Automating Fish Sound Recognition in Spawning Aggregations: Application of Passive Acoustics in Fisheries

Reconocimiento Automatizado de los Sonidos de los Peces Durante Agregaciones de Desove: Aplicación de Acústica Pasiva en las Pesquerías

Détection des sons Produits par les Poissons Pendant les Agrégations de frai: Application de L'acoustique Passive aux Pêcheries

MICHELLE SCHÄRER-UMPIERRE*[1], RICHARD APPELDOORN[1], ALI IBRAHIM[2], NURGUN ERDOL[2], HANQI ZHUANG[2], LAURENT CHERUBIN[3], and BING OUYANG[3]
[1]*Department of Marine Sciences and Caribbean Coral Reef Institute, University of Puerto Rico, Mayagüez, Puerto Rico 00681-9000. *michelle.scharer@upr.edu.*
[2]*Department of Computer and Electrical Engineering and Computer Science, Florida Atlantic University, 777 Glades Road, Boca Raton, Florida 33431 USA.*
[3]*Harbor Branch Oceanographic Institute – Florida Atlantic University, 5600 US 1 North, Fort Pierce, Florida 34946 USA.*

KEY WORDS: Red hind, spawning aggregations, passive acoustics, Puerto Rico, signal recognition

EXTENDED ABSTRACT

Fisheries independent sampling of species that form spawning aggregations is increasingly important as these fish may show hyper-stability (Erisman et al. 2011, Heppell et al. 2013). Constant landings and catch rates of species that aggregate to spawn can continue to occur even though populations may be declining overall (Robinson et al. 2014). In some cases, regulations that prohibit their capture during spawning months may affect the reliability of fishery dependent metrics. When spawning ready individuals aggregate to reproduce at predictable times and locations these stocks can potentially be assessed with surveys to determine size structure and reproductive parameters (Nemeth 2005, Nemeth and Kadison 2013,). In order to effectively and efficiently survey the population during spawning aggregations the spatio-temporal dynamics of each species reproductive behavior must be well understood. The months of spawning, days of maximum abundances, and the occurrence of multiple species at spawning aggregations can vary by location (Mann et al. 2010, Wall et al. 2014) or over time therefore, any sampling design focusing on spawning aggregations requires this basic knowledge.

The use of passive acoustics of soniferous species has been proposed as a measure of reproductive activity in some cases because the sounds produced are associated with the formation of spawning aggregations (Amorim et al. 2015, Lobel et al. 2010, Lobel, 1992). Many groupers that occur in the US Caribbean have been recorded interacting with other individuals and producing species specific sounds during territorial displays, courtship and antagonistic events (Mann et al. 2010; Schärer et al. 2014, Schärer et al. 2012a, Schärer et al. 2012b). Courtship associated sounds (CAS) are detected during days prior to spawning and the low-frequency sound levels can be measured *in situ* to detect differences in reproductive activity and relative density (Rowell et al. 2012, Schärer et al. 2012b). In groupers these sounds are generated by the rapid contraction of cranial muscles against the gas filled swim bladder (Hazlett and Winn, 1962). The dominant calls of each species may have slight differences in peak frequency generally based on the size of the swim bladder of each fish. Smaller swim bladders can produce sounds of higher frequency, while larger ones tend to produce low frequency sounds.

Hydrophones (sensitivity = −186 dBV μPa−1; frequency range 2 to 37 kHz) are coupled with micro-computer, circuit board, and a battery power source in a sealed tube. Each hydrophone requires a specific calibration value based on factory specifications. These instruments (DSG from Loggerhead Instruments) are programmed to record sounds on a memory card in low frequencies over time in intervals that can range up to a year. Sampling rates are 10,000 Hz for digitizing the audio record. Each file is compressed to .DSG format with a date and time stamp in the file header. These files are then uncompressed into .wav files that can be listened to or analyzed with additional methods.

To date, the method to collect sounds at spawning aggregations involves vessel-based recorders with hydrophones lowered in the water or by deploying long-term recorders on the seafloor, which remain there for months. Once the recordings are collected there are two methods to quantify the sound production levels over time. One approach is to measure the overall sound level of a frequency band in which that species is known to produce sound. Each file with passive acoustic recordings is filtered to eliminate the sounds outside the target frequency band and after that the received sound pressure levels ((dB re 1 μPa). This procedure is run in MATLAB with DSG Lab that incorporates the dB relative to 1 μPa, followed by a root mean square calculation and a hydrophone calibration adjustment. While this method is rapid, the range of values recorded may include sound generated by sources other than the grouper that coincide in that frequency band.

