DCLDE Rotterdam 2024

ABSTRACTS

Commando Materieel en IT
Ministerie van Defensie

Ministerie van Landbouw, Natuur en Voedselkwaliteit

living marine RESOURCES — SCIENCE • STEWARDSHIP • NAVY READINESS

TNO

Office of Naval Research — ONR — Science & Technology

United States Fleet Forces Command

Rijkswaterstaat
Ministerie van Infrastructuur en Waterstaat

SOUND & MARINE LIFE PROGRAMME

BOEM BUREAU OF OCEAN ENERGY MANAGEMENT

Fisheries and Oceans Canada — Pêches et Océans Canada

Abstracts for Full Presentations

Strengths and weaknesses of using DCLDE algorithms to track baleen whales and examine their behavior using long-term acoustic recordings with a large-scale hydrophone array

G.C. Alongi[1*], T.A. Helble[2], R.A. Guazzo[2], S.W. Martin[1], C.R. Martin[2], E.E. Henderson[2]

[1]*National Marine Mammal Foundation, San Diego, CA, USA*

[2]*Navy Information Warfare Center, San Diego, CA, USA*

**gabriela.alongi@nmmpfoundation.org*

The use of detection, classification, localization, and tracking algorithms on acoustic data recorded at the U.S. Navy's Pacific Missile Range Facility (PMRF) has resulted in a wealth of research on animal behavior. PMRF allows for persistent monitoring and currently uses 63 time-synchronized hydrophones covering a 20 km x 60 km area. Collectively, thousands of baleen whale tracks were produced from their recorded vocalizations over the last decade, including humpback whales, minke whales, Bryde's whales, and fin whales. Track prediction methods and Hidden Markov Models were applied using open-source R packages to characterize whale swimming and calling behavior on the range. Whale movement was examined relative to variables including calendar year, day of year, hour of day, wind speed, wave height, acoustic calling rates, and the presence of Navy sonar. Findings indicated that whales changed their swimming behavior based on a variety of environmental and temporal variables as well as acoustic behavioral state and exposure to mid-frequency active sonar. Even with this near best-case scenario for long-term and broad spatial data collection along with whale localization and tracking capabilities, several limitations have become apparent during these analyses. These include complications with quantifying cue rates, environmental-, behavioral-, and density-dependent probability of tracking, and data deficiencies for tracks under specific conditions, other potentially relevant environment variables, and uncommon events such as storms. We will review the strengths and weaknesses of this system within the broader context of the DCLDE community and consider the implications and what might be transferrable to other research efforts with different acoustic sampling methods.

Diving Deep: 3D Tracking of Cuvier's Beaked Whale Diving Behavior in Southern California using Fixed Hydrophone Arrays

L.M. Baggett[1*], E.R. Snyder[1], A. Solsona-Berga[1], K.E. Frasier[1], J.A. Hildebrand[1], A. Širović[2], S. Baumann-Pickering[1]

[1]*Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA, USA*

[2]*Norwegian University of Science and Technology, Trondheim, Norway*

**lbaggett@ucsd.edu*

Cuvier's beaked whales (Ziphius cavirostris) are a deep-diving cetacean species known to forage in the submarine canyons of the Southern California Bight. Although this species is a top predator in deep-sea ecosystems, much is yet to be learned about their social and foraging strategies due to their elusive behavior and extensive time spent foraging in the deep sea. In this analysis, time-difference-of-arrival (TDOA) localization is used to track the position of Cuvier's beaked whales from echolocation clicks recorded on seafloor-mounted hydrophone arrays. This approach yielded over 2000 of tracks of diving Cuvier's beaked whales from several years of acoustic data recordings collected at four long-term acoustic monitoring sites between 2018 and 2023. These tracks were used to estimate swim speed, track distance, and area covered during deep foraging dives, as well as distance from the seafloor, group size, and distance between simultaneously diving individuals. These metrics revealed distinct diving behaviors and spatial use trends driven by bathymetric features specific to each of the recording sites, notably a preference for canyon walls and deeper canyon pockets. Group size during observed simultaneous diving ranged from 2 to 9 whales with most encounters between 2 and 4 whales, although there was diel, seasonal, and spatial variability in this feature. During several of these encounters, individuals exhibited behaviors that appeared highly coordinated. This long-term acoustic monitoring effort and TDOA localization approach yielded an extensive tracking dataset which gives valuable insight into the behavior at depth of Cuvier's beaked whales in the Southern California Bight.

Fine-tuning acoustic classifiers for false killer whale populations: Insights from sensitivity analyses

Y.M. Barkley[1*], J.L.K. McCullough[2], S. Fregosi[3], E.M. Oleson[4]

[1]*Cooperative Institute for Marine and Atmospheric Research, Honolulu, Hawaiʻi, USA*

[2]*NOAA Pacific Islands Fisheries Science Center, Honolulu, Hawaiʻi, USA*

[3]*Ocean Associates, Inc., Arlington, Virginia, USA*

[4]*NOAA Pacific Islands Fisheries Science Center*

**yvonne.barkley@noaa.gov*

Obtaining accurate species identifications is essential for marine mammal conservation and management. For Hawaiian false killer whales, three genetically distinct populations occur throughout the Hawaiian Archipelago: a pelagic population broadly dispersed offshore, an insular population associated with the Northwestern Hawaiian Islands (NWHI), and an endangered insular population associated with the main Hawaiian Islands. Given the anthropogenic threats to this species and the endangered status of one population, it is imperative that false killer whales be identified to the population level - a challenging task that involves approaching animals closely enough to collect tissue samples or photo-identification data. False killer whales also emit different acoustic signals, which allows for the development of acoustic classification models using features of their whistles, clicks, and burst pulses. We applied detectors from PAMGuard to automatically detect and extract false killer whale calls from towed array recordings collected between 2010 and 2021, including the DCLDE 2022 data set. Acoustic features from all call types were included in BANTER, a supervised hierarchical classification model, to classify acoustic events to a population. Classification models produced a low out-of-bag correct classification score of 45.7%, with the NWHI population performing best. Given that each population contained several varying factors (e.g., number of subgroups, number of calls, total event duration), we performed sensitivity analyses to investigate which factors influenced classification results to potentially fine-tune the models for higher accuracy. Preliminary analyses showed unexpected results where a lower number of calls produced higher correct classification scores than the maximum number of calls available. Further investigation of different variables will reveal whether this classifier using automatically extracted features can confidently classify false killer whale events to a population.

Detection and classification of marine mammal sounds over a wide band of frequencies

M. Baumgartner[1*]

[1]*Woods Hole Oceanographic Institution*

**mbaumgartner@whoi.edu*

The development of passive acoustic detection and classification methods for marine mammals has traditionally focused on one species at a time, resulting in a myriad of algorithms that each operate in a narrow frequency band and must be applied one at a time to detect several species of interest in a single recording or real-time audio stream. When an application requires the detection of many species simultaneously, this approach is highly inefficient, and for very large recordings or for real-time processing on embedded microprocessors, it is impractical. Some detection and classification systems seek to detect a class of calls, such as tonal or pulse calls. The low-frequency detection and classification system (LFDCS) is one such system that uses (1) pitch tracking to detect and characterize all tonal sounds in a recording or real-time audio stream, and (2) a separate discriminant function analysis to classify sounds by call type and species. This allows the detection and classification of many species simultaneously in a single processing run. The LFDCS, like most tonal marine mammal detectors, relies on spectrograms produced with the short-time Fourier transform (STFT). These spectrograms characterize spectral content on a linear frequency scale, even though sound is both produced and perceived on a logarithmic frequency scale. This limits the band of frequencies that can be effectively monitored to roughly 4 octaves, yet marine mammals make tonal and pulse sounds over a range of 12+ octaves. I have developed a detection and classification system (preliminarily called the wide band detection and classification system, or WBDCS) that is conceptually similar to the LFDCS, but detection and pitch tracking is conducted using a spectrogram produced with the constant-Q transform (CQT), which allows the estimation of spectral content on a logarithmic frequency scale. The WBDCS allows the detection and classification of tonal sounds between 15 Hz and 60 kHz, and has been implemented on the digital acoustic monitoring (DMON) instrument for real-time monitoring from autonomous platforms, such as moored buoys and buoyancy-driven gliders and profiling floats. Future plans include incorporation of a time-domain detection and classification system for pulse sounds over a wide band of frequencies using the pre-processing steps involved in creating the CQT spectrogram for tonal detection in the WBDCS. My long-term goal is to create a single system capable of simultaneously detecting tonal and pulse calls ranging from the low-frequency moans of blue whales to the high-frequency clicks of beaked whales.

Listening to whales from the sky: How Royal Canadian Airforce sonobuoy data contributes to detection, classification, localization, and density estimation research

C.M. Binder[1*], D.J. Thomson[1], K. Thebeau[1], Z. Wallot-Beale[1], J.T MacDonnell[1]

[1]*Defence Research and Development Canada, Dartmouth, NS, Canada*

**carolyn.binder@forces.gc.ca*

The Royal Canadian Airforce (RCAF) has recently been undertaking a review and update of policies and procedures designed to limit the potential impact of active sonar on whales. During active sonar activities, passive sonar sonobuoys can be used to acoustically monitor for the presence of whales so that sonar operations can be shut down if an animal approaches an active sonar source. Defence Research and Development Canada (DRDC) is assisting the RCAF to develop methods to efficiently detect and localize marine mammal calls in sonobuoy data. This talk will provide a description of how the RCAF collects acoustic data using sonobuoys, and the benefits and limitations of using sonobuoy data for detection, classification, localization, and density estimation (DCLDE) research. There will be a particular focus on the 2018 North Atlantic right whale acoustic data collection effort conducted in the Gulf of St. Lawrence (i.e., the workshop data set). A selection of detection and localization results from RCAF-collected sonobuoy data will be presented as case studies for how sonobuoy data has contributed to DRDC's DCLDE research.

