

Biodiversity assessment using passive acoustic recordings from off-reef location – unsupervised learning to classify fish vocalization

Vasudev P. Mahale,^{1, a} Kranthikumar Chanda,¹ Bishwajit Chakraborty,^{1, b} Tejas Salkar,¹ and G.B. Sreekanth²

¹CSIR - National Institute of Oceanography, Dona Paula, Goa: 403 004, India

²ICAR - Central Coastal Agricultural Research Institute, Old-Goa, 403 402, Goa, India

^aAlso at AcSIR - Academy of Scientific and Innovative Research, School of Oceanography, CSIR-NIO Campus, Dona-Paula, Goa403004, India

^bElectronic mail: bishwajit@nio.org

We present the quantitative characterization of Grande Island's off-reef acoustic environment within the Zuari estuary during the pre-monsoon period. Passive acoustic recordings reveal prominent fish choruses. Detailed characteristics of the call employing oscillograms and individual fish call parameters of the segmented data include vocal groups such as Sciaenidae, *Terapon theraps*, and planktivorous as well as invertebrate sounds, e.g., snapping shrimp. We calculated biodiversity parameters i) Acoustic Evenness Index (AEI), ii) Acoustic Complexity Index (ACI), and mean Sound Pressure Level (SPLrms) for three frequency bands such as full band (20-22,050 Hz), the low-frequency fish band (100-2000 Hz), and the high-frequency shrimp band (2000-20,000 Hz). Here, ACI and AEI metrics characterize the location's soundscape data effectively indicating increased biodiversity of fish species for both, the low-frequency and high-frequency bands. Whereas variations for SPLrms are prominent for three frequency bands. Moreover, we employ unsupervised classification through a hybrid technique comprising principal component analysis (PCA) and K-means clustering for data features of four fish sound types. Employed PCA for dimensionality reduction and related K-means clustering successfully provides 96.20%, 76.81%, 100.00%, and 86.36% classification during the dominant fish chorus. Overall, classification performance (89.84%) is helpful in the real-time monitoring of the fish stocks in the ecosystem.

I. Introduction

Application of passive acoustic technique has become important to investigate variations in ocean ambient sound characteristics (Lammers *et al.*, 2008; Luczkovicz *et al.*, (2008). Marine soundscape is the aggregation of sound sources known as geophony, biophony, and anthrophony (La Manna *et al.*, 2021). The soundscape is characterized by spectral, temporal, and spatial features, which depend on location and time (Pijanowski, *et al.*, 2011a). Marine environments' primary sources of biological sounds are marine mammals, soniferous fish, and invertebrates. Soniferous fishes represent a group of vocal vertebrates that rely on acoustic signals during various social interactions and are highly diverse in species' life histories (Fine and Parmentier 2015). Looby *et al.* (2022) reported that 989 fish species from 133 families and 33 orders could produce active (i.e., intentional) sound. Many fish species rely on vocal signaling during their activities and produce sounds using sonic muscles that vibrate the swim bladder or rub bony elements (stridulation) (Kasumyan, 2008). Bioacoustics research investigate the behavior of individuals or single species, and therefore, individual fish sound characterization (Fine, 2012; McCauley and Cato, 2000) involves segmentation of recorded 60 seconds oscillogram data to acquire individual fish call parameters (temporal and spectral)(Chanda *et al.*, 2020a).The temporal parameters, such as call duration, the number of pulses, and inter call duration within the fish call, along with the spectral parameter i.e., frequency peaks of the power spectral density are important (Borie *et al.*, 2014;Tricas and Boyle, 2014; Fish and Mowbray, 1970). Such parameters of the individual calls are beneficial for fish family/species detection and necessary for soundscape analyses.

In coastal water, the ambient sound field generally consists of various sound sources such as biophony, geophony, and anthrophony due to the causes and consequences of biological, geophysical, and human-produced sounds (Pijanowski *et al.*, 2011b). Most fish calls and snapping shrimp sounds belong within the 100 to 20,000 Hz frequency ranges (Fish *et al.*, 1952; Knowlton *et al.*, 1963). Based on the published frequency ranges of fish calls and snapping shrimp sounds, we observed that these two frequencies do not tend to overlap. The low-frequency band (100 Hz to 2000 Hz) corresponds to the range in which most fish species vocalize (Bertucci *et al.*, 2016). The high-frequency shrimp-band (2000 to 20,000 Hz) encompasses the range typically dominated by snapping shrimp (Au and Banks, 1998). Within these two bands, the ambient noise is a composite of the two overlapping components, i.e., wind (50 to 20,000 Hz) and oceanic traffic (10 to 10,000 Hz). Both the components, such as wind and oceanic traffic, may interfere with both bands [(NRC, 2003) based on the Wenz (1962)]. Below 100 Hz, most contaminating flow noise has been reported (Bassett *et al.*, 2014). We categorize three

frequency bands as i) full-band (50 to 20,050 Hz), ii) low-frequency fish band (100 to 2000 Hz), and iii) high-frequency shrimp band (2000 to 20,000 Hz) (Chanda *et al.*, 2020b).

Biodiversity is assessed by calculating a series of indices (Sueur *et al.*, 2014; Farina *et al.*, 2017, Farina *et al.*, 2021; Pijanowski *et al.*, 2021). Diversity indices provide essential information about the rarity or commonness of species in the community. Quantifying biodiversity continues to be challenging, and there is no single index that can provide a complete assessment. A new discipline known as "Eco-acoustics" is currently emerging, which relates animal sound to assess ecology and monitor biodiversity to understand the location's ecology. Here, sound is considered as a tool to infer ecological information. Acoustic indices represent a convenient means (rapid, computationally intensive, but relatively straight forward code) to characterize soundscapes and assess biodiversity (Villanueva-Rivera *et al.*, 2011; Sueur, 2018). Several selected acoustic indices are employed explicitly to determine biodiversity (Bradfer-Lawrence *et al.*, 2019). Therefore, long-term data processing using automated systems, and species identification needs a prudent approach for monitoring purposes, and specific recommendation regarding ground truthing i.e., species identification (Mooney *et al.*, 2020), which is ensued in this research. Acoustic Complexity Index (ACI) (Pieretti *et al.*, 2011) calculates the spectrogram complexities, and detect the presence of animal sound but remain unresponsive to various kinds of monotonous background noise (Pieretti *et al.*, 2017). More precisely, ACI looks for changes in the complexity over time intervals and parses out biologic sounds from anthropogenic due to the cadence associated with the biological sounds. ACI is used in habitats having mud, cliff, and gravel seafloor (McWilliam and Hawkins, 2013), coral reef and off reefs (Bertucci *et al.*, 2016), coasts and various shallow sea areas (Harris *et al.*, 2016). These investigations indicate positive correlation among the ACI metric with respect to the biodiversity.

Besides ACI, there are few more acoustic indices that are used to evaluate biodiversity, and so far, these indices have been extensively applied for terrestrial applications (Villanueva-Rivera *et al.*, 2011). Interestingly, for marine environments, very limited applications are made except for ACI. Acoustic Evenness Index (AEI) is the Gini coefficient, providing a measure of evenness (Eldridge *et al.*, 2018). We chose AEI for present investigation. Acoustically rich habitats may produce lower values because there is little variation in intensity among frequency bands in saturated soundscape (Bradfer-Lawrence *et al.*, 2019). SPLrms is the root-mean-square (RMS) of the raw audio signal. SPLrms has been demonstrated to track ecologically- relevant temporal and spatial dynamics in the forest canopy and shown to be strongly positively correlated with the percentage of living coral cover in tropical reefs (Bertucci *et al.*, 2016).

Here, we calculate multiple biodiversity parameters using the recorded passive acoustic data and apply machine learning techniques (Bianco *et al.*, 2019) to uncover patterns in recorded fish vocalization (Malfante *et al.*, 2018), which in turn help in fish sound identification. Limited application is carried out using machine learning for passive acoustic data analyses and related classification. Roca and Opzeeland (2020) had successfully used twenty-three different acoustic metrics to develop a supervised machine learning approach to discriminate between two different acoustic environments. Similarly, Williams *et al.*, (2022) had also used supervised machine learning techniques for eco-acoustics parameters to carry out automated analyses. Healthy and degraded reef sites were identified through live coral cover surveys having 90–95% and 0–20% cover respectively. As our aim also include involvement of an unsupervised learning (Lee and Staneva, 2000) technique for classification of the recorded fish sound data, we use peak frequency and corresponding peak of the PSDs and computed biodiversity parameters as input for unsupervised classifier (Ozanich *et al.*, 2018). We apply classifier to the known soundscape parameters (Lin *et al.*, 2018), which benefits to check the performance of the classifier.

In order to develop a method to identify fish sound, clustering is a type of unsupervised learning that we use when we do not know beforehand what groups each observation in the data belongs to. One of the major clustering approaches is based on the sum-of-squares criterion and on the algorithm that is today well-known under the name ‘K-means’ (Bock, 2007). K-means is a commonly used partitioning based clustering technique that tries to find a user specified number of clusters (k), often do not work well for high dimensional data (Ding and He, 2004). Principal component analysis (PCA) is a technique which is used to increase the interpretability of such datasets by reducing their dimensionality (Jolliffe and Cadima, 2016). The earliest literature on PCA dates from Hotelling (1933), and its use has rapidly increased and several variants have been developed (Jolliffe, 2002). PCA is found to be effective in reducing data dimension. Therefore, applying K-means clustering technique in the reduced dimension obtained from PCA is being implemented in recognizing the patterns in the datasets (Amiri-Simkooei *et al.*, 2011). Here, we utilize a combination of PCA with K-means clustering techniques, also known as a hybrid clustering technique. The present work include: i) application of the acoustic biodiversity metrics such as ACI and AEI, and SPLrms for three frequency bands, and ii) development of an unsupervised learning - driven methodology using these metrics to classify fish sounds which have been analyzed using the parameters that have been described above.

II. MATERIALS AND METHOD

A. Study location and recorded data

Goa's major coastal island ecosystems include Grande Island. This island is situated near the mouth of the Zuari estuary (Ansari *et al.*, 1995) off South Goa in the Arabian Sea, having diverse flora and fauna (Fig. 1). A report from Grande Island indicates a comparatively higher live coral cover in the mid-shelf zone (Manikandan *et al.*, 2016; Hussain *et al.*, 2020). During the pre-monsoon session, we carried out data acquisition in the present study location (15° 20.160' N; 073° 45.277' E) during 10-14 March 2016. The location is away from the reef system, and the distance from the nearest landmass of the Grande Island is 1870m, where water depth is around 20m. For present application, computed transmission loss (TL) at this water depth and range due to sound propagation are negligible (0.022 dB to 0.563 dB) for 500 to 5000 frequencies respectively (Ainslie and McColm, 1998). Therefore, to protect our recorder during windy weather, our measurements are carried out at 20 m water depth. We used, measured values of 20°C water temperature, 35 ppt salinity, and eight pH values for our computation.

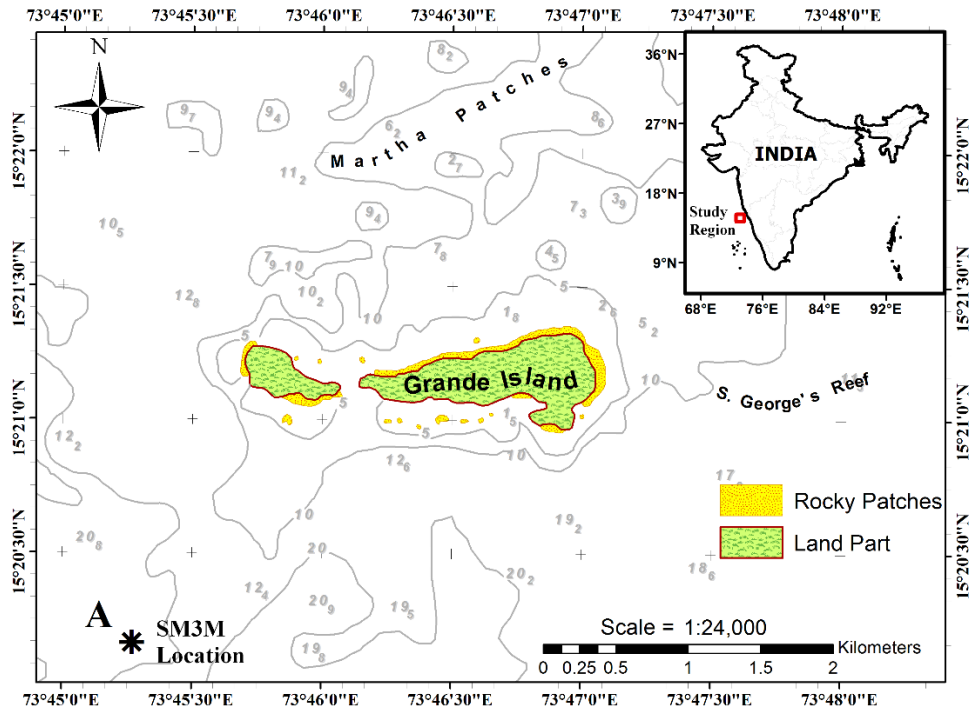


FIG. 1: Map showing the study location around Grande Island from Central West Coast of India. SM3M system¹ is deployed at 20m water depth which comprised of hydrophone and associated recording electronics.

B. Acoustic instrumentation

We used the Song Meter SM3M systems¹ to acquire passive acoustic data (Fig. 2a). The Song Meter SM3M system possesses standard broadband hydrophones (2- 192000Hz bandwidth). The sampling frequency was 96000Hz for field data acquisition. However, to maintain comparability among the acquired data, we resampled the data obtained at the present location to 44100Hz. We used the “resample” function of the MATLAB² signal processing toolbox by incorporating the FIR anti-aliasing filter. For data acquisition, we have programmed our recorder to acquire passive acoustic data for 60 seconds at every 15-minute intervals. The dataset under investigation was recorded a total of 92:30 h, and considering the battery capacity, we optimized data acquisition length of 60 seconds at an interval of 15-minute for our data acquisition. We employed moorings that have positive buoyant passive acoustic recording system i.e., Song Meter submersible (SM3M) tied to a 40 kg deadweight, which is lying on the seafloor. We used ESSO–National Institute of Ocean Technology (NIOT) calibration facility³ for hydrophone calibration.

C. Acoustic analysis

1. Concatenated Power Spectral Density

We present soundscape plots using concatenated PSDs (i.e., linking PSDs for 60 seconds passive acoustic data) at 15-min intervals for the study location on the time-frequency axis. The soundscape plot indicates the duration of data acquisition and intensity depicted by the colour bar. We computed PSD using the “pwelch” MATLAB function² with 50% overlapping Hanning windows with a length of 2048 samples (Fig. 3a). We also used online software Audacity (Audacity Development Team)⁴ to visualize and segregate the 60 seconds oscillograms to undergo comparisons with the archived *.wav files^{5,6}. A preliminary auditory comparison between data and archived sound file parameters allowed us to know about the fish species/families (Table I). The matching of the frequency peak (PSDs) and the auditory comparison between the datasets with a wav file is an initial comparison.