The second method to quantify the sounds produced by fish is by visually inspecting the spectrograms generated for each .wav audio file. Various software applications are available to convert audio files into spectrograms such as *Ishmael*, *Adobe Auditon*, and *Audacity*. Each spectrogram is generated with a Hamming window with a frame size of 2,048 samples. A trained observer proceeds to verify each spectrogram and can determine the number of sound signals that are representative of each species. The visual inspection of each file is labor intensive and may incorporate observer error that can only be corrected by a second observer.

To overcome the limitations of the above approaches and to provide a method that is rapid, quantitative and species specific, an automated method previously used in marine acoustic research for mammals is now being applied to quantify fish sound signals. The grouper sound recognition system was first developed to classify each sound file based on the presence of CAS. This system is composed of two main procedures; training and recognition. The system's block diagram is presented below:

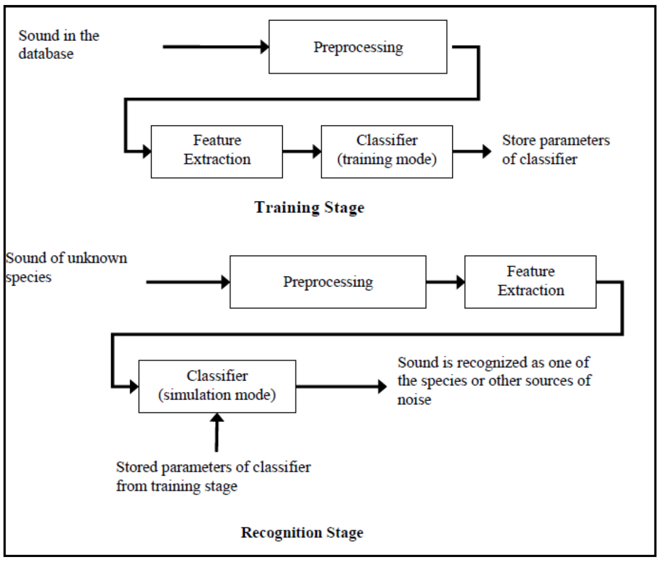

Figure 1. System Block Diagram

The Function of each block is explained below:

i) **Pre-processing:** In order to flatten sound spectrum, a pre-emphasis filter has been used before spectral analysis.
ii) **Feature Extraction:** Mel Frequency Cepstral Coefficients (MFCC) has been used in this stage.
iii) **Classifier:** In the classification stage, popular classifiers such as Support Vector Machine (SVM) and K-Nearest Neighbor (KNN) are to be tested to detect and classify Grouper species sounds.
iv) The processing in both training and identification stages is similar except for the classifier, which works in training mode in the training stage and in simulation mode in the recognition stage.

Mel Frequency Cepstral Coefficients (MFCC)

Feature extraction of acoustic signals is an important task to produce a better recognition performance. The efficiency of this phase is important for the next phase since it affects its behavior. Mel Frequency Cepstral Coefficients (MFCC) is one of the most popular features for classification of acoustic signals, including those generated by marine mammals (Roch et al. 2007). The success of MFCC is due to the use of Mel spaced filter banked processing of the Fourier transform which provides robustness to the system. Figure 2 describes the block diagram of the MFCC algorithm.

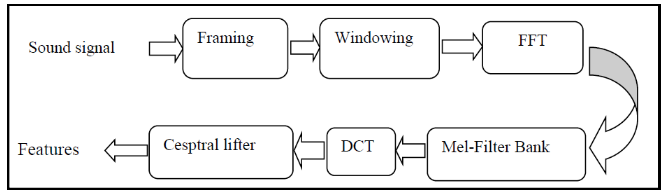

Figure 2. MFCC extraction algorithm.

The following steps were used to evaluate MFCCs:

i) The signal is divided into a small frame with length within the range of 200 to 500 msec.
ii) Hamming Window is used to discard the effect of discontinuities at edges of frame.
iii) Fast Fourier Transform (FFT) is applied to convert each frame of N Samples from time domain into frequency domain.
iv) The result is passed through a set of mel filter banks.
v) Discrete Cosine Transform (DCT) is used to convert the log mel spectrum into time domain.
vi) Finally, Sinusoidal lifter has been used for reweighting process of cepstral coefficients.

The result of conversion is called Mel Frequency Cestrum Coefficients.