Density estimation of Indo-Pacific finless porpoises using passive acoustic monitoring off the Sindhudurg coast of India

I. Bopardikar[1,2*], D. Harris[3], T. Yack[4,5], V.V. Robin[1], H. Klinck[2]

[1]*Indian Institute of Science, Education and Research, Tirupati, India*

[2]*K. Lisa Yang Center for Conservation Bioacoustics, Cornell University, USA*

[3]*Centre for Research into Ecological and Environmental Modelling University of St Andrews, Scotland*

[4]*Duke University, USA*

[5]*EcoSound Bioacoustics, LLC, USA*

**isha.bopardikar@gmail.com*

Coastal cetaceans in India, several listed as vulnerable or threatened by the IUCN, are exposed to increasing anthropogenic pressures related to unregulated and overexploitative fishing. The Indo-Pacific finless porpoise (Neophocaena phocaenoides) is one such species in need of conservation action. A major cause of concern is direct and indirect interactions with fishing gear, as this species is one of the most common cetacean species encountered in by-catch. However, data on abundance, distribution and trends of finless porpoises are severely lacking for Indian waters, hindering the formulation of effective long-term conservation and management measures. Gathering visual abundance estimates of finless porpoises is hampered by their cryptic behaviour, elusive surfacing patterns, and small group sizes. However, this species regularly produces narrowband high-frequency echolocation clicks for communication, navigation, and prey detection, making passive acoustic monitoring (PAM) an efficient tool for population-level monitoring. Our study utilised an autonomous towed array to detect, classify, localise and subsequently produce density estimates for finless porpoises along a 376 km2 area off the Sindhudurg coast of India.

Following a distance sampling protocol, we conducted 38 acoustic line transect surveys between January 2020 to March 2023. A linear four-channel array was towed behind the survey vessel at a depth of ~5 meters. The array consisted of four HTI-96 min hydrophones connected to a SoundTrap ST4300. All channels were sampled simultaneously and continuously at a 288 kHz sampling rate and 16-bit resolution. This setup allowed localisation of individual vocalising animals using time difference of arrival and target motion analysis.

A total of 114 hours of acoustic data were post-processed in PamGuard using the click detector and classifier modules to detect, classify, and track porpoise click trains. Click trains were localised to estimate distances measured perpendicularly to the trackline to vocalising animals. These distances were then used to estimate a detection function.

A hazard-rate detection function was selected to calculate a detection probability of 0.6 (CV = 13%). A preliminary analysis of data between 2020 to 2022 estimated a porpoise density of 1.4 individuals/km2 (95% CI: 0.9 - 2.2). Density and abundance estimates from 2020 to 2023 using the calculated detection function will be presented and discussed further.

Our study provides the first acoustic-based population estimates for finless porpoises along the Sindhudurg coast of India. Further, we highlight the need for a similar assessment of finless porpoise populations across their entire distribution in Indian waters using PAM.

First glimpses into the frequency response of Distributed Acoustic Sensing to blue and fin whale calls as a function of gauge length

L. Bouffaut[1*], H. Klinck[1]

[1]*K. Lisa Yang Center for Conservation Bioacoustics, Cornell Lab of Ornithology, Cornell University*

**lea.bouffaut@cornell.edu*

Distributed Acoustic Sensing (DAS) converts fibers in existing underwater telecommunication cables into dense listening arrays of strain sensors by connecting an interrogator to the cable on shore. Previous studies have shown that DAS can be used to detect waterborne acoustic sources, including the low-frequency sounds of blue and fin whales (Balaenoptera musculus and B. physalus) over tens of km of fiber. DAS is instrumented from land, collects data in near-real time, and can offer region-wide spatial coverage. Thus, this technology may provide a unique opportunity for robust monitoring of endangered and threatened migratory baleen whales at ecologically meaningful spatial and temporal scales.

Conceptually, the DAS interrogator uses the Rayleigh backscatter of laser sweeps (typically, centered around 1,550 nm) sent through the fiber to interpret the displacement of randomly distributed defects present along the fiber under an incoming pressure wave. Then, it calculates the average time-differentiated phase change over short fiber sections —the gauge length— at regularly spaced intervals along the fiber, creating virtual channels. Finally, the phase shift is converted into longitudinal strain waveforms for each corresponding fiber section. The resulting strain data, sampled in both time and space, is accessible in near-real-time.

The strain response to waterborne pressure waves for a specific fiber/interrogator combination is a function of the user-defined gauge length, signal frequency, wave velocity, and the angle of arrival of the incoming wave. Because the phase averaging over the gauge length introduces a frequency- and bearing-dependent channel response, the gauge length is the most critical parameter for DAS data acquisition. In seismology, it is usually set on the order of ten meters. Averaging over longer fiber sections leads to increased signal-to-noise ratios (SNR) but limits the highest solvable frequency, therefore limiting the bandwidth of the collected data. After a brief introduction to DAS, this presentation will summarize our findings on the impact of the gauge length on the frequency response of different DAS configurations evaluated using low-frequency whale calls. It will be followed by a roadmap of the research and tests needed for the technology to reach its full potential.

Learning from birds to find whales: Efficacy of transfer learning for detection and classification of North Atlantic right whale upcalls

R.E. Cohen[1*], S. Kahl[1,2], H. Klinck[1], A.N. Rice[1]

[1]*K. Lisa Yang Center for Conservation Bioacoustics, Cornell Lab of Ornithology, Cornell University, Ithaca, NY, USA*

[2]*Chemnitz University of Technology, Chemnitz, Germany*

**r.cohen@cornell.edu*

Deep learning models have become standard tools in the bioacoustics detection/classification workflow but achieving good performance on specialized tasks with limited datasets remains challenging. A primary consideration when training a model from scratch is the need for sufficient labeled training examples, which in bioacoustics applications are often scarce and highly time- and labor- intensive to construct. One approach to mitigating the data-hungry nature of these models is to employ transfer learning, in which a model learns effective feature representations (embeddings) from a generic labeled data set, and then the final few model layers are fine-tuned with a small training set customized for the particular discrimination task at hand. This allows the utilization of very large, publicly available datasets for training, even if they do not contain the particular classes of interest to the end user. However, the efficacy of the embeddings learned from the generic training data, and therefore the performance of the model on a new custom discrimination task, will vary based on the extent of differences between the characteristics of the original training classes and those of the custom target classes (domain shift). We explored the utility of the transfer learning approach for bioacoustics applications in the case of a substantial domain shift: can a highly parameterized model trained on an extensive bird call dataset be leveraged for detecting low-frequency North Atlantic right whale calls? The large embeddings model (BirdNET V2.4) was trained on data drawn from Xeno-canto, a publicly accessible archive of primarily avian vocalizations which has been widely used for training deep learning classification models. We customized the final few layers of the embeddings model for the desired low-frequency classification task using a modestly sized NARW upcall training set. This transfer learning technique was compared with two alternative classifiers: an embeddings model completely retrained with the inclusion of several low frequency classes, and a small custom model trained specifically on NARW calls. Here we compare the performance of these models and appraise their respective advantages and shortcomings.

Classification of dolphin whistles from the Adriatic Sea

A. Constaratas[1*], O. Özgöbek[2], D. Holcer[3,4], A. Širović[1]

[1]*Department of Biology, Norwegian University of Science and Technology, Trondheim, Norway*

[2]*Department of Computer Science, Norwegian University of Science and Technology, Trondheim, Norway*

[3]*Croatian Natural History Museum, Department of Zoology, Zagreb, Croatia*

[4]*Blue World Institute of Marine Research and Conservation, Veli Lošinj, Croatia*

**alexandra.constaratas@ntnu.no*

Bottlenose dolphins (Tursiops truncatus), striped dolphins (Stenella coeruleoalba), and Risso's dolphins (Grampus griseus) are delphinid species known to occur in the Adriatic Sea, while common dolphins (Delphinus delphis) are considered regionally extinct but occur occasionally. The acoustic repertoire of all these species consists of echolocation clicks and whistles. To get a better understanding of the presence of delphinids in this region, robust detection and classification approaches are needed for efficient detection of their signals. We developed a training dataset for classification of delphinid whistles from this region based on, as much as practicable, autonomous recordings from the Adriatic Sea areas known to be habitat for just a single species, as well as from hydrophone recordings during encounters with single-species groups in 2021 and 2023. Presence of whistles was annotated using Raven Pro during encounters spaced at least 15 minutes apart, with varying signal-to-noise ratios to capture different versions of potential patterns. Bottlenose dolphin whistles were annotated in the autonomous datasets from two locations: Lošinj (579 whistles during 115 encounters) and Blitvenica (579 whistles during 110 encounters), where their habitats have been reported. Striped dolphin whistles were annotated from recordings collected during encounters with single-species group offshore Vis in 2023 (32 whistles during 3 encounters). To augment data from the Adriatic Sea, striped dolphin annotations were also created from encounter in the Northwest Atlantic in 2021 (146 whistles during 3 encounters, part of the Atlantic Marine Assessment Program for Protected Species). Likewise, common and Risso's dolphin whistles were annotated during encounters with single-species group from the Watkins Marine Mammal Sound Database: 129 whistles during 52 encounters from the Atlantic collected between 1958 and 1987 for the common dolphins and 88 whistles during 31 encounters from the Atlantic and the Ionian Sea collected between 1959 and 1987 for the Risso's dolphin. We used with an equal number of whistles from each species for balanced performances, and 80% of the annotated dataset was used to train and 20% of the dataset to test the performance of two classification algorithms: ResNet, a Convolutional Neural Network (CNN) algorithm and a Classification and Regression Trees (CART). Results of the comparison across these two approaches will be discussed. Ultimately, the whistle classification will be used in conjunction with echolocation

click detection to determine the occurrence of these four species of delphinids on recordings collected offshore Dubrovnik between 2018-2020.

Estimating Rice's Whale (Balaenoptera ricei) Call Density in the Northeastern Gulf of Mexico Using Spatially Explicit Capture-Recapture Analysis

A. Cook[1*], E. Babcock[1], A. Debich[2,3], H. Frouin-Mouy[3], L. Tenorio-Hallé[3], M. Soldevilla[3]

[1]*University of Miami, Miami, Florida, USA*

[2]*Cooperative Institute for Marine and Atmospheric Studies, University of Miami, Miami, Florida, USA*

[3]*National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Southeast Fisheries Science Center, Miami, Florida, USA.*

**atc82@miami.edu*

Rice's whales (Balaenoptera ricei) are an endangered baleen whale species inhabiting the Gulf of Mexico, with a population of fewer than 100 individuals. Estimating the density of whale populations is essential for effective conservation and management as it provides valuable information for monitoring population trends, understanding habitat use, and mitigating threats. Spatially explicit capture-recapture (SECR) is a developing method applied in passive acoustic monitoring (PAM) of cetaceans, allowing estimation of animal abundance and density by integrating mark-recapture and distance sampling techniques with passive acoustic recording methods.