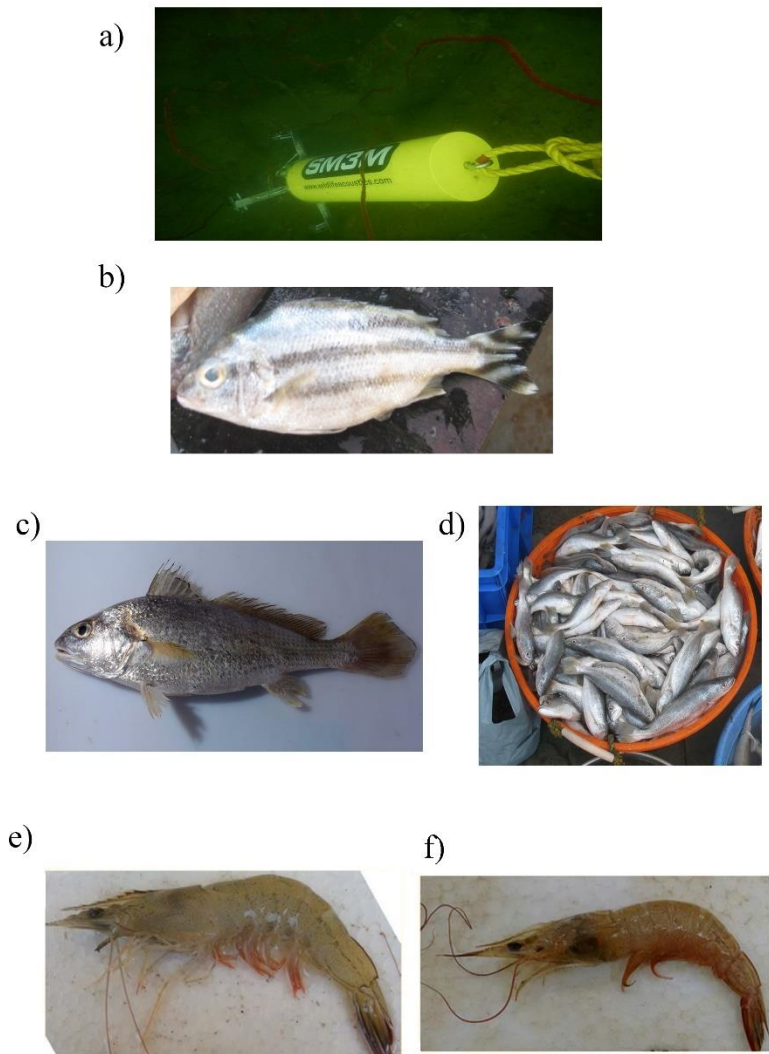


FIG. 2: (Color online) Photographs of (a) Song Meter system is getting deployed. (b) *Terapon theraps*. (c) Sciaenidae. (d) Live Sciaenidae caught in Zuari after the gillnet experiment. (e) Indian white shrimp (*F. indicus*) and (f) Flower tail prawn (*M. dobsonii*) recovered near the Grande Island¹⁵.

TABLE I. Soundscape information based on 60sec oscillogram length acquired from Grande Island, Arabian Sea from 10-16 March 2016.

Fish type	Timing (h)	Mean peak Frequency (Hz)	Mean Peak level (dBre 1 Pa ² /Hz)	Total no. data
Type A ^a	00:30 to 02:45	1137.20 ± 95.68	81.38 ± 2.46	22
Sciaenidae ^b	14:15 to 19:30	857.41±22.32	93.48 ± 4.69	84
Tiger Perch ^c	19:45 to 22:30	1505.87 ± 29.53	89.99 ± 13.86	69
Planktivorous	22:45 to 23:30	561.66 ± 29.73	96.18 ± 7.05	28
Boat	14:00 to 13:45	287.11 ± 65.78	83.41 ± 6.76	29
Snapping Shrimps	14:00 to 13:45	3629.65±1011.26 (max 5490.96 – min.2045.65)	75.95±6.79	370

^aUnknown fish categorized as Type A fish sound details

^bScientific name of drums and croakers is Sciaenidae

^cSpecies name *Terapon theraps* of Tiger Perch family having scientific name Terapontidae.

2. Computing Biodiversity

To assess biodiversity, we use two acoustic indices (i) Acoustic Evenness Index (AEI), and (ii) Acoustic Complexity Index (ACI). We employed packages to compute eco-acoustics indices in the R computing environment⁷. To calculate AEI and ACI parameters, we have utilized the ‘soundecology’ package⁸.

AEI is the Gini coefficient, providing a measure of evenness. We employ the Gini coefficient, a statistical measure that helps define inequality, which is a statistical measure that help define inequality. Moreover, species evenness is a preferred index used by ecologists and is computed utilizing Gini function⁹. It investigates occupancy of the function in each frequency band, and the index varies within the 0 and 1 range (Villanueva-Rivera *et al.*, 2011). AEI is extracted by dividing the spectrogram into n number of frequency bins and computing Gini index over the proportion of signals in each bin above a threshold (-50dBFS). AEI can be used for comparison between the sites.

ACI⁸ can quantify complex biotic sound patterns by calculating the variability of the acoustic energy in audio recordings (Pieretti and Farina,2013). The intensities registered in acoustic data were derived from its spectrogram by dividing the values of the spectrogram into bins of frequency and time. For ACI, spectrogram for each 60 seconds record was computed with 1024 point Fast Fourier Transform algorithm without overlap. This corresponds to 43.1 Hz frequency resolution and provides time resolution as high as 0.023 second.

We calculated the root-mean-square sound pressure level (SPLrms) for 60 seconds recorded data files. The SPLrms recorded at every 15-minute interval are used for comparison with the two computed acoustic metrics. The expression for SPLrms in dB re 1μPa is given by

$$SPL_{rms} = 20 \log_{10} \left(\sqrt{\frac{1}{T} \int_{t=0}^{t=T} P(t)^2 dt} \right) \quad (1)$$

Here P (t) is a root-mean-square (RMS) pressure level, and T is the call duration.

As we know that the ambient sound fields consist of biophony, geophony and anthrophony, and they are dominating within the different frequency ranges. In this investigation, we chose three biodiversity parameters: ACI, AEI, and SPLrms, and calculated them for three frequency bands: full-band, low-frequency fish band, and high-frequency shrimp bands. Therefore, we partitioned acoustic spectrum into frequency bands i.e., low-frequency fish band and high-frequency shrimp band and full-band. We employ band-pass filtering through use of the four-pole Butterworth filters for three frequency bands (Kaplan *et al.*, 2015).

D. Unsupervised learning

To develop unsupervised learning technique for soundscape characterization, we chose twelve parameters from sound files. Out of twelve, nine parameters related to vocalizations are acoustic metrics (AEI, ACI, and SPLrms) for three bands. Besides, three more parameters are employed. They are the peak frequency peaks and dominant peak PSDs, as well as the time of recorded vocalization.

PCA is a dimensionality reduction method used to reduce the dimensionality of large data sets. Mathematically, PCA provides an orthogonal linear transformation of the dataset into another coordinate system. We computed Spearman’s rank correlation coefficient, which is denoted by ‘rho’(ρ). Spearman’s correlation is a nonparametric (distribution-free) measure of rank correlation (Schober *et al.*, 2018). It assesses how well the relationship of the variables can be described using a monotonic function. After that, to form a feature vector, eigenvalues and eigenvectors of the covariance matrix are computed. The largest variance in the dataset is along the first PC1, and it successively diminishes for PC2, PC3, and so on. These PCs are obtained from the eigenvectors of the covariance matrix, and they define the new feature space (Jolliffe and Cadima, 2016). We draw scree plot i.e., line plot of the cumulative eigenvalues of factors or principal components, which helps us to decide number of features in data to be reduced, while keeping most of the variance.

Next, we utilize a combination of PCA with K-means clustering technique. Applying K-means clustering on the reduced dimensions of PCAs was found effective in recognizing the patterns in the data set (Hoppe and Roan, 2011). However, one of the significant challenges in K-means algorithm

implementation is to provide an initial number of possible clusters to the K-means algorithm, which has been overcome in the present study by implementing the Elbow method (Liu and Deng, 2021). Determining the appropriate number of clusters that should be acknowledged in a data set is an analytical and exploratory exercise. This is because, increasing the number of clusters will always reduce the distance between individual feature and associated cluster centers.

K-means clustering algorithm segments the observations into k-sets to minimize the within-cluster sum of squares (WSval). WSval is the sum of distances of each point in the cluster to the centre of cluster across k clusters (Kval). Five major steps for clustering through the combined use of PCA and K-means algorithm are: i) correlation matrix, ii) computing eigenvalues and vectors, iii) dimensionality reduction, iv) use of Z-score normalization to dimensionally reduced data and v) K-means algorithm-based clustering and use of elbow method. Z-score normalization¹⁰ is implemented when data variables have different units of measurement. Moreover, if each variable to receive equal weight in the analysis, then the variables should be standardized before conducting K-means. The application of unsupervised leanings for soundscape studies was carried out using five steps mentioned here (Lin *et al.*, 2017).

A confusion matrix¹¹ is a table that is used to define the performance of a classification algorithm. It is also known as an error matrix formed by checking the actual values and predicted values of observations in the datasets. In the confusion matrix, the column represents predicted label (class) i.e., output of the K-means algorithm, while row represent the actual label. Here, we used confusion matrix to determine extent of the classification success utilizing our labelled and classification data.

III. RESULTS

A. Fish assemblages and passive acoustic studies of the area

The Grande Island archipelago off the coast of Goa, India, is known for its fish diversity, among the few island ecosystems in India. Under the program "Assessment of estuarine fish communities," the Central Coastal Agricultural Research Institute (CCARI), a fisheries wing of the Institute of Agriculture Research (ICAR), uses continuous motorized gillnet experiments in the estuaries to prepare a checklist of fisheries resources from the estuaries around Goa (Sreekanth *et al.*, 2018). Here, number of fish species were recorded at the reefs, and they contribute to the abundance of commercial food fish species, including ornamental fish species. The greater species richness and abundance and several exclusive fish species recorded suggest that the reef sites constitute a unique habitat (Sreekanth *et al.*, 2017).

Investigation of sound analysis of fish families close to the reef sites along with the diver-assisted underwater visual census reported a total of 21 soniferous fish species (Chary *et al.*, 2020). Besides, off reef fishes such as Terapontidae family (Tiger Perch) (Sreekanth *et al.*, 2018), Sciaenidae (croakers and drums) (Velip and Rivonkar, 2018) (Fig. 2b-d), and planktivorous (Padate *et al.*, 2010), are also reported. Moreover, vocalizations of the Tiger Perch (Chanda *et al.*, 2020a), Sciaenidae and planktivorous (Mahanty *et al.*, 2018) are also recorded in the area. We also recorded snapping shrimp sounds, which are commercially important catch and belong to the Penaeid family (during the pre-monsoon, 15.47% out of the total annual fish catch) (Sreekanth *et al.*, 2019). (Fig. 2e).

B. Soundscape component

The concatenated PSD plots of the recorded oscillogram of 60 seconds) at every 15-minute are presented (Fig. 3a) in time and frequency axes of passive acoustic data recorded off-reef location of Grande Island (Fig.1) at 20m water depth. PSDs were initially scrutinized to identify biological sounds from the site. Our auditory comparison of 371 sound files and analyses of the concatenated PSDs reveal prominent fish choruses within 14:00 to 02:45 h during 10-14 March 2016 (Fig. 3a, c). We observe a pattern in the soundscape data, i.e., dominant ambient sounds due to fish vocalizations within the period (14:00 to 02:45 h). Interestingly, similar sounds due to the fish vocalizations were absent from (03:00 to 13:45 h) (Fig. 3a and c). The peak PSDs, and corresponding peak frequency of the concatenated PSDs are presented (Fig. 3b, c). The mean values of the dominant peak PSDs and corresponding frequency peaks of each 60 seconds recorded data for entire data acquisition days are $(86.20 \pm 11.15 \text{ dB re } 1 \text{ Pa}^2/\text{Hz})$ and $(1174 \pm 429.14 \text{ Hz})$ respectively. A significant difference in peak PSDs between fish chorus timing (14:00 - 02:45 h) and remaining timings (03:00 -13:45h) are found to be $(93.49 \pm 09.05 \text{ dB re } 1 \text{ Pa}^2/\text{Hz})$ and $(78.12 \pm 06.90 \text{ dB re } 1 \text{ Pa}^2/\text{Hz})$ respectively.

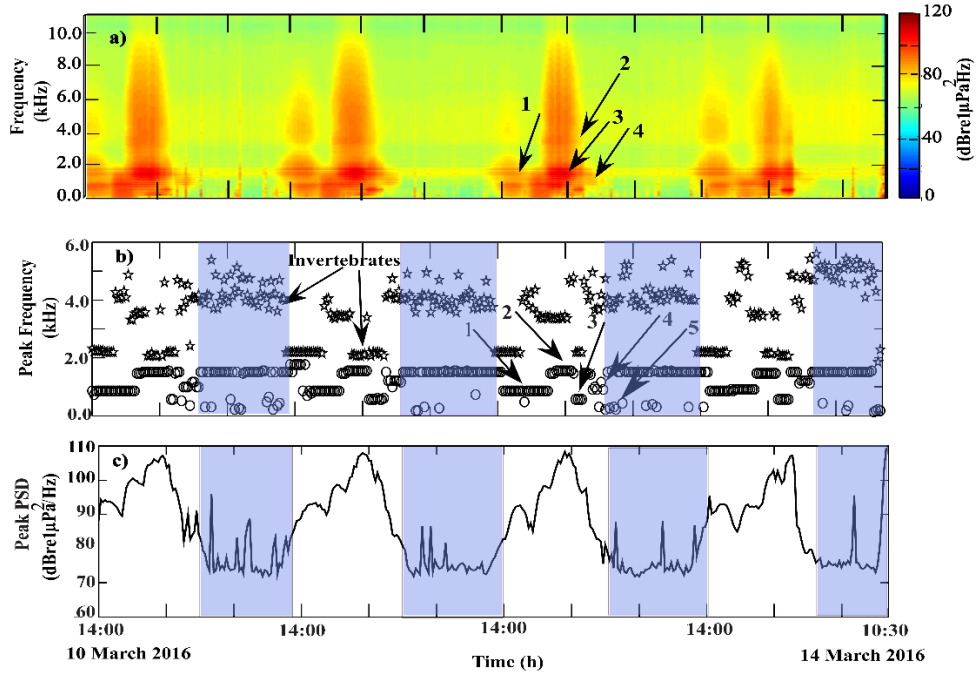


FIG. 3: (a) (Color online) Concatenated PSD (dB re $1 \mu\text{Pa}^2/\text{Hz}$) plot for 60 seconds passive acoustic time series records acquired at 15-minute interval with respect to the Time (h). (b) Frequency peaks of the PSD (kHz) plots of 60 sec oscillograms. (c) peak PSD (dB re $1 \mu\text{Pa}^2/\text{Hz}$) of the full band data, and high frequency shrimp band data (data points are shown as invertebrates). Arrows with number indicate 1. Sciaenidae, 2. *Terapon theraps*, 3. planktivorous, 4. Type A fish, and 5 boat sounds (only in b).

Presence of four types of fish sounds is observed within the chorus period, indicated as 1, 2, 3, and 4 (Fig. 3a-b). Accordingly, here we identified Sciaenidae (common name: croakers and drums) (14:15-19:30 h), Tiger Perch (19:45-22:30 h) and planktivorous (22:45- 23:30 h). Timing wise the trend of calls is always not appropriate and mixing of calls are observed occasionally especially for planktivorous. Also, we identified an unknown ‘buzz’ sound (00:30-02:45 h) and estimated its peak frequency using PSDs and have named them as a ‘Type A’. Luczkovich and Sprague (2022) had reported similar distant fish chorus of making the unknown “buzz”, although individual fish calls could not be discerned. They reported the pulse repetition rate and peak frequencies around 102.09 Hz and 1128 Hz respectively for this buzz sound.

Sciaenid (Fig. 2c, d) (drums and croakers) sounds [indicated as ‘1’ in (Fig. 3a, b)] match well with the reported (wav format) and archived sound data (short head drum) through auditory comparison⁵ and by matching the frequency peak of the PSDs of oscillograms (Ramcharitar *et al.*, 2006).