K-Nearest Neighbor

The idea behind the K-Nearest Neighbor (KNN) algorithm is straightforward and easy to understand. Individuals are classified based on the distance from their nearest neighbors in the feature space. The two tasks that lead to the decision are: a) to find *k* closest points in set Z to a query point or set of points X, and b) to assign a class to an unknown point based on the dominant class among those nearest neighbors (Cover and Hart 1967). Various metrics can be used to determine the distance between data points in Z and query points in X but the Minkowski metric is defined as:

$$d = \sqrt[p]{\sum_{j=1}^{n} \left| z_j - x_j \right|^p} \qquad (1)$$

where Euclidean distance is given for a special case of $p = 2$ and Minkowski metric calculates the city block metric when $p = 1$. Although other algorithms exist to make the decision making more robust, the method suffers from a judgment of fairness and is open to many arbitrary parametric decisions (Coomans and Massart 1982). KNN classifiers fail for objects that are on or near class boundaries unless the margin between the class boundaries is sufficiently large, and also misclassified objects bias subsequent classifications.

A preliminary test of this algorithm was employed with data collected at a known red hind (*Epinephelus guttatus*) aggregation site off western Puerto Rico. The site at Abrir la Sierra has been studied since 2007 with passive acoustic techniques (Mann et al. 2010, Rowell et al. 2012).

A dataset comprising the spawning season of 2013 to 2014 was analyzed with all three methods described above. The result of each method was very similar, picking up days with high sound production associated with two presumed spawning events (January and February). Figure 3 summarizes the results with each method, displayed as separate starting points and axes to facilitate comparison.

The development of this new methodology will increase the applicability and efficiency with which passive acoustic data can be used to assess the spatio-temporal variability of FSA or soniferous species.

LITERATURE CITED

Amorim, M.C.P., R.O. Vasconcelos, and P.J. Fonseca. 2015. *Sound Communication in Fishes* (Volume 4). http://doi.org/10.1007/978-3-7091-1846-7.

Coomans, D. and D.L. Massart, 1982, Alternative k-nearest neighbour rules in supervised pattern recognition: Part 1. k-Nearest neighbor classification by using alternative voting rules," *Analytica Chimica Acta*, Volume 136, pp: 15–27. doi:10.1016/S0003-2670(01)95359-0

Cover, T. M. and P.E. Hart, 1967, "Nearest neighborhood pattern classification," IEEE Transactions on Information Theory, vol. 13. no. 1.

Erisman, B.E., L. Allen, J.T. Claisse, D.J.P. Ii, E.F. Miller, and J.H. Murray. 2011. The illusion of plenty: hyperstability masks collapses in two recreational fisheries that target fish spawning aggregations, *1716*, 1705–1716. http://doi.org/10.1139/F2011-090.

Hazlett, B. and H.E. Winn. 1962. Sound producing mechanism of the Nassau grouper, *Epinephalus striatus*. *Copeia* **1962**(2):447-449.

Heppell, S.A., B.X. Semmens, C.P. Semmens, P. Bush, C. McCoy, and B. Johnson. 2014. Behavior, hyperstability , and population declines of an aggregating marine fish. *Proceedings of the Gulf and Caribbean Fisheries Institute* **66**:379-380).

Lobel, P.S. 1992. Sounds produced by spawning fishes. *Environmental Biology of Fishes*, 351–358. Retrieved from: http://link.springer.com/article/10.1007/BF00010947.

Lobel, P.S., I.M. Kaatz, and A.N. Rice. 2010. Acoustical behavior of coral reef fishes. Pages 307-348 in: K. Cole (ed.) *Reproduction and Sexuality in Marine Fishes.* University of California Press, Oakland, California USA.

Mann, D.A., J.V. Locascio, M.T. Schärer, M.I. Nemeth, and R.S. Appeldoorn. 2010. Sound production by red hind *Epinephelus guttatus* in spatially segregated spawning aggregations. *Aquatic Biology* **10** (2):149-154. http://doi.org/10.3354/ab00272.

Nemeth, R.S. 2005. Population characteristics of a recovering US Virgin Islands red hind spawning aggregation following protection. *Marine Ecology Progress Series* **286**:81-97.

Nemeth, R.S. and E. Kadison. 2013. Temporal patterns and behavioral characteristics of aggregation formation and spawning in the Bermuda chub (*Kyphosus sectatrix*). *Coral Reefs*. http://doi.org/10.1007/s00338-013-1050-0.

Robinson, J., N.A.J. Graham, J.E. Cinner, G.R. Almany, and P. Waldie. 2014. Fish and fisher behaviour influence the vulnerability of groupers (Epinephelidae) to fishing at a multispecies spawning aggregation site. *Coral Reefs*. http://doi.org/10.1007/s00338-014-1243-1.