In our study, we deployed an array of 18 bottom-mounted hydrophones, spanning a 40 by 360 km area along the upper slope of the West Florida Shelf, adjacent to the De Soto Canyon, during a twelve-month period from May 2021 to April 2022. We quantified cue density from detected acoustic cues of Rice's whales using a deep-learning-based call detector, accounting for false positives and probability of detection. False positive rates were determined through manual validation, while the probability of detection was calculated by employing SECR methods on a subset of data in which we identified the hydrophones that detected each call, yielding call capture history vectors.

We present monthly estimates of cue densities in long-moan calls per hour per km2. While converting long-moan cue densities to animal densities currently remains challenging due to limited knowledge of Rice's whale population call rates, this aspect may be explored in future studies as call rate data become available. Our study provides valuable insights into relative trends in seasonal density and distribution, which is necessary for assessing the Rice's whale population within their core habitat. These findings contribute significantly to the conservation and protection efforts aimed at preserving this endangered species.

Use of machine learning and dynamic time warping to categorise large datasets of bottlenose dolphin whistles

C.R. Day[1*], J.N. Oswald[1], L.S. Sayigh[2], V.M. Janik[1]

[1]*Scottish Oceans Institute, School of Biology, University of St Andrews, St Andrews, Fife KY16 8LB UK*

[2]*Biology Department, Woods Hole Oceanographic Institution, Falmouth, MA 02543, USA*

**crdd@st-andrews.ac.uk*

Comparisons of whistle repertoires across species and populations may improve our ability to identify them in acoustic recordings. It is important to examine full repertoires, as some whistle types may carry more identifying information than others. Additionally, recognising biologically relevant whistle categories such as signature whistles can be a useful tool for density estimation. ARTwarp uses an unsupervised neural network and dynamic time warping to categorise tonal signals based on frequency content and contour shape (Deecke VB, Janik VM (2023) https://doi.org/10.5281/zenodo.7713615). This is achieved by comparing signals to be categorised to a set of reference contours and placing them into the categories to which they are most similar. However, there is uncertainty regarding the level of similarity (the Vigilance Parameter, VP) required to achieve biologically relevant categories. Empirically derived VPs of 91% and 96% have been reported in previous studies, however, these were based on small sample sizes. This study assessed ARTwarp's ability to create biologically relevant categories using the Sarasota Dolphin Whistle Database of bottlenose dolphin (Tursiops truncatus) signature whistles. We used 1800 whistles from 60 signature whistle types evenly split into three shape types: upsweep, arch and oscillatory to assess the impact of contour shape on categorisation success. The optimal VP was identified for each signature whistle type by running ARTwarp using 56 different VPs to find the lowest VP before the whistle type was split into two categories. Optimal VPs were 96%, 94% and 94.5% for upsweep, arch and oscillatory whistle types respectively, and the overall optimal VP was 95%. Categorisation was then performed using VP=95% on the pooled 1800 whistles and ARTwarp created signature whistle categories with a Normalised Mutual Information (NMI, a measure of categorisation success) value of 0.88. ARTwarp was not as successful as categorisation by human observers (NMI=0.9-0.99, Kershenbaum et al. (2013) PLoS One 8, e77671). However, it was more successful than all algorithms tested in Kershenbaum et al., including using a correlation matrix, a similarity matrix of time-warped whistles, and a Parson's code matrix, clustered using k-means and hierarchical clustering and another adaptive resonance theory algorithm (NMI=0.52-0.79). These findings show that ARTwarp is a consistent and quantifiable method for categorizing whistle types and determining whistle repertoires. Knowledge of the optimal VP for bottlenose dolphin whistles will allow more accurate quantification and description of their whistle repertoires and facilitate the use of signature whistles in methods for acoustic density estimation.

Robustness assessment of a dolphin whistle detector in the Red Sea and on the DCLDE11 dataset

R. Diamant[1*], A. Testolin[2]

[1]*University of Haifa & University of Zagreb*

[2]*University of Padova*

**roee.d@univ.haifa.ac.il*

The monitoring of dolphin populations is of interest for marine ecological studies. Their distribution in almost all tropical and temperate regions and in both coastal and offshore waters makes them ideal representatives, as they exhibit high plasticity in their habitat choice and can migrate to a different environment in case of severe disturbance. A key technology for monitoring dolphin populations is the detection of dolphin whistles. By analyzing opportunistic dolphin whistles, we can classify species, determine the number of dolphins in a pod, and study their territorial behavior. However, without accurate modeling of the structure of dolphin whistles reliable detection of these signals remains a challenge, since they have an unknown time varying duration and frequency band. In addition, the distribution of ambient noise can vary between marine environments, making adaptation difficult without intensive parameter calibration. While simple blind detection techniques such as variants of the energy detector seem to be the ideal solution to the problem at hand, these methods yield significant false alarms in noisy environments that include shipping noise. In this talk, we will present the results obtained by adopting an alternative detection approach based on deep learning. The proposed detector is based on the popular VGG16 convolutional architecture used in machine vision, which we fine-tune through supervised learning on a dataset containing human-annotated dolphin whistles. We benchmark the performance of our detector in two extremely different marine environments: the Red Sea, using a dataset we collected, and the DCLDE2011 dataset including bottlenose dolphin whistles, which we manually labeled as part of a citizen science project. To assess the robustness of our approach, the deep learning detector was only trained on the former dataset. As expected, the detection accuracy was fairly high on the test set sampled from the Red Sea recordings, while it decreased on the test set sampled from the DCLDE dataset. Nevertheless, our results also suggest that the proposed method can generalize to new recordings to some extent, and thus has significant potential for the goal of robust detection without parameter calibration.

Using stereophonic passive acoustics to study humpback singers' interactions

Y. Doh[1*], B. Denis[1], B. Ecalle[1], M. Ovize[1], G. Nolibé[2], O. Adam[3]

[1]*ABYSS NGO, 1 Rue du Quai Berthier, 97420, La Réunion, France*

[2]*CeSigma and Systems, France*

[3]*Sorbonne Université, Centre National de la Recherche Scientifique, UMR 7190, Institut Jean Le Rond d'Alembert, France*

**yanndoh.m2@gmail.com*

During autumn and winter acoustic interaction between humpbacks individuals are very intense and take and important place into the global mating strategy. In the breeding season, Male humpbacks are known to produce organized vocalizes within the time in sequences and phrases called leitmotiv of a "song". These sound productions are called sound units with a variable frequential complexity, of a given duration and followed by a given silence. A part of the Time/Frequency pattern is related to the anatomy of the laryngeal area and their behavior motivations. About the purpose of these songs, the main hypothesis of female attraction and territory defense were emitted but precise mechanism remains an open question today. One explanation to the lack of knowledge could be that most of song studies dealt with isolated singers recordings. Despite, these studies were interesting for primary describing the sound units and patterns of individual song, it was unfortunately disconnected from the communication aspect of these song productions. Thus, to replace the song in its context of interactions with other individual, passive acoustic devices equipped with just 2 hydrophones can be useful to track different singers vocalizing simultaneously, extracting various features and observe or not change/reaction into respective sound production. The BAOBAB dataset collected in 2012 in the channel of Sainte Marie (Madagascar), represents 9 days of continuous stereophonic recordings. Previous study has shown that stereophony allowed to estimate the density of acoustically active humpbacks in the channel. More than 10 years after, this dataset is still investigated and this presentation deals precisely with the application of an innovative stereophonic data processing for multi-singers analysis. Beside the method development, an example of acoustic interaction is highlighted and precise reactions in sound production are described for the first time.

Improving automatic detection with supervised contrastive learning: application with low-frequency vocalizations

G. Dubus[1*], O. Adam[1], D. Cazau[2]

[1]*Institut d'Alembert (Sorbonne Universite, CNRS), Paris France*

[2]*Lab STICC, ENSTA Bretagne, Brest France*

**gabriel.dubus@dalembert.upmc.fr*

Automatic detection of cetacean vocalizations from passive acoustic monitoring (PAM) datasets using machine learning remains a challenging task. The large diversity of sound types and variability of surrounding soundscapes and marine environments figure among the numerous complexity factors. The scarcity of vocal activity of many species within the long-term time series recordable with PAM adds the difficulty of addressing imbalanced datasets. Furthermore, the complexity of manual annotation results in incomplete and/or noisy labels, which further complicate the classification/detection task. While open-source annotated datasets have already been released in the DCLDE community, classical methods based on supervised learning are still limited, especially in their capability of generalization.

Recently, contrastive learning methods have shown great performance in automatic detection and classification of images and sounds, and outperforms classical supervised learning (i.e. the classical CNNs + MLP), notably in cases of imbalanced datasets, common in PAM datasets. Those methods focus on training models to acquire valuable representations, emphasizing interpretability and latent features. The objective is to cultivate an embedding space where similar data samples are positioned closely, while dissimilar ones are distanced from each other. In a second step, the detection/classification task is done from this embedding space.

In this study, we explore the potential interest of such methods in underwater PAM applied to low-frequency vocalizations of Antarctic fine whale and blue whale, within the largest public labeled acoustic dataset, especially dedicated to those species (Miller, et al., 2021). The supervised contrastive learning is compared with the most commonly used methods in the DCLDE community. This application is done using validation sets from different geographical areas based on several underwater PAM datasets (with examples based on DCLDE datasets from previous editions), to evaluate the generalization capacity of such methods.

The general objective of this work is to put on the scope of the PAM community one of the newest methods for automatic detection and classification.

[1] Miller, B.S., The IWC-SORP/SOOS Acoustic Trends Working Group., Balcazar, N. et al. An open access dataset for developing automated detectors of Antarctic baleen whale sounds and performance evaluation of two commonly used detectors.

Bounding Box-Based Object Detection for Whale Vocalizations

E. Ferguson[1*], G.C. Alongi[1], P.C. Sugarman[2]

[1]*Ocean Science Analytics LLC, San Diego, California, 92129, USA*

[2]*Acoustic Interactions, LLC, Bellevue, Washington, 98008, USA*

**eferguson@oceanscienceanalytics.com*

In the context of conserving the critically endangered North Atlantic right whale, the method of call detection can serve a multitude of purposes and applications. Passive acoustic call detection can encompass various forms, including timestamp recording, contour tracing, or the creation of a bounding box around a call of interest. However, to obtain measurements from a call or localize a sound with a simple timestamp, additional processing steps are required, and contour tracing can pose challenges in capturing the full extent of complex calls. Automated, accurate creation of a bounding box around a call of interest provides temporal and spectral boundaries that are computationally efficient and can be leveraged by measurement and localization tools to facilitate facilitating detailed analyses of whale vocalizations. By leveraging the extensive DCLDE 2013 dataset and a small subset of the 2024 sonobuoy data to incorporate the diverse ambient and instrument noise, a series of YOLOv4 framework with a CSP-DarkNet-53 architecture neural network were trained to detect right whale upcalls. Our aim was to quantify the degree and quality of overlap between ground-truthed and predicted call bounding boxes, by reporting the Intersection over Union (IoU) and Average Precision (AP) for test calls from the DCLDE 2024 dataset. This form of call detection accommodates post-detection tasks, as a higher degree of accuracy in IoU and AP can be instantly leveraged for measurement and localization using interoperable data products, further emphasizing its utility in whale conservation efforts.