Comparison between the recorded data with the online Tiger Perch (Fig. 2b) sound file^{6,12} was also found to be matching. Recording of ‘trumpet ‘like sound mostly during the period (19:45 – 22:30 h) are indicated as ‘2’ (Fig. 3a, b) (McCauley and Cato, 2000). The daily spectral pattern shows the *Terapon theraps* chorus during dusk followed by the planktivorous during midnight, which is indicated as ‘3’ (Fig. 3a, b) (McCauley, 2012; Parsons *et al.* 2017 and Mahanty *et al.* 2018). The planktivorous chorus was observed for a short period generally after the *Terapon theraps* chorus timing. The concatenated PSDs of time series within also show fish sounds (indicated as ‘4’) [Fig. (3a, b)], and we could match them with available data (Luczkovich and Sprague, 2022). They had discussed about recordings of unknown sounds using glider system and named them as “buzz” call. The reported characteristics are very much in accordance with our ‘Type A’ data.

Out of the 371 recorded 60 seconds sound files, 195 belong to the fish chorus period (14:00 – 02:45 h) when mean peak intensity levels of the PSDs are higher than the levels obtained within the timings (03:00 -13:45h). Beside three boat sounds, we have identified 79, 69, 22, 22 fish sound samples named as Sciaenidae, *Terapon theraps*, planktivorous and Type A respectively during the fish chorus period i.e., during dominant fish sounds. Whereas within (03:00-13:45h) we possess 176 data files, and out of them we identified 139 calls as *Terapon theraps* as their peak frequencies are found to be matching with the *Terapon theraps* though the dominant peak PSD levels of these *Terapon theraps* fish sounds are much less than the data obtained during the fish chorus timing (14:00-02:45h). Besides, limited number of samples for Sciaenidae and planktivorous sounds are found to be 05 and 06 respectively. Here, we have also identified 26 boat sounds based on the frequency ranges of the available archived data. Overall, we identified 84, 69, 28, 22, fish calls which are belonging to Sciaenidae, *Terapon theraps*, planktivorous and Type A respectively within the entire period. Also, we have identified 139 *Terapon theraps* fish calls (low peak PSDs level) and 29 boat sounds.

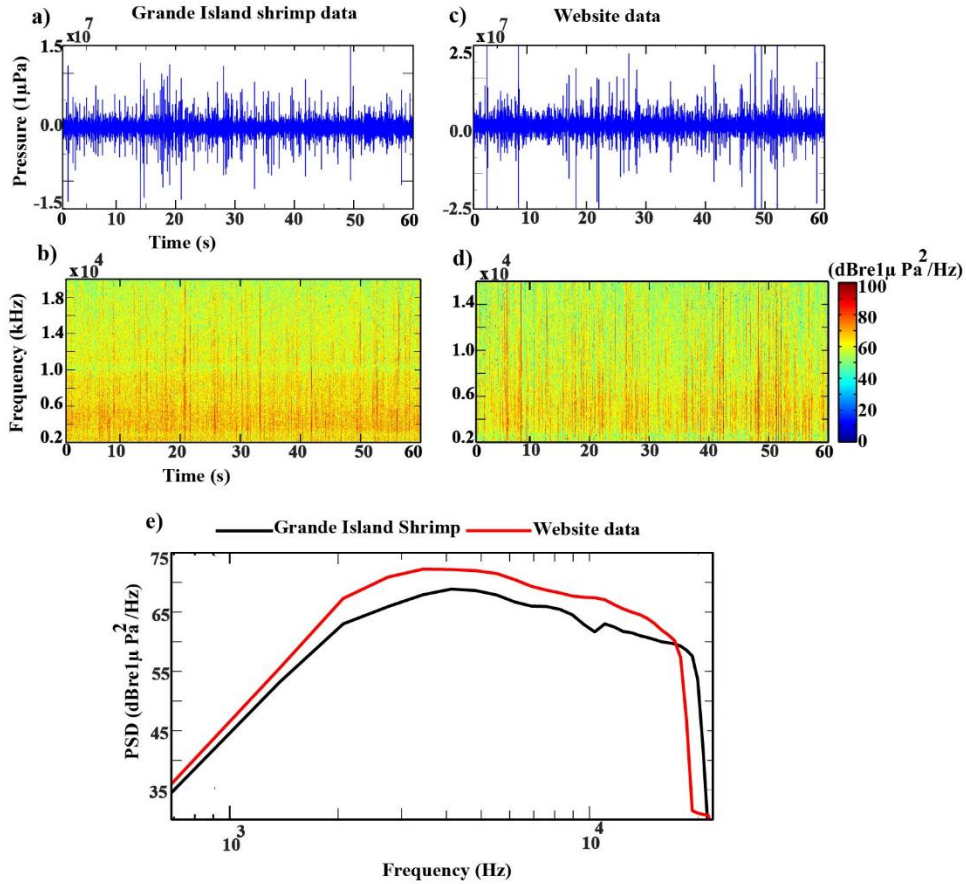


FIG. 4: Representative 60 seconds call from snapping shrimp and downloaded snapping shrimp data and their Spectrograms and PSDs (dB re $1 \mu\text{Pa}^2/\text{Hz}$): (a) representative data from present location and (b) data from website¹⁴, c) and d) spectrograms and e) PSDs of both the time series.

Distinctly different peak frequencies for Sciaenidae, *Terapon theraps*, Type A and planktivorous fishes are found to be 864.40 ± 28.19 Hz, 1502.95 ± 35.53 Hz, 1121.68 ± 95.35 Hz and 553.71 ± 53.35 Hz respectively. PSD peaks for Sciaenidae, *Terapon theraps*, Type A and planktivorous fishes are observed to be 92.54 ± 5.15 dB re $1 \text{ Pa}^2/\text{Hz}$, 97.55 ± 10.15 dB re $1 \text{ Pa}^2/\text{Hz}$; 80.68 ± 02.78 dB re $1 \text{ Pa}^2/\text{Hz}$ and 94.31 ± 08.35 dB re $1 \text{ Pa}^2/\text{Hz}$ respectively. Whereas, mean values of the peak PSD and frequency for *Terapon theraps* samples recorded during the timing (03:00-13:45h) provide (76.32 ± 02.78 dB re $1 \text{ Pa}^2/\text{Hz}$; 1521.68 ± 53.52 Hz). A significant difference in peak PSDs (20.23 ± 07.73 dB re $1 \text{ Pa}^2/\text{Hz}$) is obtained in comparison with dominant *Terapon theraps* fish sounds during (14:00 – 02: 45h).

We identified invertebrate sounds e.g., snapping shrimps¹³ (Fig. 4) based on the comparison between the auditory records and archived data. We observed high-frequency peaks along with peaks related to fish vocalizations (Fig. 3b). Mean level of the dominant peak shows (75.95 ± 6.79 dB re $1 \text{ Pa}^2/\text{Hz}$) and its corresponding mean peak frequency (3629.65 ± 1011.26 Hz) within the range (2045.65

Hz to 5490.96Hz). We know that marine soundscapes are generally dominated by broadband sources when associated with aggregation of snapping shrimp (Bohnenstiehl *et al.*, 2016; Lammers *et al.*, 2008). Therefore, unlike narrow band fish sound, segmentation of the snapping shrimp calls is complex. Further, we have shown plots of the representative oscillogram, spectrogram of high-frequency shrimp band data. Oscillogram and spectrogram of the representative data from the Grande Island are shown (Fig. 4a, b), along with the oscillogram and spectrogram of the archived data¹⁴ of the snapping shrimps (Fig. 4c, d). A typical spectrum of a snapping shrimp has a peak frequency within the 2 and 5 kHz (Au and Bank, 1998; Radford *et al.*, 2008) is depicted in the presented PSDs (Fig. 4e) of the representative calls from both the locations. Knowlton and Moulton (1963) presented spectral characteristics of the snapping shrimp from 56 locations. Frequency peaks were reported to be within the 2000 Hz to 4000 Hz, and have six dominant shrimp species, including black tiger shrimps (*Penaeus monodon*), which is also reported¹⁵ from the present study area (Fig. 2e, f).

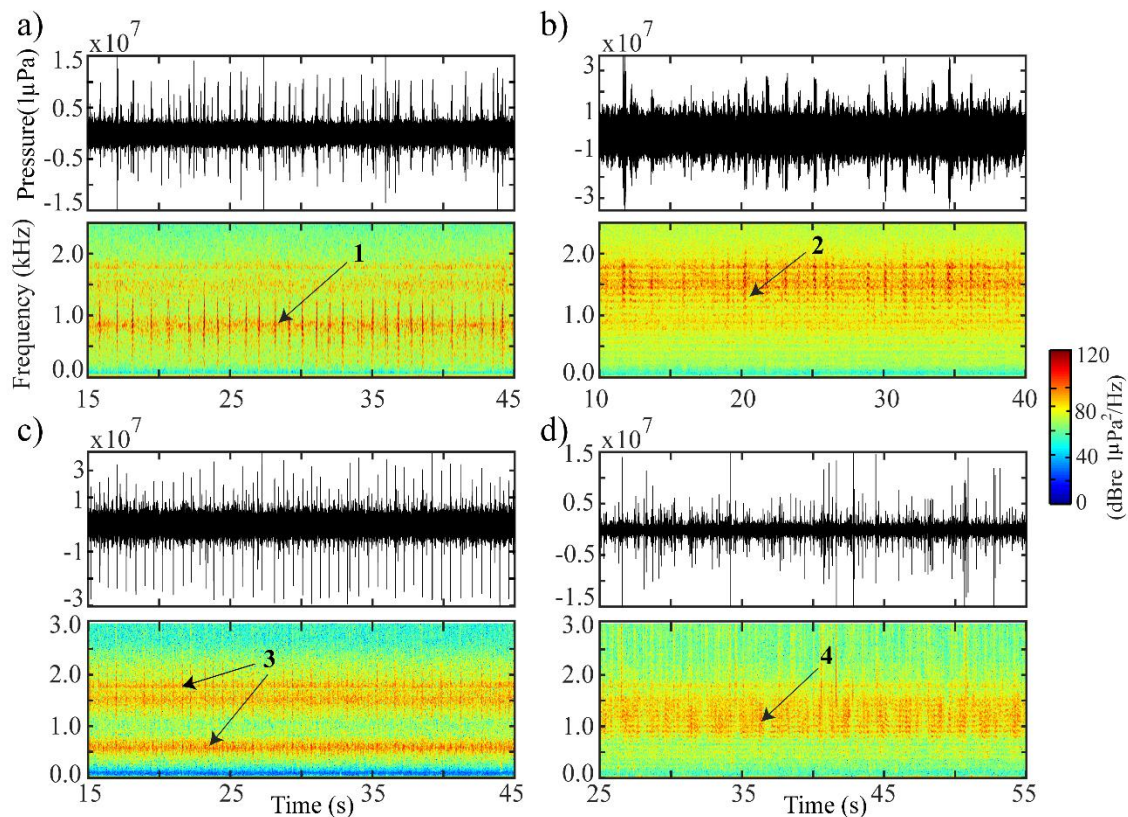


FIG. 5: (Color online). Pressure ($1\mu\text{Pa}$) versus time (sec), and Frequency (kHz) versus time (sec) for (a) Sciaenidae, (b) *Terapon theraps*, (c) planktivorous, and (d) Type A fish. Two arrows in figure (Fig. 5c) suggest planktivorous sound and its image reflection.

C. Segmentation of oscillogram

Once we identify different fish sounds, we apply segmentation technique to 60 seconds oscillogram to get individual fish species using temporal and spectral techniques. For visualization purposes, we present the spectrogram of 30 seconds oscillogram through the employment of the pwelch MATLAB function² (Fig. 5). Here, recorded fish sounds from 14:00 to 02:45 h show vocalizations from four fish families. In Fig. 5 (a-d), we show the oscillogram and spectrogram of the representative time series data for the fish sounds. The segmentation technique to obtain a single fish sound call is described in detail (Chanda *et al.* 2020a). The segmented method is applied here to limited numbers of representative oscillograms to get individual call details (call duration, number of pulses per call, inter-call durations, peak frequency of the PSDs) (Table II). Also, the PSD of a segmented calls have been used to estimate the dominant frequency peak levels of the calls (Fig. 6). Similarly, PSD and corresponding peak frequencies of the segmented calls for four identified fish sounds are shown in Fig. 6(a-d).

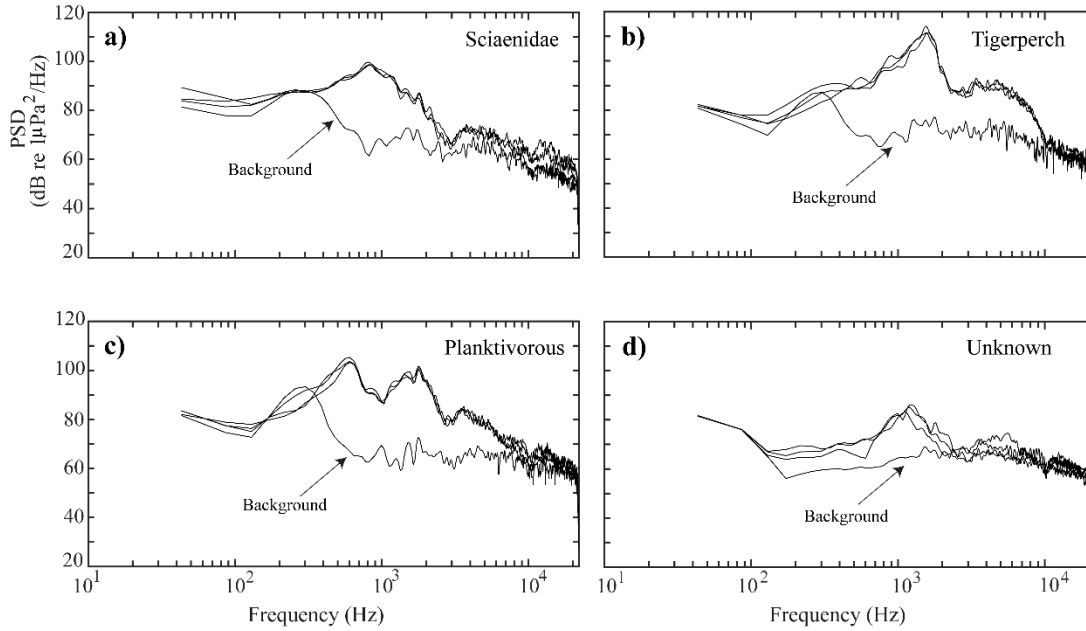


FIG. 6: PSD (dB re 1 $\mu\text{Pa}^2/\text{Hz}$) for (a) Sciaenidae; (b) *Terapon theraps*; (c) planktivorous, and (d) Type A fish.

Segmented calls' temporal and spectral parameters reveal Sciaenidae (Ramcharitar *et al.*, 2006). We have compared the parameters (call duration, number of pulses per call, inter-call durations, peak frequency of the PSDs) of present data with the archived Sciaenidae calls¹⁶ (Table II). The mean peak frequency of the PSDs of segmented calls is $(848 \pm 20.70 \text{ Hz})$ (Fig. 6a). Moreover, temporal and spectral parameters of the representative data for Sciaenidae calls are found to be matching with the

parameters of the archived data (Table II). Similarly, the oscillogram, spectrogram (Fig. 5b) and PSD of segmented call (Fig. 6b) confirm that the sounds were from *Terapon theraps* fish species of the Tiger Perch family (McCauley and Cato, 2000; McCauley, 2012; Mahanty *et al.*, 2018; Parsons *et al.*, 2017; Chanda *et al.*, 2020a) (Table II). The PSDs of segmented calls (Fig 6b) provide the mean value of the frequency peak of the PSD of individual call (1541.40 ± 18.13 Hz) (Table II).