Roch, M.A., M.S. Soldevilla, J.C. Burtenshaw, E.E. Henderson, and J.A. Hildebrand. 2007. Gaussian mixture model classification of odontocetes in the Southern California Bight and the Gulf of California. *The Journal of the Acoustical Society of America* **121**:1737.

Rowell, T.J., M.T. Schärer, R.S. Appeldoorn, M.I. Nemeth, D.A. Mann and J.A. Rivera. 2012. Sound production as an indicator of red hind density at a spawning aggregation. *Marine Ecology Progress Series* **462**:241-250. http://doi.org/10.3354/meps09839,

Schärer, M.T., M.I. Nemeth, D.A. Mann, J.V. Locascio, R.S. Appeldoorn and T.J. Rowell. 2012. Sound production and reproductive behavior of yellowfin grouper, *Mycteroperca venenosa* (Serranidae) at a Spawning Aggregation. *Copeia* **1**:135-144.

Schärer, M.T., M.I. Nemeth, T.J. Rowell, and R.S. Appeldoorn 2014. Sounds associated with the reproductive behavior of the black grouper (*Mycteroperca bonaci*). *Marine Biology* **161**:141-147. http://doi.org/10.1007/s00227-013-2324-3.

Schärer, M.T., T.J. Rowell, M.I. Nemeth, and R.S. Appeldoorn. 2012. Sound production associated with reproductive behavior of Nassau grouper *Epinephelus striatus* at spawning aggregations. *Endangered Species Research* **19**(1):29-38. http://doi.org/10.3354/esr00457.

Wall, C.C., P. Simard, M. Lindemuth, C. Lembke, D.F. Naar, C. Hu, and D.A. Mann. 2014. Temporal and spatial mapping of red grouper *Epinephelus morio* sound production. *Journal of Fish Biology* **85** (5):1470-88. http://doi.org/10.1111/jfb.12500.

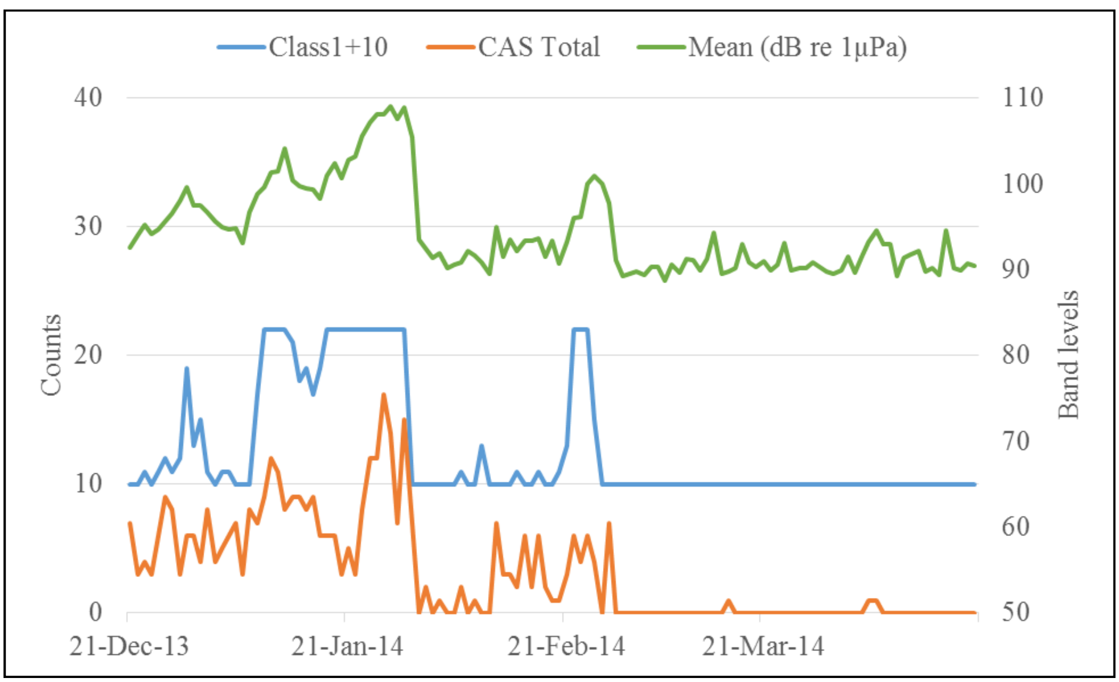

Figure 3. Results of red hind sound levels per day at Abrir la Sierra with three different methods; band level estimation (green), grouper recognition system files with red hind CAS present (blue), and total counts of red hind CAS based on visual inspection (red).