An open-source deep learning model for North Atlantic right whale gunshot identification

F. Frazao[1,2*], B. Padovese[2], R. Gehrmann[3], O.S. Kirsebom[1,4], R. Joy[2], M. Dowd[5]

[1]*Faculty of Computer Science, Dalhousie University, Halifax, NS, Canada*

[2]*School of Environmental Science, Simon Fraser University, Burnaby, BC, Canada*

[3]*Defence Research and Development Canada, Dartmouth, NS, Canada*

[4]*Open Ocean Robotics, Victoria, BC, Canada*

[5]*Department of Mathematics and Statistics, Dalhousie University, Halifax, NS, Canada*

**fsfrazao@dal.ca*

Deep artificial neural networks are now staple algorithms for developing bioacoustic tools. These models present several useful characteristics for building DCLDE systems, including the ability to adapt to new environments given enough data, and a modular architecture that enables parts of trained networks, or their entirety, to be combined and repurposed. Several studies have used neural networks to detect and classify characteristic sounds of marine mammals, including North Atlantic right whales (NARW). The NARW upcall, a stereotypical contact call produced by individuals of all sexes and ages, is the most common target signal for automated passive acoustic monitoring efforts. However, NARWs produce a variety of other calls that can also be useful for acoustic monitoring. In this contribution, we present a neural network trained to identify NARW gunshots, the short broadband sounds produced by adult males, and believed to be related to reproductive behavior. We extracted annotated gunshot segments from the Gulf of Maine recordings provided in the 2013 DCLDE workshop along with upcalls and ambient noise segments. 3s-long mel-scaled spectrograms ranging from 0 to 1KHz were used to train a ResNet network. To evaluate the model's performance in identifying gunshots, we assembled 10 balanced test sets with 1000 gunshot samples and 1000 non-gunshot samples (500 upcalls and 500 ambient noise samples). The model achieved a mean recall of 0.89 (std:0.003) and a mean precision of 0.83 (std:0.11). We followed a similar testing procedure to evaluate performance in identifying upcalls and obtained a recall of 0.72 (std:0.009) and a mean precision of 0.91 (std:0.005). We note that our previous efforts to detect NARW upcalls yielded slightly better results, and highlight the combination of multiple models optimized for different call types as an option for monitoring efforts. The model trained for this contribution will be open-sourced and available to be integrated into future DCLDE systems.

Considerations when applying classification models across recording platforms: A case study with Hawaiian false killer whales

S. Fregosi[1*], J. Trickey[2], M. Wood[2], J.L.K. McCullough[3], Y. Barkley[4], A.E. Simonis[2], E.M. Oleson[3]

[1]*Ocean Associates, Inc. contracted to NOAA Pacific Islands Fisheries Science Center, Honolulu, USA*

[2]*Saltwater, Inc. contracted to NOAA Pacific Islands Fisheries Science Center, Honolulu, USA*

[3]*NOAA Fisheries: Pacific Islands Fisheries Science Center, Honolulu, USA*

[4]*Cooperative Institute for Marine & Atmospheric Research, Honolulu, USA*

**selene.fregosi@noaa.gov*

An effective two-step detection and classification approach was developed for identifying false killer whales (Pseudorca crassidens) from towed array recordings (DCLDE 2022 dataset). This approach uses Pamguard for detection of clicks, whistles, and burst pulses and then uses the banter R package for classification using a two-stage random forest approach. Events are classified as either 'Pc' (false killer whales) or 'UO' (unidentified odontocete). It would be useful to apply this approach to additional acoustic datasets that are not visually verified. But applying a model trained on one recording system will likely not result in robust classifications. Alternatively, a new model could be trained for each recording system, but in many cases would lack visual verification of species. We investigated the tradeoffs of applying a classification model trained on a different recording system, but with visually verified events, to a model trained on the same acoustic system but relying on expert assessment of manual species identification from recordings alone. We applied the model trained on towed array data (towed array model) to two additional datasets: a miniaturized HARP (High-frequency Acoustic Recording Package) deployed on longline fishing gear and a PMARXL (Passive Miniaturized Acoustic Recorder XL) installed in an underwater glider. Two trained analysts independently assessed over 200 cetacean events on both the longline and glider datasets to serve as the 'ground truth' dataset. Separate models were then trained for both the longline and glider datasets. Each model (towed array, longline, and glider) was subsequently applied to each dataset to predict species. Accuracy of the model predictions were assessed at three classification thresholds (80%, 60%, and 50%). Finally, a combined model was trained on all datasets and compared to the individual-recorder models. Preliminary results show that the accuracy varies by recorder type and is sensitive to detection and filtering steps implemented before model training as well as the number of signals of each type (clicks, whistles, and burst pulses) present in each event. Additional model improvement could likely be achieved through exploration of additional factors, such as a larger sample size or more diverse recording conditions. Because these automated approaches are invaluable in processing large amounts of passive acoustic data,

understanding even small differences in model performance across recorder types is essential.

Diving behavior and acoustic-based detection range inferred from three-dimensional tracking of beaked whales in the Gulf of Mexico

H. Frouin-Mouy[1,2*], K.E. Frasier[3], E. Snyder[3], J.A. Hildebrand[3], S. Wiggins[3], L. Garrison[2], M.S. Soldevilla[2]

[1]*Cooperative Institute For Marine And Atmospheric Studies (CIMAS), University of Miami*

[2]*NOAA Southeast Fisheries Science Center*

[3]*Scripps Institution of Oceanography, UCSD*

**heloise.frouin-mouy@noaa.gov*

The Gulf of Mexico (GOM) is a semi-enclosed large marine ecosystem inhabited by at least 20 cetacean species, including several species of beaked whales. Knowledge about beaked whales in the GOM is limited due to their pelagic habitat, the difficulty in identifying them visually, and their elusive behavior, with prolonged deep dives and short surface intervals. Fortunately, beaked whales produce species-specific echolocation clicks for both navigation and foraging. Fixed stations for long-term passive acoustic monitoring (PAM) are becoming more prevalent in the study of cetacean population dynamics, facilitating the estimation of animal densities. To use distance-sampling approaches for estimating density from passive acoustic detections, a critical component is evaluating the probability of detecting an acoustic signal given an animal's distance from an acoustic recorder. However, accurate detection probability estimation from single-sensor acoustic recorders requires simulation models that incorporate detailed population-specific information on subsurface behavior of both individuals and groups of animals. Unfortunately, such information is currently unavailable due to the absence of tag data for GOM beaked whales. Passive acoustic tracking, an alternative tool to tagging, can be used to study the diving behavior of beaked whales and provides the ability to characterize key aspects of their echolocation clicks. As part of the LISTEN GoMex study, which aims to improve understanding of cetacean population trends following the DWH oil spill and through ongoing restoration activities, two 4-channel High-Frequency Acoustic Recording Packages (HARPs) were deployed at Green Canyon (GC, depth ~1100 m) in 2021. Echolocation pulses from beaked whales detected on both recorders at GC were used to localize individuals in 3D to study their diving behavior. Twenty-six dives for Cuvier's, 25 dives for Gervais' and only 2 dives for Blainville's, due to their rare presence at this site, were selected for tracking. Results show that Cuvier's, Gervais' and Blainville's were detected at mean depths of 981 m (standard deviation = 97 m), 863 m (sd = 70 m), and 795 m (sd = 36 m), respectively. Some Cuvier's beaked whales spent at least some time foraging at or near the seafloor. The probability of a near-seafloor single-sensor HARP detecting individual beaked whale clicks was estimated using a cue-based detection probability simulation method. This Monte Carlo simulation utilized acoustic and behavioral parameters that were deduced from the acoustic tracking data.

Horizontal detection range, received level and elevation distributions from localized encounters at GC sites were compared with the model predictions.

Using Variational Auto-Encoders and Temporal Convolutional Networks to classify bioacoustics data from a weakly labelled training set

L. Garrobé[1*], D. Gillespie[1], L. Stankovic[2], V. Stankovic[2], L. Rendell[1]

[1]*Sea Mammal Research Unit, School of Biology, University of St. Andrews, KY16 9TH, St. Andrews, UK*

[2]*Electrical and Electronic Engineering, University of Strathclyde, James Weir Building Level 5, 75 Montrose Street Glasgow, G1 1XJ, Glasgow, UK*

**lgf3@st-andrews.ac.uk*

In the last decade, Deep Learning (DL) methods have gained popularity in tasks such as the detection of animal sounds from large acoustic datasets. Nonetheless, underwater acoustic datasets pose a unique set of challenges when it comes to training DL methods. One such challenge is the prevalent use of weak labelling in bioacoustics datasets, where annotations are assigned to time intervals rather than individual vocalizations. This practice is mainly adopted for efficiency reasons, as it is less consuming to use weakly labelled data and thus it allows us to have a larger and more diverse dataset, all of which helps prevent overfitting in the training process. However, using long audio segments to train DL models significantly increases computational demands, often rendering it practically infeasible.

In this work we propose a two-step solution to leverage weakly annotated data for training DL detection models. Our case study involves binary classification of the presence/absence of sperm whale (Physeter macrocephalus) click trains in 4-minute-long recordings from a dataset comprising diverse sources and deployment conditions, with different types of background noise from other sea creatures and boats, to reduce overfitting and maximise transferability. We (1) explored methods for extracting crucial acoustic features from lengthy audio segments without additional labels and (2) integrated Temporal Convolutional Networks (TCNs) trained on the extracted features for sequence classification. In the feature extraction step, we introduced an innovative approach using Variational AutoEncoders (VAEs) to extract information from waveforms and spectrograms. This approach eliminates the necessity for manual threshold setting or time-consuming strong labelling. For classification, TCNs were trained separately on sequences of either VAE embeddings or handpicked acoustic features extracted from the time series and spectrogram representations using classical methods. The TCN demonstrated robust classification capabilities on a validation set, achieving accuracies exceeding 85% when applied to 4-minute acoustic recordings. This performance was observed using both VAE embeddings and manually selected features. Notably, TCNs trained on handpicked acoustic features exhibited greater variability in performance across recordings from diverse deployment conditions, whereas those trained on VAEs showed a more consistent performance, highlighting the robustness of VAEs for feature extraction across different deployment conditions.