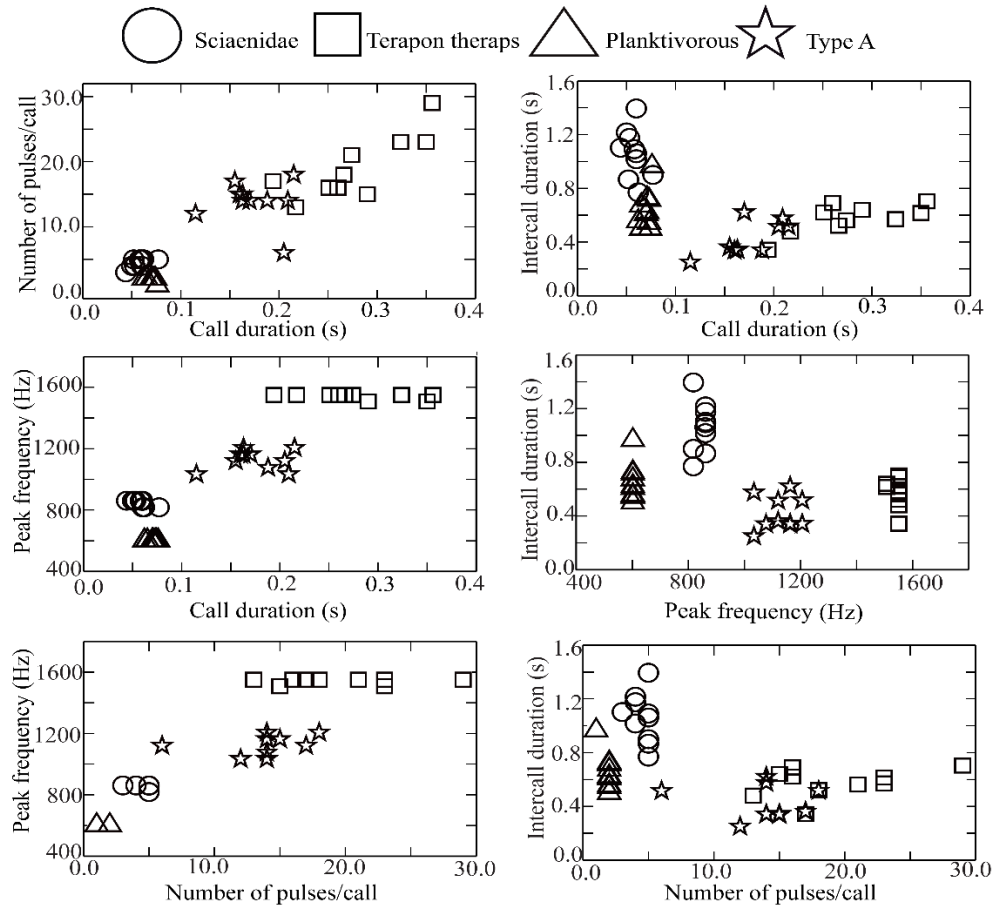


FIG. 7: Scatter plot of the parameters of the four fish sound types obtained using segmentation techniques to 60 seconds oscillograms. Four parameters are: Call duration (s), no. of pulses/call, intercall duration (s), and peak frequency (Hz).

Planktivorous fish type recorded here is a small fish species and it vibrates by its swim bladder with a single strike (McCauley and Cato, 2000). They produce sounds during feeding processes and drive their swim bladder to track plankton patches (McCauley, 2012). Planktivorous fish chorus investigated here are similar to those recorded around the South Eastern Arabian Sea (Mahanty *et al.*, 2018), Northern and Western Australia (McCauley, 2012; Parsons *et al.*, 2017). The PSD of a single segmented call of the planktivorous fish shows the spectral peak at (602.00 ± 0.0 Hz) (Fig. 6c), which matches well with the data of South Eastern Arabian Sea (Mahanty *et al.*, 2018), and Maret Island (McCauley, 2012). The oscillogram and spectrogram of planktivorous fish chorus recorded in this

study are shown (Fig. 5c), and single call details are tabulated (Table II). The planktivorous fish chorus was strongly influenced by the moon phase. We need to have continuous data acquisition at least for one month to carry out such investigations. The presence of notches in the call spectrum suggests the fishes' position in the water column during sound signal production (Fig. 6c). We observed notches at around 1000Hz in our study location (Fig. 5c). In present analyses, we reported the depth of the sound-producing planktivorous to be 1.5 m below the sea surface, which is calculated based on the expression given in McCaulay, (2012). Here, we present the oscillogram, spectrograms (Fig. 5d), and PSDs (Fig. 6d) for the representative calls of a fish considered Type A in this study. The PSD of the segmented call shows a frequency peak at $(1128.60 \pm 63.45 \text{ Hz})$, and call duration, number of pulses per bout, and inter-call duration are also given (Table II).

The scatterplots [Figs. 7a–7f] between the temporal (call duration, number of pulses/calls, and inter-call duration) and spectral (PSD peak frequency) parameters for the segmented call of the ten representative data show cluster. Here, the importance of temporal characteristics of the fish sounds are depicted, and these differences are due to the differences between the calls of the individual fish species' temporal characteristics, including the structure and rise time of the single pulses for the Sciaenidae, *Terapon theraps*, planktivorous and Type A fish families. Differences in temporal and spectral characteristics for the four fishes from the present location are observed.

D. Assessing biodiversity

In this section, we compute two acoustic metrics (ACI and AEI) and root-mean-square sound pressure levels (SPLrms) for the three frequency bands (full, fish, and shrimp bands). We perform a statistical test comparing datasets of SPLrms. A Kruskal-Wallis H test¹⁷ shows a statistically.

TABLE II. Average call parameters of the individual fishes estimated using segmentation technique and data from literatures to support present findings

Fish type	Frequency peak (Hz)	Highest Peak level (dBre 1 Pa ² /Hz)	Inter call duration (sec)	No. of pulses /call	Call duration (sec)
Type A	1128.6±63.46 ^a	83.10±1.15 ^a	0.420±0.12 ^a	13.90±3.25 ^a	0.174±0.03 ^a
Sciaenidae	848±20.7 ^a	101.10±0.875 ^a	1.0593±0.18 ^a	4.50±0.71 ^a	0.0576±0.01 ^a
	627.0±35.83 ⁸		0.9040±0.44 ⁸	3.83±1.60 ⁸	0.0583±0.02 ⁸
<i>Teraponthera ps</i>	1541.4±18.13 ^a	110.10±0.88 ^a	0.5742±0.107 ^a	19.10±4.8 ^a	0.2782±0.053 ^a
	500-1500 ^b	94-118 ^b	-----	9-16 ^b	0.079-0.105 ^b
	1000-2000 ^c	-----	-----	-----	~0.20 ^c
	600-1500 ^d	-----	-----	35-40 ^d	0.35-0.45 ^d
	1500 ^e	-----	-----	20-40 ^e	0.10-0.15 ^e
	500-3000 ^f	-----	-----	~51 ^f	~0.32 ^f
	1791.5±106.5 ^g	-----	-----	22.0±2.77 ^g	0.26±0.03 ^g
1768±73.7 ^h	-----	-----	19.50±2.0 ^h	0.25±0.03 ^h	
Planktivorous	Peaks at 602±0.00 and 1583±22.2; notch at 1000 ^a .	102.70 ±1.25 and 101.06±0.50 ^a	0.641±0.140 ^a	1.90±0.32 ^a	0.070±0.005 ^a
	Peaks within 400-700 no notch ^b .		-----	-----	-----
	Peak at 900; notch at 1200 (Maret Island) ^c	126-135 ^c	-----	-----	~0.01 ^c
	Peak at 600; notch at 1500 (Scott reef) ^c	94-118 ^c	-----	-----	-----
	Peak at 900; no notch ^e	-----	-----	-----	-----
	Peak at 550; notch at 1000 ^f	-----	-----	-----	-----

References: ⁸<https://www.fishbase.se/physiology/FishSoundsList.php?ID=1180>

^aData from present study, ^bMcCauley and Cato (2000);

^cMcCauley (2012); ^dMahanty *et al.*, (2015), ^eMahanty *et al.* (2018); ^fParsons *et al.* (2017);

^gChanda *et al.* (2020a); ^hChanda *et al.* (2020b).

significant difference between the SPLrms data. The test statistics result reveals (Chi-square = 101.242, $p = 1.036136e-22$, $df = 2$). Since the p-value is less than 0.05, it indicates a statistically significant difference between the three independent SPLrms values. We carried out the Dunn Post Hoc test and found p-values between the bands (full vs. fish: $p = 4.044192e-13$), (full vs. shrimp: $p = 1.572406e-21$) and (fish vs. shrimp: $p = 0.019561$). Similar tests conducted for AEI profiles for three bands reveal (Chi-square = 723.7256, $p = 6.998042e-158$, $df = 2$) and we found p-values using Dunn Post Hoc between the bands (full vs. fish: $p = 9.416350e-141$), (full vs. shrimp: $p = 2.223392e-06$) and (fish vs. shrimp: $p = 1.046795e-93$). Again, the test conducted for ACI values for three bands revealed (Chi-square = 935.111954, $p = 8.770394e-204$, $df = 2$), and calculated p-values using Dunn Post Hoc between the groups (full vs. fish: $p = 6.857810e-205$), (full vs. shrimp: $p = 1.788483e-52$) and (fish vs. shrimp: $p = 1.788483e-52$). Statistical tests allow us to ensure that there are significant differences in the datasets.

Temporal analysis of soundscape data using acoustic indices (two-hour average for no. of survey days) of the study region is presented (Fig. 8). This observation suggests two different regimes in terms of fish vocalization, i.e., dominance and reduction of SPLrms values with respect (02:00-04:00h) (Fig. 8). The variations of the mean SPLrms for the three frequency bands also support the view expressed in the previous section regarding the clear-cut demarcation in ambient sounds while describing the soundscape components. We found that the highest value of the mean SPLrms value (135.50 ± 2.73 dB re 1 μ Pa) within the (20:00-22:00 h) for full-band, which reaches its lowest value (113.35 ± 1.06 dB re 1 μ Pa) within (04:00-06:00h) (Fig. 8a). We have also observed a similar trend for SPLrms values for fish and shrimp-bands. We present box plots of the SPLrms for the full, fish, and shrimp bands for dominant ambient fish sound within the timings (14:00 – 02:45 h) and low ambient sounds within (03:00-13:45 h) (Fig. 9). For dominant ambient sound, we observe that the comparable mean SPLrms value of the full-band data (125.45 ± 7.34 dB re 1 μ Pa) in comparison with the fish-band (123.72 ± 8.55 dB re 1 μ Pa) and shrimp-band (117.49 ± 6.46 dB re 1 μ Pa) (Fig. 9a-c). Here, higher SPLrms suggest the contribution of the higher energy due to the fish vocalizations for full and fish- bands.

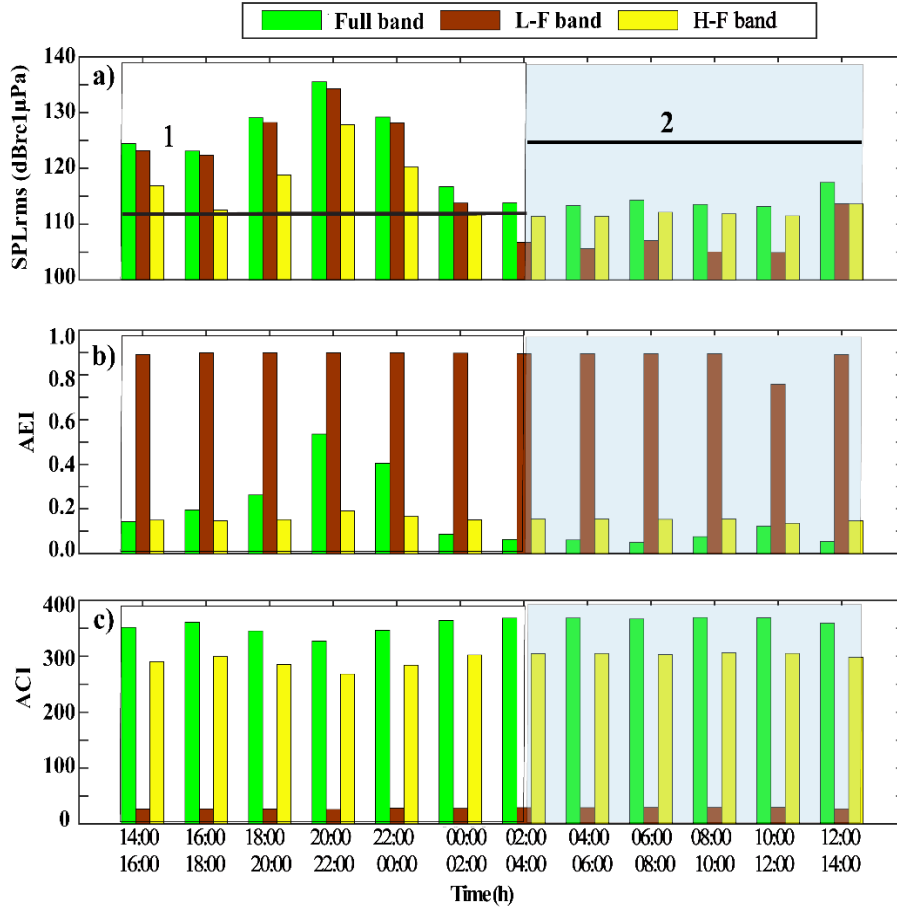


FIG. 8: Temporal variation of (a) SPLrms (dB re $1\mu\text{Pa}$), (b) AEI and (c) ACI for three different bands. Numbers 1 shows timings for a particular fish sound when ambient sound is dominant, and 2 (shaded area) shows the timings when ambient sound is low.

Box plots of the SPLrms within the timing (03:00 – 13:45 h) indicate mean SPLrms values for full-band (115.34 ± 4.47 dB re $1\mu\text{Pa}$) are higher than the fish-band (108.43 ± 6.62 dB re $1\mu\text{Pa}$) and shrimp band data (112.41 ± 2.61 dB re $1\mu\text{Pa}$) (Fig. 9a-c). However, these values are lower than the higher ambient sound (14:00-02:45 h). We also observe that the distribution of the SPLrms for full and fish bands is normal (Fig. 9a, b), whereas the difference between the median and mean values for the shrimp band suggest data is non-normal for the shrimp band within the (14:00 – 02:45 h) (Fig. 9c). During the period (03:00 – 13:45 h), variabilities are non-normal for all the three bands (Fig. 9a-c). Considerably lower SPLrms values for full- and fish band data within the period (03:00 – 13:45 h) suggest poor fish vocalizations. Whereas, the SPLrms of the shrimp band sounds (112.41 ± 2.61 dB re $1\mu\text{Pa}$) during the period (03:00- 13:45h) (Fig. 9c) is lower than the SPLrms (117.49 ± 6.46 dB re $1\mu\text{Pa}$) of the fish chorus (14:00 – 02:45 h).

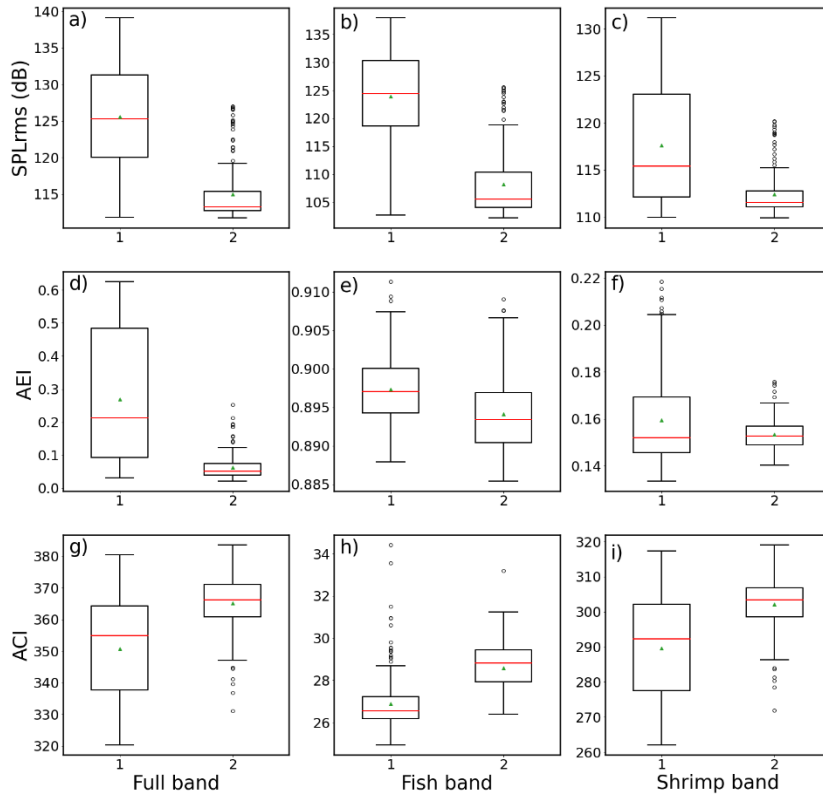


FIG. 9: Boxplots of the acoustic metrics: (a) SPLrms (dB re $1\mu\text{Pa}$), (b) AEI and (c) ACI for full-, low-frequency and high-frequency bands for two different timings: 1. When ambient sound is dominant (14:00- 02:45 h), 2. Low ambient sound (03:00 – 13:45 h).

