Comput. Biol. Med. **18**, 145–156.

Kohonen, T. (**1982**). "Self-organized formation of topologically correct feature maps," Biological Cybernetics **43**, 59–69.

MacLeod, C.D., Perrin, W.F., Pitman, R., Barlow, J., Ballance, L., D'Amico, A., Gerrodette, T., Joyce, G., Mullin, K.D., Palka, D.L., and Waring, G.T. (**2006**). "Known and inferred distributions of beaked whale species," J. Cetacean Res. Manage. **7**, 271–286.

MacQueen, J.B. (**1967**). "Some methods for classification and analysis of multivariate observations," Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability. University of California Press. pp. 281–297.

Madsen, P.T., Kerr, I., and Payne, R. (**2004**). "Echolocation clicks of two free-ranging, oceanic delphinids with different food preferences: false killer whales Pseudorca crassidens and Risso's dolphins Grampus griseus," J. Experimental Biology **207**, 1811–1823.

Madsen, P.T. and Wahlberg, M. (**2007**). "Recording and quantification of ultrasonic echolocation clicks," J. Deep-Sea Res I. **54**, 1421–1444.

Maragos, P., and Potamianos, A. (**1999**). "Fractal dimensions of speech sounds: Computation and application to automatic speech recognition," J. Acoust. Soc. Am. **105**, 1925–1932.

Moehl, B., Surlykke, A., and Miller, L.A. (**1990**). "High intensity narwhal click," In: Thomas, J., Kastelein, R. (Eds.), *Sensory Abilities of Cetaceans*. Plenum Press. pp. 295–304.

Møller, M.F. (**1993**). "A scaled conjugate gradient algorithm for fast supervised learning," J. Neural Networks **6**, 525–533.

Nielsen, M. (**2001**). "On the construction and frequency localization of finite orthogonal quadrature

filters," J. of Approximation Theory **108**, 36–52.

Parada, P.P., and Cardenal-López, A. **(2014)**. "Using Gaussian mixture models to detect and classify dolphin whistles and pulses," J. Acoust. Soc. Am. **135**, 3371–3380.

Parnum, I., McCauley, R.D., Cato, D.H., Thomas F.P., Duncan A.J., and Johnson, M. **(2011)**. "Detection of beaked whale clicks in underwater noise recordings," Australian Acoustical Society Conference 2011.

Rényi, A. **(1961)**. "On measures of information and entropy," Proceedings of the fourth Berkeley Symposium on Mathematics, Statistics and Probability. pp. 547–561.

Shannon, C.E. **(1948)**. "A mathematical theory of communication," Bell System Technical Journal **27**, 623–666.

Schorr, G.S., Falcone, E.A., Moretti, D.J., and Andrews, R.D. **(2014)**. "First long-term behavioral records from Cuvier's beaked whales (Ziphius cavirostris) reveal record-breaking dives," PLoS ONE **9**, e92633.

Sousa-Lima, R.S., Norris, T.F., Oswald, J.N., and Fernandes, D.P. **(2013)**. "A review and inventory of fixed autonomous recorders for passive acoustic monitoring of marine mammals," Aquatic Mammals **39**, 23–53.

Stimpert, A., Au, W., Parks, S., Hurst, T., and Wiley, D. **(2011)**. "Common humpback whale (Megaptera novaeangliae) sound types for passive acoustic monitoring," J. Acoust. Soc. Am. **129**, 476–82.

Svozil, D., Kvasnicka, V., and Pospichal, J. **(1997)**. "Introduction to multilayer feed-forward neural networks," Chemom. Intell. Lab. Syst. **39**, 43–62

Thomas, L., and Marques, T.A. **(2012)**. "Passive acoustic monitoring for estimating animal density," Acoustics Today, **8**, 35–44.

Vesanto, J. and Alhoniemi, E. **(2000)**. "Clustering of the self-organizing map," IEEE Trans. Neural Netw. **11**, 586–600.

Yack, T.M., Barlow, J., Roch, M.A., Klinck, H., Martin, S., Mellinger, D.K., and Gillespie, D. **(2010)**. "Comparison of beaked whale detection algorithms," J. Applied Acoustics **71**, 1043–1049.

Yack, T.M., Barlow, J., Calambokidis, J., Southall, B., and Coates, S. **(2013)**. "Passive acoustic monitoring using a towed hydrophone array results in identification of a previously unknown beaked whale habitat," J. Acoust. Soc. Am. **134**, 2589–95.

Zhao, Y., and Karypis, G. **(2001)**. "Criterion functions for document clustering: experiments and analysis," (Technical Report #01-40). Dept. of Computer Science, University of Minnesota.

Zimmer, W.M.X., and Tyack, P.L. **(2007)**. "Repetitive shallow dives pose decompression risk in deep-diving beaked whales," Marine Mammal Science **23**, 888–925.

Vita

The author was born on November 28, 1992 in Oakland, California, and moved to Covington, Louisiana at an early age. During high school, he was dually enrolled at Mandeville High School and The New Orleans Center for Creative Arts. After graduating, and attending The California Institute of the Arts for a semester, he decided to pursue a degree in science. Returning to Louisiana, he attended The University of New Orleans, and in 2015 received a bachelor's degree in physics. During his senior undergraduate year, he implemented computational skills as a student researcher for The Department of Physics, as well as The Department of Earth and Environmental Sciences. Following graduation, he worked at The Naval Oceanographic Office as a computer scientist. He returned to UNO in the spring of 2017 to pursue a master's degree in applied physics under the guidance of Dr. Juliette Ioup. While performing his thesis research, he also engaged in research for the university departments of environmental sciences, and biological sciences.