North Atlantic right whale detection and localisation using deep learning, nonlinear Bayesian inversion, and sound propagation modelling

R. Gehrmann[1*], O.S. Kirsebom[2], B. Padovese[3], F. Frazao[3], K. Thebeau[5], D. Barclay[5], S. Pecknold[1]

[1]*Defence Research and Development Canada, Dartmouth, NS, Canada*

[2]*Open Ocean Robotics, Victoria, BC, Canada*

[3]*School of Environmental Science, Simon Fraser University, Burnaby, BC, Canada*

[4]*Faculty of Computer Science, Dalhousie University, Halifax, NS, Canada*

[5]*Dalhousie University, Halifax, Canada*

**romina.gehrmann@forces.gc.ca*

A deep learning model for NARW upcall detection, trained with openly available acoustic datasets, is adapted and optimized to the DCLDE 2024 dataset. Upcalls identified by the model are filtered using an automated algorithm that selects upcalls suitable for localization estimates based on a set of criteria. Two different approaches to localizing the calling whale are explored. The first approach, applicable when the same call is detected in multiple hydrophones, estimates the location from the time-difference of arrival using nonlinear Bayesian inversion. The second approach localizes the whale by combining the bearing estimate from the directional hydrophone channels with a range estimate based on the sound-pressure level, the latter relying on estimates of NARW source levels and accurate sound propagation modelling. The results of the two approaches are compared and their relative merits discussed. The work presented will utilize and extend the open-source Python packages Ketos and Kadlu.

Comparison of Detection Techniques for Fin Whale Calls in a Distributed Acoustic Sensing Dataset: A Step Towards Automated Localization.

Q. Goestchel[1*], W.S.D. Wilcock[1], S. Abadi[1]

[1]*School of Oceanography, University of Washington, Seattle, WA 98195*

**qgoestch@uw.edu*

Fin whales are a vulnerable species that was dramatically impacted by commercial whaling and now faces increasing anthropogenic occupation of the oceans. Monitoring fin whales is key to quantifying the recovery of their populations and developing plans to mitigate the impacts of ongoing human activities. Fin whales are well suited for acoustic monitoring because male fin whales produce a stereotyped, 1-second-long 20-Hz down-swept chirp that is incorporated into songs associated with breeding and observed in non-song vocalizations in transiting groups. Fin whales also produce a higher frequency call associated with feeding. Distributed Acoustic Sensing (DAS) is an emerging technology that injects repeated laser pulses into an optical fiber and uses Rayleigh scattering to measure changes in optical length that result from strain (on short time scales), which makes it suitable for passive acoustic monitoring. The main advantage of DAS is that it behaves like a linear array of receivers with an unprecedented spatial resolution of tens of meters over distances up to ~100 km. For 4 days in November 2021, a public domain DAS dataset was collected on the two submarine cables of the Ocean Observatories Initiative Regional Cabled Array, that extend offshore Pacific City in Oregon. This experiment occurred during the breeding season for fin whales, and tens of thousands of 20-Hz calls are present in the data. The dataset comprises DA measurements on three fibers extending out to 65-95 km with a channel spacing of 2 m and is 26 TB large. Automated detection and localization techniques are needed to support studies of fin whale behavior, make density estimates, and understand the performance of DAS for marine mammal monitoring. Our study aims to develop scalable methods to automatically identify and localize fin whale calls in DAS data. As a first step, we are considering small representative samples of data and comparing different detection techniques. Specifically, we are considering match-filtering and spectrogram cross-correlation applied to single channels and sums of adjacent channels and image filtering and enhancement applied to record sections of the waveform envelope. We plan to compare both sensitivity and computational efficiency of these approaches and their performance in picking the arrival times of calls. This work will then form a basis for developing approaches to automatically localize fin whales using the time difference of arrival method. [work supported by ONR]

Automatic detection of spectra-banded echolocation clicks in Skagerrak, North Sea

E.T. Griffiths[1*], S. Sveegaard[1], L.A. Kyhn[2], C. Marcolin[1], J. Teilmann[1], J. Tougaard[1]

[1]*Department of Ecoscience, Aarhus University, Roskilde, 4000, Denmark*

[2]*NIRAS A/S, Kæhlersbakken 5, 4700 Næstved, Denmark*

**emilytgriffiths@ecos.au.dk*

The white-beaked dolphin, Lagenorhynchus albirostris, is the second most common cetacean in the North Sea , yet our knowledge of its habitat use and distribution is poorly understood. A few publications have documented the white-beaked dolphin's echolocation spectrum and summarized different acoustic characteristics. Reported peak frequency and bandwidth vary both between and within studies, however most studies report a distinct pattern of bands/notches in the spectrum, between 20-80 kHz on 75-90% of clicks recorded in the wild. We present a simple, computationally efficient technique to automatically identify the occurrence of spectral bands/notches in echolocation clicks and burst pulses most commonly associated with white-beaked dolphins in northern European waters. Four broadband recorders (sample rate: 384 kHz) were deployed in Skagerrak, northern North Sea, between May-August 2021, two each in the Natura 2000 conservation sites Gule Rev and Store Rev. Recordings were analyzed with the PAMGuard Whistle & Moan and Click detectors, with the resulting binaries further processed in R to identify timeframes of likely dolphin activity. These results were manually audited to create a dataset of presumed white-beaked dolphin acoustic events (n=23). Waveforms for all clicks and burst pulses were extracted from the PAMGuard Click Detector binaries in R. A template with spectral bands/notches between 20-80 kHz was created using a nonlinear least squares model from 35% of audited dolphin clicks, from events that contained no less than 100 clicks with a signal-to-noise ratio of at least 40 dB (click peak level over the broadband noise floor (10 Hz-10 kHz)) (n=21). This resulted in a sinusoidal template, which mapped the spacing and amplitude of the spectral bands/notches. Using Pearson correlation, the template was tested against the mean spectra of each dolphin and not-dolphin event (n=313), allowing for sine wave phase shift to match the spectrum bands/notches. At the 5% significance level, all audited white-beaked dolphin were positively correlated with the template, along with 20% of not-dolphin events (n=68). However, at 1% significance, 87% of dolphin events (n=20) had a positive correlation with the template, with only 3% of not-dolphin events mis-categorized (n=12). Our method would still benefit from a manual audit of results, however it dramatically improves the capacity to detect broadband signals with bands/notches in the spectrum, thereby facilitate detection and identification of white-beaked dolphins. Such a capacity is an important step to automate broad-scale mapping of a common, yet still poorly known species in Danish waters.

MAMBAT: a framework to track and localize multiple marine mammals with wide baseline, stationary arrays

P. Gruden[1,2*], E.-M. Nosal[2], E. E. Henderson[3]

[1]*Cooperative Institute for Marine and Atmospheric Research, Honolulu, HI 96822, USA*

[2]*Ocean and Resources Engineering, University of Hawai`i at Manoa, Honolulu, HI 96822, USA*

[3]*Whale Acoustic Reconnaissance Project, NIWC Pacific, San Diego, CA 92152-5001, USA*

**pgruden@hawaii.edu*

Multiple sources, spurious clutter and missed detections are a challenge common to all localization methods. Various approaches that use single animal localization methods have been used to localize in multiple animal situation. These usually rely on Time-Difference-Of-Arrivals (TDOAs) and/or call association steps. Although more generalized methods also exist, they are typically computationally intensive and employ case-by-case or ad-hoc steps.

This presentation will introduce our Multiple-Animal Model-Based Acoustic Tracking (MAMBAT) framework and demonstrate its capabilities on the US Navy's AUTEC test range data from the 2nd DCLDE workshop. Unlike conventional methods, MAMBAT employs a "track-before-localize" strategy, which eliminates the need for detection, classification, and association steps. MAMBAT consists of two phases. In the first phase, TDOAs and Direct-Reflected-Time-Differences (DRTDs) are automatically tracked. In the second phase, TDOA and DRTD tracks are used to compute ambiguity surfaces and obtain spatial position estimates that consider depth-dependent sound speed profiles. Based on position estimates, multiple targets are subsequently tracked in the 3-D spatial domain. We demonstrate MAMBAT as a fully automated and computationally efficient approach that can successfully track single and multiple animals. It is useful for widely-spaced or compact bottom-mounted arrays, and can be scaled to other scenarios. [Work supported by ONR's Marine Mammals and Biology Program].

Bottlenose dolphins show clear geographic variation in whistles when controlling for repeated, stereotypical signals

C.R. Hargrave[1*], J.N. Oswald[1], V.M. Janik[1]

[1]*Scottish Oceans Institute, School of Biology, University of St Andrews, St Andrews KY16 8LB, UK*

**crh8@st-andrews.ac.uk*

Geographic variation in animal signalling has been described in several taxa, including avian, primate, and cetacean species. The presence of both micro and macrogeographic variation in vocal signals can thus present difficulties when developing classification models. Due to geographic variation in vocal characteristics within single species, classifiers need to be developed for the specific area they are to be used in. However, for many species, the presence and extent of this variation remains unknown. As such, the effective range of utilisation for many of these classifiers is also unknown. We analyse the presence of acoustic geographic variation at multiple geographic scales in whistles of a cosmopolitan delphinid, the common bottlenose dolphin Tursiops truncatus. We utilise a suite of 46 time-frequency measurements taken from the spectrograms of 980 whistles across 7 sites within the Northeast Atlantic Ocean, Gulf of Mexico and Eastern Tropical Pacific Ocean. Whilst controlling for the potentially biasing effects of highly stereotyped signature whistles, we employ a machine learning algorithm, the Random Forest classification model, to establish a) whether whistles from different sites are measurably distinct and classifiable based on acoustic evidence alone; b) whether this inter-site variation exceeds intra-site variation and c) whether this variation is present between two populations with overlapping ranges. Overall, whistles were classified to site of origin with substantial accuracy (71.6% ± 4.14 [1 SD], k = 0.66, n = 6). Intra-site variation was assessed by classifying to acoustic encounter within the East Scotland site, with corresponding photo ID data used to limit the potential for reoccurrence of individuals between encounters. This second model performed poorly (28.0% ± 3.68, k = 0.16, n = 7), implying that the inter site differences observed in model 1, outweighed those present within sites. Finally, the third classification model performed well when classifying whistles from two overlapping populations (81.3% ± 6.52, k = 0.63 when discriminating between whistles from the eastern and western Scottish coasts). This study informs our ongoing development of odontocete species classifiers for passive acoustic monitoring and highlights UK coastal regions which may require bespoke classification models. Additionally, the vocal differences between overlapping dolphin populations observed in this study can be used to infer social delineations, informing future stock assessments of bottlenose dolphins. Vocal identity forms a measure of biodiversity per the Convention on Migratory Species and thus acoustically distinct populations may constitute Evolutionarily Significant Units (ESUs) requiring individual consideration from policy makers.