In (Fig. 8a), unshaded bar graphs towards left hand side indicate vocalization timing when ambient sound is dominant. Here, a general inventory of the fish and boat sounds are reported within the timings (14:00-02:45 h) (Fig. 10). During the dominant fish chorus time (14:00-02:45h) for full-band, the mean SPLrms value for *Terapon theraps* is the highest (129.49 ± 8.59 dB re $1\mu\text{Pa}$), and thereafter we found planktivorous (127.23 ± 4.85 dB re $1\mu\text{Pa}$), Sciaenidae (124.56 ± 4.01 dB re $1\mu\text{Pa}$), and Type A (115.23 ± 1.07 dB re $1\mu\text{Pa}$) (Fig. 10a). Boat sound during the fish chorus timing for full-band is found to be (117.50 ± 5.50 dB re $1\mu\text{Pa}$). We find negligible reduction in the mean SPLrms value for fish band data (Fig. 10b), Within the timings (03:00- 13:45h), the SPLrms are low (Fig. 9a, b). Comparatively lower SPLrms for fish band data are found in comparison with the dominant ambient sound (14:00 – 02:45 h). The trend of the fish-band SPLrms in the shaded section of the bar graph (Fig. 8a) indicates the lower SPLrms values.

Temporal analysis of AEI for fish-band data shows the highest value (0.898 ± 0.002) throughout, which is nearly constant (Fig. 8b). Interestingly, the trend of the AEI values for full-band shows similar variations to that of SPLrms. There is an increase from (14:00-16:00 h), and the highest value ($0.535 \pm$

0.013) is obtained during (20:00-22:00 h) and post which reduction. However, the variations of AEI values for the high-frequency shrimp band are insignificant compared to the full-band data. We present box plots of the AEI for the three bands during the dominant fish chorus timing (14:00-02:45h) (Fig. 9d-f). We observed from the box plots that the distributions are normal for the fish band, whereas non-normal for full and shrimp-band data (Fig. 9d and f). The mean AEI value of the fish-band data is observed to be the highest (0.897 ± 0.004) (Fig. 9e) in comparison with the full (0.265 ± 0.193) (Fig. 9d) and shrimp (0.159 ± 0.019) (Fig. 9f) bands. The presented box plot within the (03:00-13:45 h) timings show the lowest value of the AEI (0.080 ± 0.124) for full band, which suggests equal distributions of sound sources (Fig. 9d). Higher fish band values (0.895 ± 0.00) are comparable with the high ambient sound conditions (Fig. 9e).

We estimated AEI values for vocalization sources like Sciaenidae, *Terapontheraps*, planktivorous, and Type A (Fig. 10 c, d). During the dominant fish chorus time (14:00-02:45h), AEI of the *Terapontheraps* fish is higher (0.405 ± 0.206) for full band data. Similarly, AEI values for planktivorous, Sciaenidae and Type A are found to be (0.335 ± 0.168), (0.182 ± 0.103), and (0.076 ± 0.033) respectively (Fig. 10c). For full-band data, we found lower values of AEI (0.087 ± 0.062) for the boat sound. Whereas for fish-band data, we find higher but comparable values of AEI for Sciaenidae, *Terapontheraps*, planktivorous and Type A and boat sound (Fig. 10d). Computed AEI values follow a different pattern for fish-band data than the full-band data.

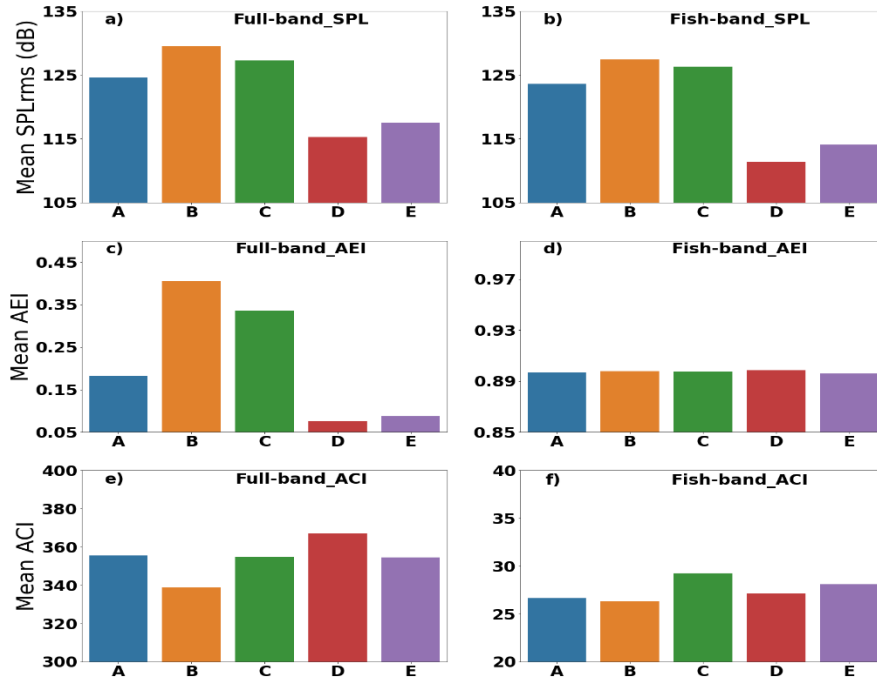


FIG. 10: Bar graph indicate a) SPLrms (dB rel μ Pa) of the full-band and b) low-frequency fish band data; c) AEI of the full-band and d) low-frequency fish band data; e) ACI of the full-band and f) low-frequency fish band data for (A) Sciaenidae, (B) *Terapon theraps*, (C) planktivorous, (D) Type A fish and (E) Boat sound (when ambient sounds are dominant within the 14:00-02:45 h).

Two-hour average values show moderate ACI for full-band and shrimp-band and the lowest value for fish-band (Fig. 8c). The temporal data of ACI show slow variation for three frequency bands. However, reciprocal behavior is seen compared to the AEI values for the fish and shrimp bands (Fig. 8b, c). For a particular band data, no significant changes in the ACI are observed irrespective of the presence or absence of fish chorus. Box plots of the ACI for the full, fish and shrimp bands are also presented (Fig. 9g-i) during the higher (14:00-0245h) and lower (03:00-13:45 h) ambient sound. Data distribution for the full, fish, and shrimp-bands show non-normal characteristics for both the ambient sound situations (Fig. 9g-i). We found that ACI values are the lowest (26.89 ± 1.36) for the fish band compared to the other two bands (Fig. 9g, i). Generally, the ACI values are the highest (350.82 ± 16.32) for full-band data, closely followed (289.62 ± 14.97) by the shrimp band sounds. Within the (03:00-13:45h) timings, we found that ACI values are to comparatively higher for full (365.79 ± 10.56) and shrimp bands (302.76 ± 9.00), whereas the lowest ACI (28.734 ± 1.445) for fish bands.

Estimated ACI values for vocalizing fishes like Sciaenidae, *Terapon Theraps*, planktivorous, and Type A during the dominant fish chorus time (14:00-02:45h) for full band data are shown (Fig. 10e). ACI values of the *Terapon theraps* fish are comparatively low (338.66 ± 18.03). For planktivorous,

Sciaenidae, and Type A the ACI values are (354.86 ± 10.09) , (355.64 ± 9.88) , and (367.12 ± 6.86) as shown in the bar graph (Fig. 10e). Boat sound during the fish chorus timing for full-band is found to be (354.59 ± 16.54) . For fish band data, we found ACI values for Sciaenidae, *Terapon theraps*, planktivorous, and Type A and boat sounds are low (within 26.64 ± 0.448 to 29.23 ± 2.088). (Fig. 10f).

E. Data clustering

Soundscape analysis reveals two distinct regimes i.e., dominant fish vocalizations within the period (14:00 to 02:45 h) and (03:00 to 13:45 h) (Fig. 3a and c) and their characteristics. Initially, we carry out classification, when fish chorus is dominant within the timing (14:00-02:45 h) i.e., classification of the fish species when ambient sound is dominant. Application of PCA on the dataset indicates that first five PCs contributes to 94.80% of the total variance (Fig. 11a, b). Kaiser's rule (Kaiser, 1960) suggests that only those PCs are retained, whose variance exceeds 1, and accordingly, PCs 1-5 only qualify this rule. Once minimum number of PCs are identified as feature vectors (best representing the dataset), the goal is to project the original data on to the derived feature space. Graphical representation of the derived data in feature space helps in identifying the underlying clusters in the original dataset (Fig. 11c). We present biplot between PC1 and PC2, where PC1 and PC2 explains 74.77% of the total variance, i.e., PC1 and PC2 contributes 63.26% and 11.51.0% respectively. In order to derive the desired subspace with reduced dimensions, PCA has successfully been implemented to the large dimensional (~12-odd features) input space to five principal components.

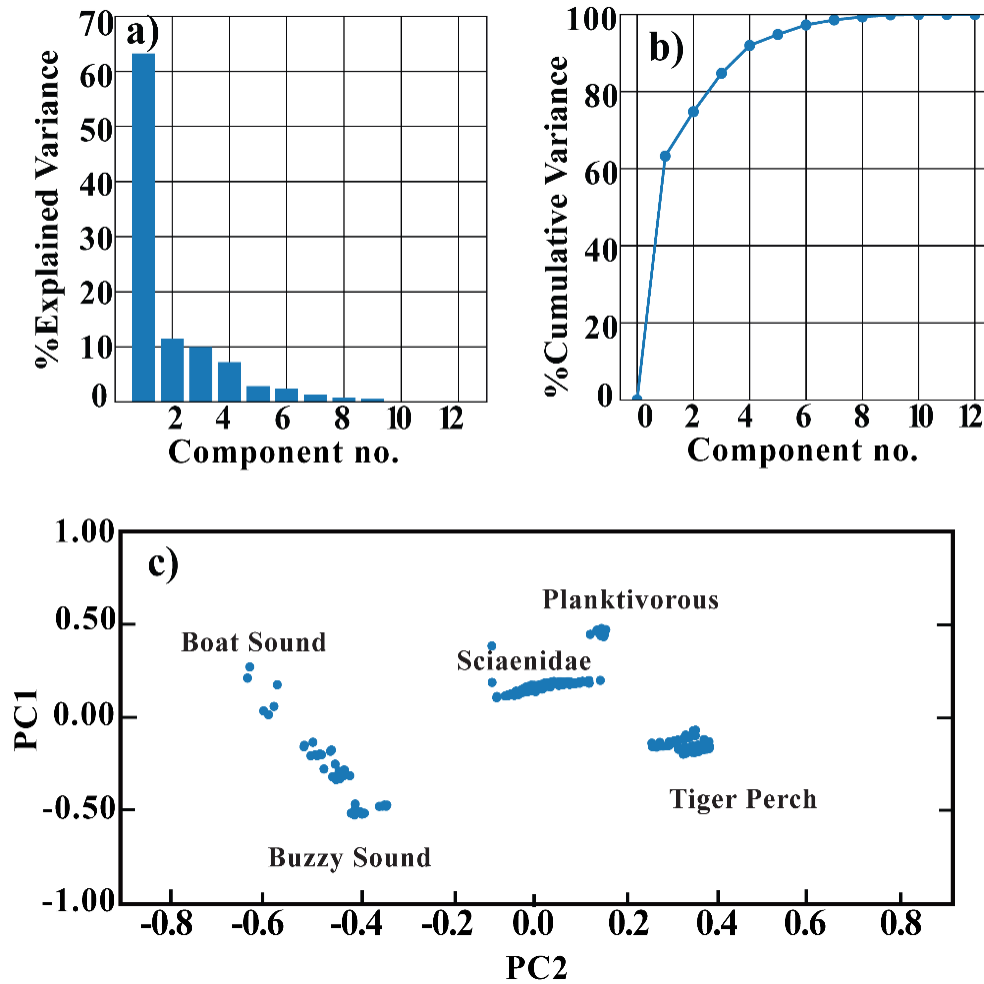


FIG. 11: (Color online) Clustering of the recorded data during the dominant ambient sound (14:00-02:45 h): (a) Principal Components versus Variances (%) are shown. Here PC1-PC5 together explain 94.80% of the variance in the data, (b) Cumulative variance, and (c) PCA based scatter plot between the PC1 and PC2 shows clustering of the four types of fishes during the dominant ambient sound.

We implemented the K-means clustering method (Ding and He, 2004) to the derived reduced subspace. K means clustering method is an algorithm where the number of clusters are decided a priori. This algorithm is generally based on defining cluster centers to represent clusters by minimizing the squared error function. Computing and plotting the WSval measures in percentage with respect to the number of clusters (Kval) enables to define the optimum number of clusters via the "elbow" method. Accordingly, five cluster centers were assigned to our dataset as a target. Implementing the K-means algorithm then allocates every data point to the nearest cluster, while keeping the centers as small as possible (Fig. 11a). The K-means clustering process here is an unsupervised method as it is not based on any predictive/response variable (Cheung, 2003). In other words, the clustering is based on relationship between all the observations without being trained by a response variable.

Confusion matrix¹¹ is a matrix formed by checking the actual values and predicted values of observations in the datasets. The actual values obtained from the label data 79, 69, 22, 22 and 3 are Sciaenidae, *Terapon theraps*, Type A, planktivorous fishes, and boat sounds respectively. Predicted values are obtained through unsupervised classification technique involving K-means clustering. For Sciaenidae, *Terapon theraps*, Type A, planktivorous fishes and boat sounds, the predicted values are 76, 53, 41, 19 and 6 respectively. We obtained reasonable classifications such as 96.20%, 76.81%, 86.36.8%, and 100% for Sciaenidae, *Terapon theraps*, planktivorous, and Type A fishes. We also obtain 100% classification for boat sound. We calculated the classification percentage based on the number of correctly classified sounds on the data employed (Table III). For example, unsupervised classification techniques successfully classify all twenty-two Type A sound samples, i.e., 100% classifications. At the same time, classification for Sciaenidae fish sound (96.20%) suggest negligible overlapping with Type A (3.80%). It is 86.36% classification for planktivorous sounds, having 13.63% overlap with boat sounds. Similarly, for *Terapon theraps* fish sound classification (76.81%), overlapping is significant (23.18%) with recorded Type A sound.