Adapting distance sampling to account for non-ranging instruments: an example with Ocean Bottom Seismometer data for baleen whale density surface estimation

D. Harris[1*], A. Pereira[2], T.A. Marques[1,3], L. Matias[2], L. Thomas[1]

[1]*Centre for Research into Ecological and Environmental Modelling, University of St, Andrews, UK*

[2]*Instituto Dom Luiz, Faculdade de Ciências da Universidade de Lisboa, Portugal*

[3]*Departamento de Biologia Animal, Centro de Estatística e Aplicações, Faculdade de Ciências da Universidade de Lisboa, Portugal*

**dh17@st-andrews.ac.uk*

Monitoring programs to determine marine mammal densities are costly, so using opportunistic data from instrumentation deployed for other purposes is one way to achieve cost efficiency. Ocean bottom seismometers (OBSs) are prime examples of such "platforms of opportunity". Designed to detect earthquakes and study other geophysical processes, OBSs also record low frequency baleen whale calls. The Comprehensive Nuclear-Test-Ban-Treaty Organisation's International Monitoring System is another example of an opportunistic marine mammal monitoring platform. The "Combining global OBS and CTBTO recordings to estimate abundance and density of fin and blue whales", or CORTADO, project is focused on developing case studies and training materials to highlight the utility of both OBS and CTBTO datasets.

However, using opportunistic data often requires adjustments be made to standard analytical methods. Distance sampling, a standard density estimation method, can be used to estimate the average probability of detecting specified whale calls by OBSs, leading to an estimate of animal density. Distance sampling relies on estimated ranges to calling whales. However, a practical issue is that some OBSs may only detect whales, but cannot estimate ranges to them. This impacts distance sampling because the distances at which whale calls can be both detected and ranged to by an OBS may be smaller than simply detecting calls. Here we use a case study from the south coast of Portugal to demonstrate how a distance sampling analysis was adapted to incorporate detections from non-ranging OBSs to work towards a density surface of fin whales (Balaenoptera physalus). Of the 24 OBSs deployed between September 2007 and August 2008, 12 of the instruments could not estimate distances to fin whale calls. Multiple-covariate distance sampling was used to estimate the detection function using data from the ranging OBSs. Depth and month were used as covariates, enabling detection functions for the non-ranging OBSs to be estimated. Next, for each non-ranging OBS, the estimated probability density function was used to scale the number of fin whale detections to remove a proportion of calls that were likely detectable but could not be ranged to. Finally, the resulting detections were combined with the data from the ranging OBSs in a spatio-temporal GAM to estimate a density surface of fin whale

calls, which can be converted to a fin whale density surface with an appropriate multiplier of fin whale call production rate.

Localization of North Atlantic Right Whales (NARWs) in the Gulf of St. Lawrence using Passive Acoustics

L.S. Hsu[1*], Y.T. Wang[1], W.C. Hu[1], C.F. Chen[1]

[1]*Department of Engineering Science and Ocean Engineering, National Taiwan University*

**r11525103@ntu.edu.tw*

The North Atlantic right whale (NARW, Neobalaenid glacialis) is a critically endangered species facing severe population decline. Understanding their distribution and behavior is crucial for their conservation. This study presents a novel approach for NARW localization in the Gulf of St. Lawrence using acoustic data collected during two 32-sonobuoy deployments in July 2018. We employed long baseline time difference of arrival (TDOA) measurements, which require at least three sonobuoy receptions of the target signal. For localization, we divided the Gulf of St. Lawrence into a 10x10 meter grid and calculated the theoretical TDOA for each grid point based on the distance between sonobuoy pairs and sound speed in the water. By comparing these calculated TDOAs with the actual TDOAs obtained from cross-correlation analysis of NARW calls (focusing on high-SNR "Gunshot" and "Scream" calls), we identified the grid point with the minimum residual for each call, pinpointing the estimated NARW location. This TDOA-based localization method offers a cost-effective and scalable approach for tracking NARWs and potentially other cetaceans.

Robust real-time detection of right whale upcalls using the Medusa acoustic buoy

M.D. Hyer[1*], T.A. Mooney[2], D.A. Mann[3], A. Anderson[3], F.H. Jensen[1*]

[1]*Department of Ecoscience, Aarhus University, Roskilde, Denmark*

[2]*Biology Department, Woods Hole Oceanographic Institution, Woods Hole, USA*

[3]*Loggerhead Instruments, Sarasota, FL, USA*

***mhyer@ecos.au.dk; fjensen@ecos.au.dk**

Animals worldwide are facing increasing ecological pressures from global climate change and increasing anthropogenic activities. To transition to a renewable energy future, extensive offshore wind development is planned globally. In the Northeast of the United States and Canada, future development sites overlap with the migratory range of critically endangered North Atlantic Right Whales (NARW) and will lead to increased risk of ship strikes, pile driving impacts, and other noise exposures. New methods to accurately detect cetaceans and provide real-time feedback for mitigation will be increasingly important to facilitate the recovery of the NARW and ensuring a sustainable green transition.

Recent developments in acoustic event detection made possible by deep learning have shown significantly improved detection performance across many different taxa but tends to be too computationally demanding to run on existing wildlife monitoring platforms. Here, we use spectrogram normalization and model compression techniques combined with an acoustic recording platform (the Medusa satellite-linked recording system) to bring real-time detection with deep learning to the edge. We test if edge-based inference using a compressed network running on a microprocessor entails significant performance loss and find that this loss is negligible.

Deployed wildlife monitoring solutions need to be robust to a variety of noise scenarios to provide robust detection. To encourage deep learning-based detectors to generalize well to diverse soundscapes, we leverage large, open-source datasets of noise from the NOAA SanctSound project for generating semi-synthetic training datasets.

We evaluate performance on benchmark DCLDE 2013 datasets in addition to other openly available datasets, and we compare performance against the published Shiu et al. 2020 upcall detection algorithm. On the 2013 DCLDE benchmark challenge, our compressed model achieves slightly better performance (AUC 0.913) compared to the baseline (AUC 0.898) despite a 100-fold reduction in model size. We achieve significantly better performance compared to the baseline across all other tested recording sites in the Western North Atlantic Ocean, demonstrating a relatively robust detection performance throughout the entire migratory range of the NARW.

AUTOMATED DETECTION AND CLASSIFICATION OF NORTH ATLANTIC RIGHT WHALE CALLS USING AN ENSEMBLE OF PAIR-WISE CLASSIFIERS

S.M. Jarvis[1*], A.E. Carroll[1]

[1]*Naval Undersea Warfare Center, Newport RI USA*

**susan.m.jarvis.civ@us.navy.mil*

The severely endangered North Atlantic right whale (Eubalaena glacialis, Eg) inhabits the Atlantic Ocean primarily between 20o and 60o north latitude. During the winter months, the population is known to migrate down the east coast of the US to an area NOAA has designated as the Southeast U.S. Critical Habitat, which runs along the coast from Cape Fear, NC to Cape Canaveral, FL. The U.S. Navy's Jacksonville Shallow Water Training Range (JSWTR) facility is installed approximately 95 km east of Jacksonville, FL, in an area just east of the identified southeastern critical habitat. The newly complete range is comprised of 258 bottom-mounted broadband (~20Hz-48kHz) hydrophones covering an area of approximately 1700 km2. The density, permanence and invariance of this sensor field offers a unique resource for passive acoustic monitoring of NARW and other species.

Unfortunately, despite much excellent work by a number of researchers, robust fully-automated NARW classification remains challenging. One reason the problem is so difficult is because it is inherently unbalanced. Across any interval of time and for any level of effort expended, NARW calls are much less likely to be present than calls from confusion species. However, one way to build confidence in the classification of a rare event, like a NARW call, is to become confident in what it is not. We will present a multi-class classification approach that builds pairwise classifiers for NARW up-calls versus calls from common confusion species including Atlantic minke whales, humpback whales and anthropogenic sounds. The aggregate results from the ensemble of pairwise decision functions are jointly scored to determine the class present. The classifier uses the low frequency (LF) FFT data from the Marine Mammal Monitoring on Navy Ranges (M3R) system to form smoothed, de-noised, decimated spectrograms of the calls and was trained using data from past DCLDE conferences and other public sources. Results from testing with the current and past DCLDE data sets as well as with data recorded from JSWTR hydrophones will be presented.

Advancing Automated Acoustic Monitoring through Self-Supervised Learning: Applications in Marine Mammal Detection and Classification

L. Kitchell[1*], M. Ogg[1]

[1]*Johns Hopkins University Applied Physics Laboratory*

**lindsey.kitchell@jhuapl.edu*

In light of recent advancements in self-supervised machine learning and the increasing availability of publicly accessible passive acoustic monitoring (PAM) data, an exciting opportunity emerges for the development of a general-purpose underwater acoustic feature extractor. Self-supervised learning, which eliminates the need for labeled data (a bottleneck for supervised training approaches), empowers models to derive meaningful and transferable representations directly from raw audio signals. Leveraging these techniques, we trained a convolution neural network model using hydrophone recordings spanning multiple years from the SanctSound PAM dataset. This dataset, containing long-term recordings from 30 locations within 8 U.S. marine sanctuaries, serves as a rich source of diverse training data. We explore the learned representations of the model and assess its performance across various downstream tasks. These tasks include the detection and classification of North Atlantic Right Whales within the 2024 DCLDE dataset, classification of multiple species from the 2022 DCLDE dataset, and classification of over 40 marine mammal species from the Watkins Marine Mammal Sound database. The results underscore the adaptability and efficacy of our self-supervised learning approach in the domain of underwater acoustic monitoring.

Cumulative prediction yields accurate species ID: presenting new acoustic classifiers for delphinids of the northeast Atlantic

T. Kleyn[1*], V.M. Janik[1], D. Risch[2], L. Hartny-Mills[3], O. Boisseau[4],J.N. Oswald[1]

[1] *Sea Mammal Research Unit, School of Biology, University of St. Andrews, Fife KY16 8LB, United Kingdom*

[2] *Scottish Association for Marine Science, Dunbeg PA37 1QA, United Kingdom*

[3] *Hebridean Whale & Dolphin Trust, Tobermory PA75 6NQ, United Kingdom*

[4] *Marine Conservation Research, Inc., Kelvedon CO5 9AG, United Kingdom*

**tk81@st-andrews.ac.uk*

While the popularity of passive acoustic monitoring continues to grow in the northeast Atlantic, monitoring of delphinids at the species level remains difficult due to the absence of reliable acoustic classifiers in this region. The vocalizations of seven commonly occurring northeast Atlantic species (Delphinus delphis, Grampus griseus, Globicephala melas, Lagenorhynchus acutus, Lagenorhynchus albirostris, Orcinus orca, and Tursiops truncatus) exhibit considerable overlap in time-frequency characteristics of the sounds they produce, complicating the task of acoustic species identification. Here, we address this gap and present a comparison of machine learning acoustic classifiers developed for predicting species identity of the seven named species using characteristics of their whistles and echolocation clicks. Over 30,000 vocalizations from 105 spatiotemporally independent encounters with delphinids were extracted from a visually confirmed dataset of recordings collected across the northeast Atlantic and in the North Sea and used to train classification models. Highest accuracy was achieved by combining click classification and whistle classification using a custom implementation of convolutional neural networks (CNN) in Python 3.10.0, on average predicting 84% (SD = 3.8%) of acoustic encounters with delphinids correctly. Here, separate whistle and click models predicted 4-second frames containing detected vocalizations, represented as time-frequency contours for whistles and mean power frequency spectra for clicks; cumulative whistle and click predictions over encounters were used by a separate Random Forest model to inform overall encounter prediction. This approach resulted in higher than 70% correct classification for all species but Delphinus delphis (64%, SD = 9.3%). The CNN + Random Forest classifier outperformed classification using Random Forest analysis alone, where time-frequency measurements of whistles and normalised band energies of clicks were used to base predictions, by more than 10% ($p < 0.001$) and showed a higher degree of consistency in accuracy across species. Across both methods, accumulating information from vocalizations over time significantly improved classification for both clicks and whistles, demonstrating the importance of context in classification and corroborating similar findings from previous studies. The classifiers developed are being integrated into open-source PAMGuard software to help progress monitoring at the species level.

Three-dimensional localization of dolphin sounds from an underwater drifter using short-aperture arrays and acoustic vector sensors

A.B. Laferriere[1*], A. Thode[1], D.A. Bevans[2], E.J. Berkenpas[3], C.M. Shepard[3], L.A. Freeman[4]

[1]*Marine Physical Laboratory, Scripps Institution of Oceanography, UCSD, 9500 Gilman Dr., La Jolla CA 92109*

[2]*NUWC Keyport, Keyport, WA*

[3]*Second Star Robotics, 2333 Distribution Cir. Silver Spring, MD 20910*

[4]*NUWC Newport, Newport, RI*

**alaferriere@ucsd.edu*

In deep-water environments marine mammal tracking is typically performed by correlating signals across hydrophone arrays deployed near the ocean surface or bottom. Here we demonstrate how a depth-controlled platform drifting in the mid-water column can obtain three-dimensional positional fixes of dolphin vocalizations using both surface and bottom-reflected multipath, using both a short aperture vertical array and an acoustic vector sensor. Two autonomous opto-acoustic drifters were deployed in 1 km water depth roughly 50 km off San Diego, CA in February, June, and September of 2023. The drifters recorded at depths between 100 and 300 meters at 12- and 24-hour intervals. Each drifter was outfitted with a CTD (Conductivity, Temperature, Depth) sensor and an acoustic recording system featuring a six-element 1.75-meter aperture vertical hydrophone array, a tetrahedral hydrophone array, and either a 2-D Geospectrum M-35 (1-15 kHz bandwidth) or 3-D Wilcoxon VS-209 (500 Hz-7 kHz bandwidth) acoustic vector sensor. Over the course of the two to three-day deployments, the sensors successfully detected numerous marine mammal calls, including a pod of common dolphins. By leveraging the measured arrival angles and time delays from multi-path propagation, we both calibrate the sensors' orientations and derive a full 3-D fix of individual dolphin whistles. Additional deployments over North Atlantic seamounts also detected sperm whale multipath in 2 km water depth. Our findings suggest that deploying sparse configurations of vector sensor platforms in the mid-water column enhances multipath detection from both the surface and ocean floor. This strategy holds promise for mapping spatial distributions across spatial scales and frequency bands that traditionally necessitate large-aperture hydrophone arrays. [Work supported by ONR Task Force Ocean.]

The MIRRF classifier for PAMGuard: Using music information retrieval (MIR) techniques and metadata to differentiate whale calls from shipping noise

H.T. LeBlond[1*], L. S. Quayle[2], H. Yurk[2]

[1]*Institute of Ocean Sciences, Fisheries and Oceans Canada*

[2]*Pacific Science Enterprise Centre, Fisheries and Oceans Canada*

**wtleblond@gmail.com*

PAMGuard's Whistle and Moan Detector (WMD) module is a frequently-used tool for detecting killer whales and certain other cetacean species. However, many if not most of the detections produced by the module are not actually caused by killer whales, as the module does not have any specific means of automatically identifying the source of each detection. In areas with high vessel traffic and several marine mammal species occurring simultaneously, such as the Salish Sea, these false detections are generally caused by either vessel noise, environmental noise, or humpback whales. Here, we describe the creation and efficacy testing of the MIRRF (Music Information Retrieval Random Forest) classifier (https://github.com/htleblond/PAMGuardMIRRF), a PAMGuard plugin that takes in WMD detection data and attempts to identify the source of the noise that produced each detection. This plugin primarily uses techniques that are common in the field of music information retrieval (MIR) and uses features derived directly from the audio, in addition to detection metadata produced by the WMD. For testing, we used hydrophone audio from 12 locations in the Salish Sea with different depths, shoreline proximities, currents, and vessel densities, and from different time periods throughout the year.

The influence of toothed whale behaviour on detection probability and the implications for passive acoustic monitoring

J.D.J. Macaulay[1*], L. Rojano-Doñate[2], M. Ladegaard[2], J. Tougaard[3], J. Teilmann[3], T.A. Marques[4], U. Siebert[3], P.T. Madsen[2]

[1]*Scottish Oceans Institute, Institiud Chuantan na h-Alba, University of St Andrews, Scotland*

[2]*Department of Biology - Zoophysiology, Aarhus University, C.F. Møllers Allé 3, building 1131, 8000 Aarhus C, Denmark*

[3]*Department of Ecoscience – Marine Mammal Research, Aarhus University, Frederiksborgvej 399, 4000 Roskilde, Denmark.*

[4]*Centre for Research into Ecological and Environmental Modelling, University of St. Andrews, Scotland, UK.*

**jdjm@st-andrews.ac.uk*

Harbour porpoises are visually inconspicuous, but highly soniferous echolocating marine predators that are regularly studied using passive acoustic monitoring (PAM). PAM can provide quality data on animal abundance, human impact, habitat use and behaviour. The probability of detecting porpoise clicks within a given area (P) is a key metric when interpreting PAM data. Estimates of P can be used to determine the number of clicks per porpoise encounter that may have been missed on a PAM device, which in turn allows for the calculation of abundance and ideally non-biased comparison of acoustic data between habitats and time periods. However, P is influenced by several factors, including noise, propagation conditions and the behaviour of the vocalising animal. Whilst the influence of environmental conditions on P can usually be modelled, animal behaviour is much more difficult to predict and there is therefore a common implicit assumption that changes in animal behaviour have a negligible effect on P between different monitoring stations and across time. Using a simulation-based approach informed by acoustic biologging data from 22 tagged harbour porpoises, we demonstrate that porpoise behavioural states can have significant (up to 3x difference) effects on P. An exploration of drivers of this change suggests this could apply to other species of toothed whales. The magnitude of the behavioural effect on P has significant implications for PAM studies. For example, if bespoke experiments are conducted to determine P for a PAM device, then any estimated value of P may not be applicable to different environments or even other time periods. Assuming behaviour has a negligible effect on P can therefore potentially result in substantial over or under estimation of the true abundance, habitat use or effects of human disturbance, especially in studies which use fixed recorders such as SoundTraps or CPODs. Here we discuss our findings, the implications for PAM studies and potential future research directions that may help in overcoming these potential biases in PAM data.

The first step in the acoustic classification of beaked whales in the Amazon Mouth Basin

R. Machado[1*], J.P. Mura[1,2], F. de Castro[1], N. Soares[2], G. Ferreira[1], L. Kascher[2], B. da Silva[2], G. Rodrigues[1], T. Amorim[2], A. Andriolo[2]

[1]*Instituto Aqualie, Juiz de Fora, Brazil*

[2]*Programa de Pós Graduação em Biodiversidade e Conservação da Natureza, Universidade Federal de Juiz de Fora, Juiz de Fora, Brazil*

**raphaelmachado@aqualie.org*

Beaked whales, deep-diving cetaceans belonging to the Ziphiidae family, exhibit oceanic habits and cryptic behavior, making it challenging to access information about their biology, ecology, and distribution. In Brazil, 9 species of beaked whales are reported, but limited information exists regarding their acoustic behavior. This study aimed to acoustically classify the ziphiids in the Amazon Mouth Basin. Acoustic data collection was obtained through SoundTrap, operating at 384 kHz and 16 bits, on a dedicated research vessel in Brazilian equatorial margin. Audio files were aurally and visually browsed through Raven Pro 1.6 searching for clicks with diagnostic characteristics already reported for beaked whales: frequency modulated pulses. The PAMpal package in R was select to compute acoustic parameters of echocolocation clicks, including peak frequency, 3dB bandwidth, minimum, maximum, and central frequencies at 3dB, 10dB bandwidth, minimum, maximum, and central frequencies at 10dB. Inter-click intervals were manually obtained using Raven's selection function. A classification analysis was conducted using the Random Forest (RF) algorithm, utilizing 60% of the data for training and 40% for validation. The confusion matrix assessed errors and accuracy between classes, and values such as Out Of Bag (OOB) error, Global Accuracy (GA), Balanced Accuracy (BA) for each class, and p-value (<0.05) were calculated to evaluate the classification model's. Additionally, an ordination plot and an importance heatmap were created using the rfPermute package to visualize the proximity between classes and the most important parameters in classification, respectively. The classification model exhibited an OOB error of 25% and GA of 65.48% with the following BA values: Z01 = 73.59%, Z02 = 65.80%, Z03 = 72.14% and Z04 = 81.33%. Misclassifications were higher between Z02 and Z03 (27.27%) followed by Z04 and Z01 (20.0%). The ordination plot confirmed closer proximity among these detections. The mean decrease in accuracy indicated that the minimum frequency of the 3dB bandwidth as the most important parameter for classification. Despite a 20% of missclassifications in the matrix, Z01 and Z04 did not exhibit significant proximity in the ordination plot. While additional data is needed for precise algorithm training and species classification, the acoustic similarities between Z02 and Z03 suggest a taxonomic affinity, potentially belonging to the same genus or species. These results may contribute to the development of autonomous acoustic classifiers for Passive Acoustic Monitoring of deep divers in Brazil. Furthermore, these data enrich our understanding of the occurrence and bioacoustics of ziphiids in the Amazon Mouth Basin.