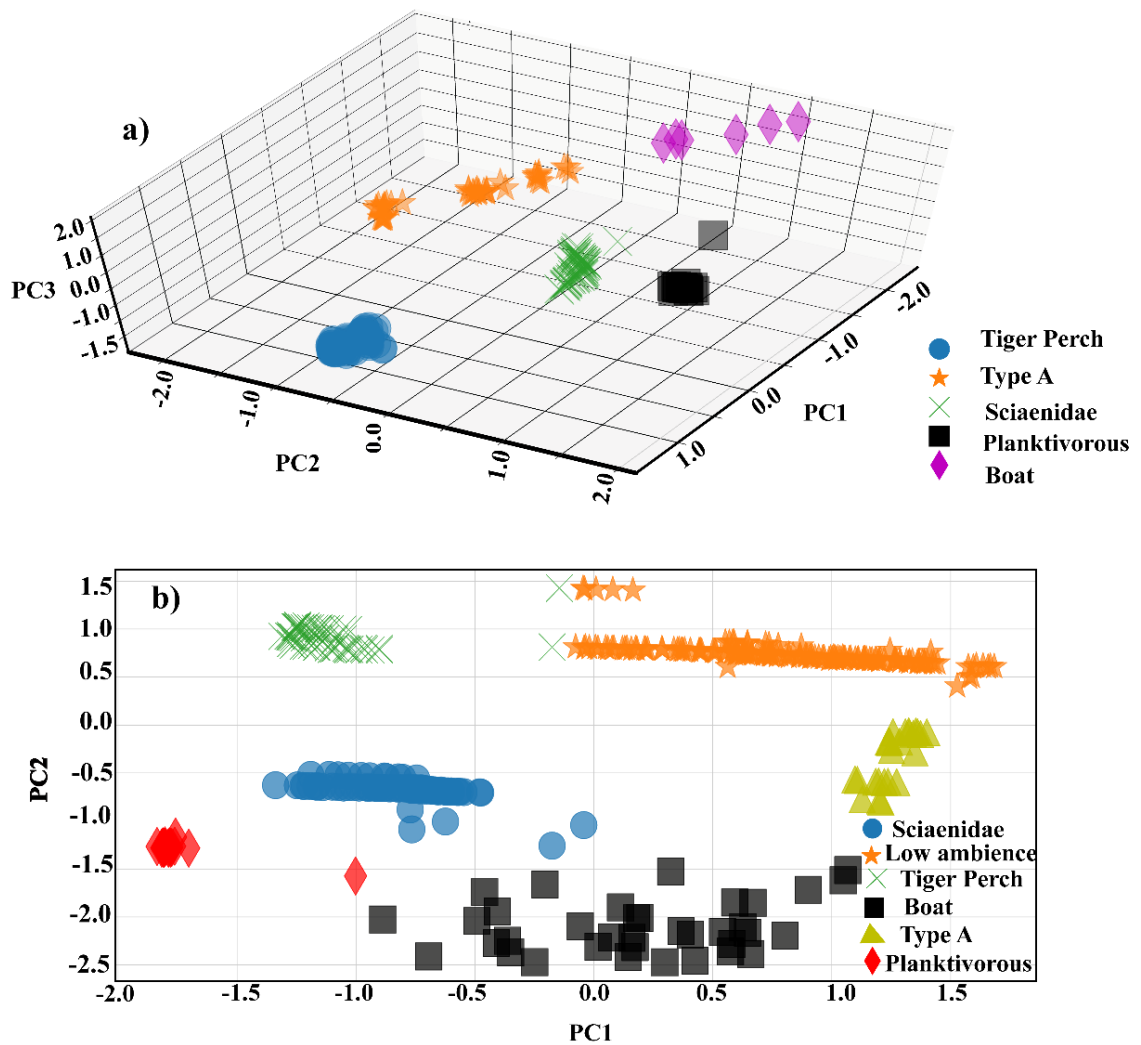


FIG. 12: (Color online) Clustering of the recorded data: (a) 3-D plot shows the clustering of the four types of fishes and boat sounds during dominated ambient sound within (14:00-02:45 h), (b) Plot between PC1 versus PC2 shows six clusters for entire data sets involving low ambient sounds.

We also presented classification results when we involve entire fish sound data i.e., 92:30 h, which include total data files considered for present study i.e., 371. PCA of the dataset indicates that first five PCs contributes to 91.75% of the total variance, and accordingly, PCs 1-5 only qualify. Graphical representation of the derived data in feature space helps in identifying the underlying clusters in the original dataset (Fig. 12b). We present biplot between PC1 and PC2. Here, PC1 and PC2 explains 69.98% of the total variance, where PC1 and PC2 contributes 58.02% and 11.953% respectively.

TABLE III. Classification results (in %) for four fish vocalizations and boat sound

Sound type	Fishes classified as in %					
	Matrix	Type A	Sciaenidae	<i>Terapon theraps</i>	Planktivorous	Boat
Type A	100	0	0	0	0	0
Sciaenidae ^a	0	96.20	3.80	0	0	0
Tiger Perch ^b	23.18		76.81	0	0	0
Planktivorous	0	0	0	86.36	13.64	
Boat	0	0	0	0	0	100

We employ K-means algorithm to obtain number of cluster center and accordingly we obtain six classes through the employment of the elbow method. The actual values obtained from the label data are Sciaenidae (84), *Terapon theraps* (69), planktivorous (28), Type A (22), and boat (29) sounds. Moreover, additional 139 datasets labeled as *Terapon theraps* sounds based on their peak frequency (1521.88 ± 53.52 Hz), however, possess considerably lower peak PSDs (76.32 ± 5.00 dB) than the *Terapon theraps* fish sounds classified during the higher ambient sound within (14: 00 – 02:45 h). These 139 values of *Terapon theraps* are acquired during the (03:00-13:45h) when ambient sound is low. Through unsupervised method involving K-means clustering, predicted values for Sciaenidae, *Terapon theraps*, Type A, planktivorous, *Terapon theraps* (recorded when ambient sound was low) and boat sounds are found to be 83, 55, 25, 19, 153 and 36 respectively. Using confusion matrix, we obtained classification results for Sciaenidae, *Terapon theraps*, Type A, planktivorous, *Terapon theraps* (low ambient sound), and boat sounds are 96.42%, 76.81%, 100%, 67.85%, 98.56% and 100% respectively. *Terapon theraps* species data within the low ambient sound, forms separate cluster than the *Terapon theraps* sounds acquired during the higher ambient sound. 98.56% classification success has been obtained for *Terapon theraps* fish sound data recorded during the lower ambient sound, which have minor overlaps (1.43%) with *Terapon theraps* sounds acquired within the higher ambient sound.

IV. DISCUSSION

The Grande Island Archipelago off the coast of Goa, is known for its fish diversity, and about 115 species of fish have been recorded in its vicinity (Sreekanth *et al.*, 2019). Chary *et al.*, (2020) had

recorded 21 out of fifty-five fish soniferous species. From the present investigation, we observe that the peak frequency of the PSDs along with other acoustic metrics for the fish vocalization is important for fish identification (Fish and Mowbray, 1970; Tricas and Boyle, 2014). Besides the peak frequency, level of the peaks of the PSDs are also important (Staaterman *et al.*, 2013), and accordingly we have labelled our data (preliminary identification) based on the peak PSDs, which is important for knowing the performance of machine learning technique. In this regard, onsite passive acoustic recordings related difficulties are also not unknown. For example, fish sounds are usually not sharply tuned. The differences in amplitude at different frequency peaks due to the potential addition of spurious variability to a dataset are quite likely (Fine and Permentier, 2015). Also, the fishes were at varying distances from the hydrophones so that other signals would have had their frequencies dispersed more than the signals emitted closer to the hydrophones. Due to the lack of systematic fish sound archival facilities, most fish sounds are not always identified at a species level (Desjonqueres *et al.*, 2015). Challenges also lie with the onsite fish photographic data to validate the acquired fish sound. At times, the employment of divers is practiced, which has restrictions up to a limited water depth and avoidance of fishes to the divers (Kaplan *et al.*, 2015). Underwater cameras installed for a long time, were always unsuitable since fishes may avoid strobe light (if utilized). Moreover, turbid water may hinder capturing the photos in the locations away from reefs and the distance between the camera and soniferous fishes (Labrosse *et al.*, 2002). Besides, there is also a common practice to acquire fish sound for characterization using a laboratory-based tank facility (Akamatsu *et al.*, 2002). Popper *et al.*, (2019) reported fluctuations in auditory threshold for fishes measured in laboratory tank facilities. However, fish sound data acquisition using tanks is also a concern due to the resonance and propagation effects (Mann and Lobel, 1997). Ladich (2015) investigated the changes in the spectral peak of fish sound and hearing sensitivity with growth, i.e., ontogenetic development of sound production and auditory sensitivity.

Within the oscillogram lengths of 60 seconds, there is a possibility of mixing between the fish family/species choruses. The peaks of the oscillograms' PSDs represent the sound energy contributed due to fish species vocalization or any other sounds (anthropogenic, geophony, and invertebrate). Investigations based on the peak frequency of the PSDs that employ a 60 seconds oscillogram indicate that the calls are truly different in terms of structure as a challenge. Therefore, to validate fish call parameters, segmentation of the 60 seconds oscillogram data is essential (Chanda *et al.* 2020a; Towsey *et al.* 2012). Detailed segmentation applications are dealt with elsewhere (Chakraborty *et al.* 2015). Segmentation of the oscillogram data (Chanda *et al.*, 2020a) is vital to characterize the individual fish call parameters (call duration, number of pulses per call, inter-call duration, and

frequency peak of the PSDs) (Rountree *et al.*, 2006) within the soundscape. Once calls are identified at a broader level, detailed characterization of the individual call allows to employ segmentation technique, and temporal and spectral call parameters of segmented calls are adopted. If soundscape lacks clarity, then segmentation of calls is also difficult. Sometimes, lack of clarity in soundscape is possible if receiver positions within the proximity of the sound sources lie in the very shallow water (Fig. 3.12b: soundscape taken at reef¹⁸). Under such nearfield conditions (Putland *et al.*, 2019), particle motion sensors are preferable, and Acoustic Vector Sensor (AVS) can be a helpful passive acoustic data acquisition system¹⁹. Interestingly, soundscape data presented in this research shows clarity (Fig. 3a), and we have no difficulty in fish call data segmentation.

Even in an open water environment where anthropogenic and substrate vibrations are dominant, the changes in vocalization quality for fishes are widespread (Hawkins *et al.*, 2021). The soundscape is the ensemble of ambient sound, including sound events associated with specific locations at a particular time, found in all vertebrates (Hawkins and Popper, 2017). Sound can be detected (a) as pressure fluctuations in the medium above and below the local hydrostatic pressure (sound pressure) and (b) by the back-and-forth motion of the medium, which is referred to as particle motion (Popper and Hawkins, 2019). A critical issue to understand and appreciate is that while many fish species (like all marine mammals) are likely to detect sound pressure, particle motion is also vital to fishes and invertebrates (Popper and Hawkins, 2018). Many fishes are not sensitive to sound pressure, and in monitoring soundscapes, it is essential to measure the particle motion levels and the SPLs. Particle motion is often generated due to substrate vibration (Hawkins *et al.*, 2021), resulting from water flow, ground rumbling, and some anthropogenic sources. Even fishes that can detect sound pressure, they may have to use particle motion to detect the direction from which sounds are arriving (Hawkins and Popper, 2018). Particle motion measurement methods are initiated (Urquhart *et al.*, 2006), and currently perfections are underway (Lumsdon *et al.*, 2018).

We categorized our recorded sounds from present location where we identified Sciaenidae, *T. theraps*, planktivorous, and unknown fish named as a Type A. Even if the temporal and spectral parameters are estimated employing acquired sound data of Type A fish (Table II), we could not confirm identification of the fish family/species using video records obtained by divers during the daytime. Due to the simultaneous presence of several fish families (Sreekanth *et al.*, 2018), identification of a particular fish and their sounds were problematic. For fish identification, comparing an archived fish sound file with video-recorded fish data was unattainable, especially from tropical water environments. Irrespective of the limitations mentioned here, we identified choruses of Sciaenidae (croakers and drums) that are unique in the diversity of their sound production mechanism

(Ladich and Fine, 2006). In this study location, populations of Sciaenidae species were high during the pre-monsoon compared to the post-monsoon (Hegde *et al.*, 2016). Velip and Rivonkar (2018) reported that 18% of fishes from the sciaenid family. Here, fish sound temporal and spectral parameters reveal matching with the shorthead drum (*Larimus breviceps*²⁰) (Borie *et al.*, 2021). Though we cannot confirm the species, our recorded sound file nonetheless confirms that the sound is from the Sciaenidae family. Interestingly, as per the coverage map, such species are available in the coastal areas of South America. However, the fish base map also indicates the suitability of the habitats for such species off Indian coasts.

Chanda *et al.* (2020a) had reported details of the Tiger Perch fish sound identifications from this location¹⁶. We could compare our representative data with the online Tiger Perch sound file through link^{6,12}. Similarly, the *Terapon theraps* species of Tiger Perch from the Terapontidae family (order Perciformes) were found in the marine coastal, brackish, and freshwaters of the Indian and West Pacific Oceans of Australia, New Guinea, Indonesia, and the Philippines (McCauley, 2012; Mahanty *et al.*, 2018; Parsons *et al.*, 2017). Moreover, the fish species maps for *T. theraps* suggest widespread coverage around the Indian coast¹². For *T. theraps*, the sonic mechanism involves the swim bladder and the extrinsic muscles (Schneider *et al.*, 1964). A helpful comparison of *T. theraps* vocalization parameters with published Australian and Arabian Sea coast data were presented (Table II). A general look at the table and given fish details in the references, supports our fish sound identification (Sreekanth *et al.*, 2018).

The planktivorous species dominate reef sites and prefer waters with high plankton availability. The abundance of plankton is influenced by primary productivity, the incidence of solar radiation, wave action, and current (Clarke *et al.*, 2009). Sreekanth *et al.* (2018) reported 44 species of planktivorous from the Zuari estuary through detailed sampling. Paramount fishery importance of the planktivorous species from the Zuari estuary belongs to demersal and pelagic classes. The percentage contribution of the planktivorous group in the Grande Island area is reported to be around 31.7% among five trophic guilds (Sreekanth *et al.*, 2019). Planktivorous chorus comprises a single broadband pulse generated from swim bladder contraction (McCauley and Cato, 2000). They are attributed to the nocturnal period in which the fish feeds during the night, and their chorus type is related to the feeding behaviour (McCauley, 2012). Interestingly, the noise notch in the PSDs of planktivorous calls from the present location is similar to the data presented from Scott Reef and the Maret Islands in northern Australia (McCauley, 2012; Parsons *et al.*, 2017) (Fig. 6c). Mahanty *et al.* (2018) recorded a similar chorus with a spectral peak at 600 Hz. The notch was often present at around 900 Hz in shallow waters off the southwest coast of India during the pre-monsoon period.

Drawn bar graph (Fig. 8) for SPLrms, AEI and ACI for three frequency bands provides interesting information. For the higher AEI (Fig. 8b) and lower ACI (Fig. 8c) values for low-frequency fish-band data and lower AEI and higher ACI values for high-frequency shrimp band data characterizes reef or shallow site with clarity (Peck *et al.*, 2021; Eldridge *et al.*, 2018). Recently, Dimoff *et al.* (2021) had indicated inconsistent and insignificant ACI values, primarily due to its closeness to high energy systems like coral reef locations (Bohnenstiehl *et al.*, 2018). We did not encounter such inconsistency in ACI values, though the distance is higher (1870m) between the present location and the Grande Island landmass. Our research finds that combined ACI and AEI values are effective in characterizing the off-reef areas. ACI values having limited fluctuations are found for full-band data, whereas AEI values vary moderately. On a similar ground, Kaplan *et al.* (2015) presented a clustering pattern for three reef areas. For a particular area, good correlations among H- parameters of the full-band and high-frequency bands are evident in the absence of anthropogenic and geophonic sounds. At the same time, a poor correlation is indicated between the H parameters of the low-frequency fish band and full-band data. The similar trend is observed for three areas. This is possible when fish sounds dominate in comparison with the anthropogenic and geophonic sounds as we have observed in the present study location for ACI parameter.

In many investigations, fish sound identification and habitat characterizations are based on the algorithms suitable for the supervised system (Lin *et al.*, 2017; Chakraborty *et al.*, 2015). Misdetections can occur in noisy or unexpected situations while using a supervised system if the right architectures or trainings are not used, which is possible if data characteristics are not properly understood (De and Chakraborty, 2012). Various models were used to understand the involved processes through interaction among the acoustic parameters in the study locations. For example, a generalized linear mixed model (Baayen, 2008) was used for diversity studies in the three temperate ponds (Desjonqueres *et al.*, 2015). Nedelec *et al.* (2015) used PCA to passive acoustic data features acquired from the coral reef area, and the use of unsupervised clustering is proposed while investigating temporal-spectral patterns in data. Such unsupervised clustering approaches are K-means clustering (Liu and Deng, 2021) or Gaussian mixture models (GMMs) (Pearson *et al.*, 1992). These models can cluster samples into groups based on the dispersion of temporal-spectral characteristics without a pre-defined rule (Lin *et al.*, 2018). These unsupervised clustering approaches have been predominantly applied in the analysis of call repertoire also (Terhune *et al.*, 1993; Lin and Chou, 2015). Here, the clustering technique employed to identify four different fish choruses is found to be successful (Lin *et al.*, 2017; Benocci *et al.*, 2021).

V. CONCLUSION

Here, we initiated an investigation related to the soundscape of the shallow-water locations based on the recorded passive acoustic data from off coral reef area near Grande Island in the Zuari estuary. Passive acoustic recordings reveal prominent fish choruses, and such vocalization was present throughout the dusk to early dawn period. Analyses of the concatenated PSD data reveal two regimes based on PSDs' highest peak level and corresponding peak frequency, i.e., dominant fish chorus timing (14:00 - 02:45 h) having dominant sounds, and absence of such dominance within (03:00-13:45 h). We found an excess of (15.37 ± 02.15 dB re $1 \text{ Pa}^2/\text{Hz}$) of the mean peak level during the dominant fish choruses. The initial requirement is to identify the recorded data through auditory comparison between the call trains of the representative data with online available fish sounds from various sources. After that, we made an effort to match the frequency peaks of the recorded data. We identified four fish species/families: Sciaenidae (drums and croakers), *Terapon theraps* (Tiger Perch), planktivorous, and unnamed fish mentioned as Type A. Detailed characterization of the individual fish call employing segmentation of the oscillogram data provides individual call parameters of the four fish species/family types. Presented call parameters (call duration, number of pulses per call, inter-call duration, and frequency peak of the PSDs) suggest useful characteristics of four fish species/families.

We also investigated biodiversity parameters such as AEI, ACI, and SPLrms for three frequency bands known as full-band, low-frequency fish band, and high-frequency shrimp band. In general, we find ACI is the most consistent metric for three frequency bands. However, such consistency is also seen for AEI, especially for low as well as high-frequency bands. Unlike variability observed for the SPLrms parameters, ACI and AEI are consistent (except for AEI of full-band data) irrespective of the soundscape condition of the study location.

We employed an unsupervised learning technique to analyze soundscape components to investigate whether the machine learning based unsupervised techniques are useful to identify fishes using their sound comprising of biodiversity parameters. Hybrid technique involving PCA and K-means-based classification technique to twelve data features proves to be successful. Initially, we employ present unsupervised technique for the data acquired during the high ambient sounds that is when fish chorus was dominant (14:00 – 02:45 h). Here, we employed PCA for dimensionality reduction and subsequent K-means clustering provides 96.20%, 76.81%, 86.36%, and 100% for four fish types and 100% for boat sound, suggesting an overall classifier performance of 89.84%. We also applied present technique successfully for entire datasets. Overall classification is found to be 89.94% for identified six classes when entire dataset is considered. The present investigation has importance from real-time passive acoustic data acquisition systems when online passive acoustic data is

transmitted to shore or ship-based systems through radio or underwater communications. In that case, using unsupervised machine learning algorithms employed to multiple parameters can help monitor additions of fish sound types and modifications in biodiversity information through the clustering patterns developed here.

ACKNOWLEDGMENTS

We are thankful to Dr. Sunil Singh, Director, CSIR-National Institute of Oceanography, for his encouragement. KKC and BC acknowledges assistance from Jawaharlal Nehru Memorial Fund and CSIR Extra Mural Scheme, New Delhi respectively. We are thankful to Dr. G. Latha, ESSO-NIOT, Chennai for her support in many ways. Thanks to the International Quiet Ocean Experiment (IQOE) for opportunities they provided to network with bioacousticians from many nations who are participating in IQOE, and to the international sponsors of IQOE, the Scientific Committee on Oceanic Research (SCOR) and Partnership for Observation of the Global Ocean (POGO). We are indebted to both the reviewers and Associate Editor for their decent and detailed review. The quality of the manuscript has improved significantly during the revision. We express gratitude to Professor Luczkovic for his support through interactions with KKC. This is CSIR-NIO publication number xxxx.

REFERENCES

¹See <https://www.wildlifeacoustics.com> (Last viewed 12th Dec 2021)

²See <https://www.mathworks.com> (Last viewed 12th Dec 2021)

³See <https://www.niot.res.in/index.php/node/index/185/> (Last viewed 12th Dec 2021)

⁴See <https://audacity.sourceforge.net> (Last viewed 12th Dec 2021)

⁵See <https://www.fishbase.se/physiology/FishSoundsList.php?ID=1180>

(Last viewed 12th Dec 2021)

⁶See <https://fishsounds.net/fish.js?id=61398f98-2f42-4157-a129-799fa71f0d1b>

(Last viewed 12th Dec 2021)

⁷See <https://R-project.org> (Last viewed 12th Dec 2021)

⁸See <https://ljevillanueva.github.io/soundecology/> (Last viewed 12th Dec 2021)

⁹See <https://boycewire.com/what-is-the-gini-coefficient/> (Last viewed 12th Dec 2021)

¹⁰See <https://mulloverthing.com/zscore>(Last viewed 12th Dec 2021)

¹¹See <https://towardsdatascience.com/understanding-confusion-matrix-a9ad42dcfd62>

(Last viewed 12th Dec 2021)

¹²See <https://m.soundcloud.com/abc-science/sounds-of-science-terapons-spawning>

(Last viewed 12th Dec 2021)

¹³ See <https://dosits.org/galleries/audio-gallery/marine-invertebrates> (Last viewed 12th Dec 2021)

¹⁴See <https://theconversation.com/baby-oysters-follow-the-crackling-sound-of-snapping-shrimp-182514> (Last viewed 12th Dec 2021)

¹⁵See <https://ccari.icar.gov.in/Technical%20Bulletin%20No.%2049.pdf>

¹⁶See <https://www.fishbase.se/summary/4829> (Last viewed 12th Dec 2021)

¹⁷ See <https://astatsa.com/KruskalWallisTest/>

¹⁸ See <http://irgu.unigoa.ac.in/drs/handle/unigoa/6110> (Last viewed 12th Dec 2021)

¹⁹See <https://dosits.org/galleries/technology-gallery/basic-technology/vector-sensors/>

(Last viewed 12th Dec 2021)

²⁰See <https://www.fishbase.se/summary/Larimus-breviceps.html>

(Last viewed 12 Dec. 2021)

Ainslie M. A., McColm J. G. (1998) "A simplified formula for viscous and chemical absorption in sea water", *J. Acous. Soc. Am.*, 103: 1671-1672.

Akamatsu, T., Okumura, T., Novarini, N., and Yan, H. Y. (2002). "Empirical refinements applicable to the recording of fish sounds in small tanks," *J. Acoust. Soc. Am.* **112**, 3073-3082.

Amiri-Simkooei, A., Snellen, M., and Simons, D. (2011). "Principal Component Analysis of Single-Beam Echo-Sounder Signal Features for Seafloor Classification," *IEEE Jour. Ocean. Eng.* **36**, 259–272.

Ansari, Z. A., Chatterji, A., Ingole, B. S., Sreepada, R. A., Rivonkar, C. U., and Parulekar, A. H. (1995). "Community structure and seasonal variation of an inshore demersal fish community at Goa, West Coast of India," *Estuarine, Coastal and Shelf Science.* **41**, 593–610.

Au, W.W.L., Banks, K. (1998). "The acoustics of the snapping shrimp *Synalpheusparneomeris* in Kaneohe Bay," *J. Acoust. Soc. Am.* 103, 41-47

Baayen RH. 2008. *Analyzing linguistic data: a practical introduction to statistics using R*. Cambridge: Cambridge University Press.

Bassett C., Thomson J., Dahl P. H. and Polagye, B. (2014). "Flow-noise and turbulence in two tidal channels," *J Acoust Soc Am*, **135**, pp. 1764–1774.

- Benocci, R., Roman, H.E., Bisceglie, A., Angelini, F., Brambilla, G., Zambon, G. (2021). "Eco-Acoustic Assessment of an Urban Park by Statistical Analysis," *Sustainability*, **13**, 7857 also available at [https:// DOI.org/10.3390/su13147857](https://DOI.org/10.3390/su13147857)
- Bertucci, F., Parmentier, E., Lecellier, G., Hawkins, A. D., and Lecchini, D. (2016). "Acoustic indices provide information on the status of coral reefs: an example from Moorea Island in the South Pacific," *Sci. Rep.* **6**, 33326.
- Bianco, M.J., Gerstoft, P., Traer, J., Ozanich, E., Roch, M.A., Gannot, S., and Deledalle, C- A. (2019). "Machine learning in acoustics: theory and applications," *J. Acoust. Soc. Am.* **146**, 3590-3628.
- Bohnenstiehl, D. R., Lyon, R. P., Caretti, O. N., Ricci, S. W., Eggleston, D. B. (2018). "Investigating the utility of ecoacoustic metrics in marine soundscapes," *Journal of Ecoacoustics*, **2**, (2):1 DOI 10.22261/JEA.R1156L
- Borie, A., Rezende, S. M., Ferreira, B. P., Maida, M., Radford, C., and Travassos, P. (2021). "Soundscape of protected and unprotected tropical Atlantic coastal coral reefs," *Scientia Marina*, **85**, 5-14.
- Bock, H. H. (2007). "Clustering Methods: A History of k-Means Algorithms," In: Brito, P., Cucumel, G., Bertrand, P., de Carvalho, F. (eds) *Selected Contributions in Data Analysis and Classification. Studies in Classification, Data Analysis, and Knowledge Organization*. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-73560-1_15.
- Bradfer-Lawrence, T., Gardner, N., Bunnefeld, L., Bunnefeld, N., Willis, S. G., and Dent, D.H. (2019). Guidelines for the use of acoustic indices in environmental research. *Methods in Ecol.Evol.* **10**, 1796-1807.
- Chakraborty, B., Menezes, A., Dandapath, S., Fernandes, W., Karisiddaiah, S.M., Haris, K., and Gokul, G.S. (2015). "Application of hybrid techniques (self-organizing map and fuzzy algorithm) using backscatter data for segmentation and fine-scale roughness characterization of seepage related seafloor along the western continental margin of India," *IEEE J. Ocean. Eng.*, **40**, 3-14.
- Chanda, K., Chakraborty, B., Mahale, V. P., Latha, G., and Fernandes, W. (2020a). "Characterizing three shallow-water locations off Goa, India, using passive acoustic data," *J. Acoust. Soc. Am.* **148**, 1536–1551.
- Chanda, K., Chakraborty, B., Vardhan, Y. V., Gracias, D., Mahanty, M. M., Latha, G., Fernandes, W., and Chaubey, A. K. (2020b). "Passive acoustic metrics to understand shallow water biodiversity off Malvan area in the west coast of India," *Ind. J. Geo-Mar. Sci.* **49**, 527-536.
- Chary, L. K., Sreekanth, G. B., Deshmukh, M. K., and Sharma, N. (2020). "Marine soundscape and fish chorus in an archipelago ecosystem comprising bio-diverse tropical islands off Goa Coast, India," *Aquat. Ecol.* **54**, 475-493.
- Cheung, Y-M. (2003). "k-means: A new generalized k-means clustering algorithm," *Pattern Recognition Letters.* **24**, 2883–2893.
- Clarke, R.D., Finelli, C. M. and Buskey, E. J. (2009). "Water flow controls distribution and feeding behavior of twoco-occurring coral reef fishes: II. Laboratory experiments," *Coral Reefs*, **28**,475-488.

- De, C., Chakraborty, B. (2012). "Estimation of mean grain size of seafloor sediments using neural network," *Mar. Geophys. Res.* **33**, 45-53.
- Desjonqueres, C., Rybak, F., Depraetere, M., Gasc, A., Le Viol, I., Pavoine, S., and Sueur, J. (2015). "First description of underwater acoustic diversity in three temperate ponds," *PeerJ*, **3**: e1393; DOI 10.7717/peerj.1393.
- Dimoff, S. A., Halliday, W. D., Pine, M. K., Tietjen, K. L., Juanes, F. and Baum, J. A. (2021). "The utility of different acoustic indicators to describe biological sounds of a coral reef soundscape," *Ecological Indicators*, **124**:107435 DOI 10.1016/j.ecolind.2021.107435.
- Ding, C., and He, X. (2004). "K-nearest-neighbor consistency in data clustering: incorporating local information into global optimization," in *SAC'04: Proc. ACM Symp. on App. Comp.* available at <https://doi.org/10.1145/967900.968021> (Last viewed 02 Dec. 2021).
- Eldridge, A., Guyot, P., Moscoso, P., Johnston, A., Eyre-Walker, Y., and Peck, M. (2018). "Sounding out ecoacoustic metrics: Avian species richness is predicted by acoustic indices in temperate but not tropical habitats," *Ecological Indicators*. **95**, 939–952.
- Farina, A., and Gage, S. H. (2017). "Ecoacoustics: A New Science," in *Ecoacoustics: The Ecological Role of Sounds*, edited by A. Farina, and S. H. Gage (John Wiley and Sons, NJ), pp.1-11.
- Farina, A., Eldridge, A. & Li, P. (2021). "Ecoacoustics and Multispecies Semiosis: Naming, Semantics, Semiotic Characteristics, and Competencies," *Biosemiotics* **14**, 141–165. <https://doi.org/10.1007/s12304-021-09402-6>
- Fine, M. L. (2012). "Swim bladder sound production: the forced response versus the resonant bubble," *Bioacoustics*, **21**: 5-7.
- Fine M. L., and Parmentier E. (2015). "Mechanisms of fish sound production," in *Sound Communication in Fishes*, edited by F. Ladich (Springer, Wien), pp. 77-126.
- Fish, M. P., and Mowbray, W. H. (1970). *Sounds of Western North Atlantic Fishes: A Reference File of Underwater Biological Sounds* (The John Hopkins University Press, Baltimore, MD).
- Fish, M. P., Kelsey, A. S. J., Mowbray, W. H. (1952). "Studies on the Production of Underwater Sound by North Atlantic Coastal Fishes," *J. Mar. Res.* **11**: 180-193.
- Harris, S. A., Shears N. T., and Radford C. A. (2016). "Ecoacoustic indices as proxies for biodiversity on temperate reefs," *Methods in Ecology & Evolution*. **7**, 713-724.
- Hawkins, A. D., Hazelwood, R. A., Popper, A. N., and Macey, P. C. (2021). "Substrate vibrations and their potential effects upon fishes and invertebrates," *J. Acoust. Soc. Am.* **149**, 2782–2790.
- Hawkins, A. D. and Popper, A. N. (2018). "Directional hearing and sound source localization by fishes," *J. Acoust. Soc. Am.* **144**, 3329-3350.
- Hawkins, A. D., and Popper, A. N. (2017). "A sound approach to assessing the impact of underwater noise on marine fishes and invertebrates," *ICES J. Mar. Sci.* **74**, 635–651.

- Hegde, M.R., Padate, V.P., and Rivonker, C. U. (2016). "Seasonal variations in habitat selection and catch trends of Sciaenids (Family: Sciaenidae) from the tropical waters off Goa, west coast of India," *Ind. J. Geo-Mae. Sci.* **45**, pp. 943-951.
- Hoppe, E., and Roan, M. (2011). "Principal component analysis for emergent acoustic signal detection with supporting simulation results," *J. Acoust. Soc. Am.* **130**, 1962–1973.
- Hotelling, H. (1933). "Analysis of a complex of statistical variables into principal components," *J. Educ. Psychol.* **24**, 417–441. DOI:10.1037/h0071325.
- Hussain, A., and Ingole, B. (2020). "Massive coral bleaching in the patchy reef of Grande Island, along the eastern Arabian Sea during the 2015/16 global bleaching event," *Reg. Stud. Mar. Sci.* **39**, 101410.
- Jolliffe, I. T. (2002) "Principal component analysis, 2nd ed. New York, NY: Springer-Verlag
- Jolliffe, I. T., and Cadima, J. (2016). "Principal component analysis: a review and recent developments," *Phil. Trans. R. Soc. Series A.* **374**, 20150202.
- Kaiser, H. F. (1960). "The application of electronic computers to factor analysis," *Educational Psychological Measurement*, **20**, 141-151.
- Kaplan, M. B., Mooney, T. A., Partan, J., & Solow, A. R. (2015). "Coral reef species assemblages are associated with ambient soundscapes," *Mar. Ecol. Prog. Ser.* **533**, 93-107.
- Kasumyan, A.O. (2008). "Sounds and sound production in fishes," *J. Ichthyol.* **48**, 981–1030. <https://doi.org/10.1134/S0032945208110039>
- Knowlton, R.E., Moulton, J. M. (1963). "Sound Production in the Snapping Shrimps *Alpheus* (*Crangon*) and *Synalpheus*," *Biol. Bull.*, **125**, 311-331
- Labrosse, P., Kulbicki, M., and Ferraris, J. (2002). "Visual Fish Census Surveys: proper use and implementation," (Secretariat of the Pacific Community, Noumea Cedex, New Caledonia). available at <http://www.spc.int> (Last viewed 18 May 2022).
- La Manna G, Picciulin M, Crobu A, Perretti F, Ronchetti F, Manghi M, Ruiu A, Ceccherelli G. (2021). "Marine soundscape and fish biophony of a Mediterranean marine protected area," *PeerJ*, **9**: e12551 DOI 10.7717/peerj.12551.
- Ladich, F. and Fine, M. L. (2006). "Sound generating mechanisms in fishes: A unique diversity in vertebrates," in *Communication in Fishes*, edited by F. Ladich, S.P. Collin, P. Moller and B.G. Kapoor (Book 1, Science Publishers, Enfield, NH), pp. 3-43.
- Ladich, F. (2015). "Ontogenetic Development of Sound Communication in Fishes," in *Sound Communication in Fishes*, edited by F. Ladich (Springer-Verlag Wien), pp. 127-148.
- Lammers, M. O., Brainard, R. E., Au, W. W. L., Mooney, T. A., and Wong, K. B. (2008). "An ecological acoustic recorder (EAR) for long-term monitoring of biological and anthropogenic sounds on coral reefs and other marine habitats," *J. Acoust. Soc. Am.* **123**, 1720–1728.

- Lee, W.-J., and Staneva, V. (2020). "Compact representation of temporal processes in echosounder time series via matrix decomposition," *J. Acoust. Soc. Am.* **148**,3429-3442.
- Lin, T.-H., Tsao Y., and Akamatsu, T. (2018). "Comparison of passive acoustic soniferous fish monitoring with supervised and unsupervised approaches," *J. Acoust. Soc. Am.* **143**, EL278-EL284.
- Lin, T.-H., Fang, S.-H., & Tsao, Y. (2017). "Improving biodiversity assessment via unsupervised separation of biological sounds from long-duration recordings," *Sci. Rep.* **7**, 4547.
- Lin, T.H. and Chou, L.S. (2015). "Automatic classification of delphinids based on the representative frequencies of whistles," *J. Acoust. Soc. Am.* **138**,1003-1011.
- Liu, F., and Deng, Y. (2021). "Determine the number of unknown targets in open world based on elbow method," *IEEE Trans. Fuzzy Systems.* **29**, 986-994.
- Looby, A., Cox, K., Bravo, S., Rountree, R., Juanes, F., Reynolds, L.K., Martin, C. W. (2022). "A quantitative inventory of global soniferous fish diversity," *Rev Fish Biol Fisheries*, **32**, 581–595.
<https://doi.org/10.1007/s11160-022-09702-1>
- Luczkovicz, J. J., Mann, D. A., and Rountree, R. A. (2008), "Passive acoustics as a tool in fisheries science," *Trans. Am. Fish. Soc.* **137**, 533–541.
- Luczkovich J. J., and Sprague, M.W. (2022). "Soundscape Maps of Soniferous Fishes Observed from a Mobile Glider," *Front. Mar. Sci.*, **9**: 779540 DOI 10.3389/fmars.2022.779540
- Lumsdon, A. E., Artamonov, I., Bruno, M. C., Righetti, M., Tockner, K., Tonolla, D., and Zarfl, C. (2018). "Sound peaking—Hydro peaking induced changes in river soundscapes," *River Res. Appl.* **34**, 3-12.
- Mahanty, M., Latha, G., and Govindan, R. (2018). "Soundscapes in shallow water of the eastern Arabian Sea," *Prog. Ocean.* **165**, 158-167.
- Mahanty, M., Latha, G., and Harikrishnan, C. (2015). "*Terapon theraps* chorus observed in a shallow water environment in the southwestern Arabian," *Indian J. Geo-Mar. Sci.* **44**, 150-155.
- Malfante, M., Mars, J. I., Mura, M. D., and Gervaise, C. (2018). "Automatic fish sounds classification," *J. Acoust. Soc. Am.* **143**, 2834–2846 (2018).
- Manikandan, B., Ravindran, J., Mohan, H., Periasamy, R., Mani Murali, R., & Ingole, B. S. (2016). "Community structure and coral health status across the depth gradients of Grande Island, Central west coast of India," *Reg. Stud. Mar. Sci.* **7**, 150-158.
- Mann, D.A., and Lobel, P.S. (1997). "Propagation of damselfish (Pomacentridae) courtship sounds. *J. Acoust. Soc. Am.* **101**, 3783-3791.
- McCauley, R. D., and Cato, D. H. (2000). "Patterns of fish calling in a near shore environment in a Great Barrier Reef," *Phil Trans. R. Soc. Lond. B.* **355**, 1289–1293.

- McCauley, R. D. (2012). "Fish choruses from the Kimberley, seasonal and lunar links as determined by long term sea noise monitoring," in *Conference Proceedings of Acoustics 2012 Fremantle (Nov. 21-23), Fremantle, Western Australia*, edited by T. McMinn, available at http://www.acoustics.asn.au/conference_proceedings/AAS2012/index.htm (Last viewed 02 Dec. 2021).
- McWilliam, J. N., and Hawkins, A. D. (2013). "A comparison of inshore marine soundscapes," *J. Exp. Mar. Biol. Ecol.* **446**, 166–176.
- Mooney, A., Di Iorio, L., Lammers, M., Lin, T.-H., Nedelec, S., Parsons, M., Radford, C., Urban, E., and Stanley, J. (2020). "Listening forward: approaching marine biodiversity assessments using acoustic methods," *Royal Society Open Science.* **7**, 201287.
- National Research Council (2003). *Ocean Noise and Marine Mammals*. Washington D.C.: National Academic Press.<http://nap.nationalacademies.org/10564>
- Nedelec, S. L., Simpson, S. D., Holderied, M., Radford, C. A., Lecellier, G., Radford, A. N., and Lecchini, D. (2015). "Soundscapes and living communities in coral reefs: Temporal and spatial variation," *Mar. Ecol. Prog. Ser.* **524**, 125–135.
- Ozanich, E. R., Thode, A.M., Gerstoft, P., Freeman, S., and Freeman, L.A. (2018). "Unsupervised clustering of coral reef fish calls," *J. Acoust. Soc. Am.* **148**, 2729 (Meeting Abstract), available at <https://doi.org/10.1121/1.5147579> (Last viewed 02 Dec. 2021).
- Padate, V. P. (2010). *Biodiversity of demersal fish along the estuarine - shelf regions of Goa*. PhD thesis, Goa University.<https://shodhganga.inflibnet.ac.in/simple-search?query=padate>(Last viewed 02 Dec. 2021).
- Parsons, M. J. G., Salgado Kent, C. P., Recalde-Salas, A., and McCauley, R. D. (2017). "Fish choruses off Port Hedland, Western Australia," *Bioacoustics.* **26**, 135–152.
- Pearson, J. D., Morrell, C. H., Brant, L. J. (1992). "Mixture models for investigating complex distributions," *J. Quan. Anthropol.*, **3**, 325-345.
- Peck, M., Tapilatu, R.F., Kurniati, E. and Rosado, C. (2021), "Rapid coral reef assessment using 3D modelling and acoustics: acoustic indices correlate to fish abundance, diversity and environmental indicators in West Papua, Indonesia," *PeerJ*. **9**: e10761; DOI: 10.7717/peerj.10761.
- Pieretti, N., Lo Martire, M., Farina, A., and Danovaro, R. (2017). "Marine soundscape as an additional biodiversity monitoring tool: A case study from the Adriatic Sea (Mediterranean Sea)," *Ecological Indicators.* **83**, 13–20.
- Pieretti, N, Farina, A. (2013). "Application of a recently introduced index for acoustic complexity to an avian soundscape with traffic noise," *J. Acoust. Soc. Am.* **134**, 891-900.
- Pijanowski, B. C., Farina, A., Gage, S. H., Dumyahn, S. L., and Krause, B. L. (2011a). "What is soundscape ecology? An introduction and overview of an emerging new science," *Landscape Ecology.* **26**, 1213–1232.

- Pijanowski B. C., Villanueva-Rivera L. J, Dumyahn S. L., Farina, A., Krause, B., Napolitano, B. M., Gage, S. H., Pieretti, N. (2011b). "Soundscape ecology: the science of sound in the landscape," *BioScience* **61**: 203–216.
- Pijanowski, B. C., Rodríguez-Buriticá, S., Ulloa, J. S. (2021). "Tooting the Latin American horn: advances in the scholarship of ecoacoustics and soundscape ecology is occurring with vigor," *Biota Colomb.*, **22**: 2-6.
- Popper, A. N., and Hawkins, A. D. (2019). "An overview of fish bioacoustics and the impacts of anthropogenic sounds on fishes," *J. Fish Biol.* **94**, 1–22.
- Popper, A. N., Hawkins, A. D., Sand, O., Sisneros, J.A. (2019). "Examining the hearing abilities of the fishes," *J. Acoust. Soc. Am.* **146**, 948–955.
- Popper, A. N., and Hawkins, A. D. (2018). "The importance of particle motion to fishes and invertebrates," *J. Acoust. Soc. Am.* **143**, 470-488..
- Putland, R. L., Montgomery, J. C., and Radford, C. A. (2019). "Ecology of fish hearing," *J Fish Biol.* **95**, 39–52.
- Radford, C. A., Jeffs, A. G., Tindle, C. T., Montgomery, J. C. (2008) "Temporal patterns in ambient noise of biological origin from a shallow water temperate reef," *Oecologia*, **156**: 921–929. DOI 10.1007/s00442-008-1041-y.
- Ramcharitar, J., Gannon, D., and Popper, A. (2006). "Bioacoustics of fishes of the family Sciaenidae (Croakers and Drums)," *Trans. Amer. Fish. Soc.* **135**, 1409–1431.
- Roca, I. T., Van Opzeeland, I. (2020). "Using acoustic metrics to characterize underwater acoustic biodiversity in the Southern Ocean," *Remote Sensing in Ecology and Conservation* **6**: 262-273. doi:10.1002/rse2.129
- Rountree, R.A., Gilmore, R.G., Goudey, C. A., Hawkins, A.D., Luczkovich, J., Mann, D.A. (2006). "Listening to Fish: Application of passive acoustics to Fisheries Science," *Fisheries Magazine*, **31**, 433-446. available at [https://doi.org/10.1577/1548-8446\(2006\)31\[433:LTF\]2.0.CO;2](https://doi.org/10.1577/1548-8446(2006)31[433:LTF]2.0.CO;2).
- Schneider, H. (1964). "Physiologische und morphologische Untersuchungen zur Biokustik der Tigerfische (Pisces, Theraponidae)," *Z. Vergl. Physiol.*, **47**, 493-588.
- Schober, P., Boer, C., Schwarte, L. A. (2018) "Correlation Coefficients: Appropriate Use and Interpretation," *Anesthesia and Analgesia, Anesthesia & Analgesia*, **126**: 1763-1768. DOI: 10.1213/ANE.0000000000002864.
- Sreekanth, G. B., Chakraborty, S. K., Jaiswar, A. K., and Zacharia, P. U. (2018). "An inventory on the coastal finfish and shellfish species of Zuari estuary southwest coast of India," *Ind. J. Geo-Mar. Sci.* **47**, 945-958.
- Sreekanth, G. B., Manju Lekshmi, M., and Singh N. K. (2017). "Temporal patterns in fish community structure: environmental perturbations from a well-mixed tropical estuary," *Proc. Natl. Acad. Sci., India, Sect. B Biol. Sci.* **87**, 135–145.
- Sreekanth, G. B., Jaiswar, A.K., Zacharia, P. U., Pazhayamadam, D.G., and Chakraborty, S.K. (2019). "Effect of environment on spatio-temporal structuring of fish assemblages in a monsoon-

influenced tropical estuary," *Environ Monit. Assess.* **191**, 305 available at doi.org/10.1007/s10661-019-7436-x(Last viewed 02 Dec. 2021).

- Staaterman, E., Rice, A. N., Mann, D.A., and Paris, C.B. (2013). "Soundscapes from a Tropical Eastern Pacific reef and a Caribbean Sea reef," *Coral Reef*. **32**, 553-557.
- Sueur, J., Farina, A., Gasc, A., Pieretti, N., and Pavoine, S. (2014). "Acoustic indices for biodiversity assessment and landscape investigation," *Acta Acustica United with Acustica*. **100**, 772-781.
- Sueur J. (2018). "Indices for Ecoacoustics," in: *Sound Analysis and Synthesis with R*, (Springer, Cham), 479-518. available at https://doi.org/10.1007/978-3-319-77647-7_16(Last viewed 02 Dec. 2021).
- Sueur, J. and Farina, A. (2015). "Ecoacoustics: The Ecological Investigation and Interpretation of Environmental Sound," *Biosemiotics*, **8**, 493-502.
- Terhune, J. M., Burton, H., and Green, K. (1993). "Classification of diverse call types using cluster analysis techniques," *Bioacoustics* **4**, 245–258.
- Towsey, M., Planitz, B., Nantes, A., Wimmer, J., and Roe, P. (2012). "A toolbox for animal call recognition," *Bioacoustics*. **21**, 107- 125.
- Tricas, T. S., and Boyle, K.S. (2014). "Acoustic behaviors in Hawaiian coral reef fish communities," *Mar. Ecol. Prog. Ser.* **511**, 1–16.
- Urquhart, D., Hawkins, A. D., and Hall, C. (2006). "Monitoring of pile driving at Aberdeen Harbour 2005—A report prepared for the River Dee Trust and Aberdeen Harbour Board," (Aberdeen Harbour Board, Aberdeen).
- Velip, D., and Rivonker, C. (2018). "Population structure and reproductive biology of selected sciaenid species along the fishing grounds of Goa, west coast of India," *Curr. Sci.* **115**, 146–152.
- Villanueva-Rivera, L. J., Pijanowski, B. C., Doucette, J., and Pekin, B. (2011). "A primer of acoustic analysis for landscape ecologists," *Landscape Ecology*. **26**, 1233-1246.
- Wenz, G.M. (1962). "Acoustic ambient noise in the ocean: Spectra and sources," *Journal of the Acous. Soc. of Am.* **34**:1936-1956.
- Williams, B. et al. (2022). "Enhancing automated analysis of marine soundscapes using ecoacoustic indices and machine learning," *Ecological Indicators* **140**. <https://doi.org/10.1016/j.ecolind.2022.108986>