ACCURATE passive acoustic monitoring density estimation: are we there yet?

T.A. Marques[1,2*]

[1]*Centre for Research into Ecological and Environmental Modelling, The Observatory, University of St Andrews, St Andrews, KY16 9LZ, Scotland*

[2]*Departamento de Biologia Animal, Centro de Estatística e Aplicações, Faculdade de Ciências da Universidade de Lisboa, Portugal*

**tiago.marques@st-andrews.ac.uk*

Estimates of wildlife density were traditionally dependent on animal sightings. For cetaceans which are hard to detect visually but produce sounds that can easily be detected by hydrophones, passive acoustic monitoring (PAM) density estimation (DE) has become a possible and even cost-efficient alternative. A major approach to PAM DE is cue counting, where a sample of detected sounds produced by the animals are used to estimate animal density. Under such a setting, as in most others, we must care about detector performance, namely false negatives via detectability, and false positives, induced by error in detection or classification of sounds. Above and beyond those multipliers, and specific to cue counting, a cue production rate – the number of sounds produced per animal per unit time, is required to convert the number of sounds detected into animal density. Cue counting PAM-based reliable estimates of animal density therefore depend upon using accurate and unbiased cue production rate estimates, i.e., cue rates that are representative of the average cue rate for to the time and place the survey took place. The US Navy Living Marine Resources program funded ACCURATE project, a large project involving several partners, was designed to investigate acoustic cue production, with a focus on marine mammals, to inform PAM DE cue counting exercises. I will present an ACCURATE overview highlighting some of the main results. The project focusses on estimating cue rates from animal borne tags but includes also results from bottom mounted hydrophones and towed arrays. Results include an overview of the existing information for marine mammals, cue rate estimates for several species of deep divers, including sperm whales, beaked whales and narwhals, exploring different approaches to do so, and how to estimate cue rates for times and places for which tag data might not be available, as well as an evaluation of the scenarios under which baleen whale cue rates - challenging to obtain from animal borne tag data given difficulty in assigning low frequency sounds to the tagged individual versus their nearby conspecifics - might be possible to be obtained. I hope to promote the discussion some of the pending challenges and about under what scenarios are we likely to be able to get reliable cue rates to support accurate density estimates.

Rough-toothed dolphin exposures to U.S. Navy mid-frequency active sonar at the Pacific Missile Range Facility, Hawaiʻi

C.R. Martin[1*], E.E. Henderson[1], R.W. Baird[2], M.A. Kratofil[2], S.W. Martin[3]

[1]*Naval Information Warfare Center Pacific, San Diego, CA, USA*

[2]*Cascadia Research Collective, Olympia, WA, USA*

[3]*National Marine Mammal Foundation, San Diego, CA, USA*

**cameron.r.martin9.civ@us.navy.mil*

Eight rough-toothed dolphins satellite tagged between 2011 and 2021 before the start of U.S. Navy training at the Pacific Missile Range Facility (PMRF), Kauaʻi, Hawaiʻi were exposed to mid-frequency active sonar (MFAS) from multiple sources. Received levels from surface ship hull-mounted sonar, sonobuoy, and helicopter dipping sonar transmissions were estimated with propagation modeling. Tag positional data were fitted to a continuous time correlated random walk model to predict locations at 5-minute intervals, and incorporated x and y positional error. Positional error was based upon a 95% confidence interval error ellipse of the northern and easting standard errors at each 5-minute location. Dive behavior data were available for five tags and were utilized to develop dive models that could be used to estimate the receiver depth for propagation modeling when depth data were available. When animal depth data were not available (i.e., tags without depth sensor or dive data not valid for a 5-minute interval), two depth regimes were utilized to represent potential animal depths. A shallow depth regime was used to represent exposures when the animal was within 54 m of the surface where surface ducting was often present, and a deep depth regime covered remaining possible animal depths from 54 to 400 m. Error ellipses were sampled with multiple 2-dimensional radial slices (range vs. depth) taken systematically in azimuth from a source for select sonar transmissions to represent 3-dimensional propagation modeling. Estimated received levels in 5-minute bins are presented as the maximum median level +/- twice the standard deviation and characterize the range of potential exposures. To date, estimated MFAS exposures to these satellite tagged rough-toothed dolphins includes a maximum median received level of 159.8 dB re 1uPa with a range of estimated received levels +/-2*standard deviation of 149-170.6 dB re 1µPa from surface ship hull-mounted sonar. Based on photo-identification, tagged dolphins were part of a resident, island-associated population that broadly overlaps with PMRF, and thus are likely regularly exposed to MFAS.

Cetacean multi-species detection, classification, localization, and contact collation routine using PAMGuard

M. Matei[1*], D. Gillespie[2], J. Macaulay[2]

[1]*SMRU Consulting, Scottish Oceans Institute, University of St Andrews, East Sands, St Andrews KY16 8LB, Scotland, United Kingdom*

[2]*Sea Mammal Research Unit, Scottish Oceans Institute, University of St Andrews, East Sands, St Andrews KY16 8LB, Scotland, United Kingdom*

**mim2@st-andrews.ac.uk*

During real-time passive acoustic surveys of cetaceans, there can often be multiple target species, whose vocalisations can vary vastly in terms of frequency range, duration, production rate, and other signal characteristics. This means that multiple approaches are required to identify these different signal types, and that these approaches need to be able to run concurrently in real time. This study produced a PAMGuard configuration which includes 24 data management, signal processing, and DCL modules. These comprised two instances of the whistle and moan detector (searching different parts of the spectrum) and three combinations of click and click train detectors searching for fin and North Atlantic minke whale pulses, as well as the higher frequency echolocation clicks of odontocetes.

The acoustic datasets used to set up each DCL module had either been collected by the authors or were sourced from the MobySound website or previous DCLDE datasets. The fin and minke whale detectors were fine-tuned using parameters extracted from annotated data via custom MATLAB scripts. Additional software was developed to rapidly assess the performance of detectors, tested with multiple parameters, by automatically comparing PAMGuard output with databases of annotations. The performance of each detector and classifier was assessed using encounter efficiency and false alarm rate rather than metrics for single vocalisations. Overall, all but one classifier had an encounter efficiency of 100%, the exception being beaked whales (80%). False alarm rates for most species were below 0.5%, the exceptions being sperm whales, fin whales and minke whales where vessel-related sounds sometimes triggered the classifiers for these species. These performance metrics were subsequently used to investigate how many detections are required, and over what period, to identify a cetacean acoustic encounter while maintaining high efficiency and low false alarm rate. The optimal values for each module were used as parameters for a new PAMGuard module called 'Contact Collator'. This module can receive data from different detector types simultaneously, and it generates summary information (detector name, time, species group, bearing information) whenever the detection count and temporal thresholds are exceeded for a given signal type. Instead of having to monitor five different detectors during real-time monitoring activities, this module allows the operator to view a single summary data display, which provides encounter details such as spectrogram clips and bearing summaries for each DCL module in a common format.

A Convolutional Neural Network to detect bowhead whale vocalizations in passive acoustic data from the Arctic Ocean

M. Meister[1*], P. Keil[2,3,4], K. Thomisch[1]

[1]*Ocean Acoustics Group, Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany*

[2]*German Climate Computing Center, Hamburg, Germany*

[3]*Helmholtz Center Hereon, Geesthacht, Germany*

[4]*Helmholtz AI*

**marlene.meister@awi.de*

Bowhead whales (Balaena mysticetus), endemic to the Arctic, possess a complex and temporally variable acoustic behavior that is utilized in reproductive and social contexts. They produce single calls, usually frequency modulated vocalizations between 50 and 500 Hz, as well as structured series of vocalizations, referred to as songs. These songs are predominantly produced during winter and spring and display great intra- and inter-annual variability, with diverse songs being sung by a population within one season. Since manual detection of bowhead whale vocalizations in continuous Passive Acoustic Monitoring (PAM) data is a challenging and time-consuming task, the Ocean Acoustics group of the Alfred Wegener Institute (AWI) teamed up with the Helmholtz Artificial Intelligence Cooperation Unit to develop an AI-based algorithm for bowhead whale detection. To this end, we train a Convolutional Neural Network (CNN) to recognize vocalization signatures of bowhead whales in spectrograms generated from PAM data. The algorithm divides data into short-duration snippets, indicating the presence or absence of bowhead whale signals for each snippet. This approach has the potential to significantly streamline the analysis process, while enhancing objectivity of call identification. The network will be applied for the analysis of an extensive acoustic dataset (spanning 2104 recording days) collected by AWI in Fram Strait between 2012 and 2021. For training we use more than 4000 humanly labeled individual whale calls over several days. In the future, we aim to provide easy operational inference from the trained network for new data. Analyzing this acoustic data will further our understanding of trends in bowhead whale occurrence, hence, contribute to the development of effective conservation strategies for a species heavily affected by climate-change related habitat loss.

Near-real-time detection of odontocete echolocation clicks from a glider

D.K. Mellinger[1,2*], J. Turnbull[3,4], S. Fregosi[4], C. Jones[6], E.M. Oleson[7], E. Norris[7], M. Wood[9], H. Matsumoto[10]

[1]*Cooperative Institute for Marine Ecosystems and Resources Studies and Marine Mammal Institute, Oregon State University*

[2]*NOAA Pacific Marine Environmental Laboratory (PMEL)*

[3]*NOAA Pacific Marine Environmental Laboratory*

[4]*Cooperative Institute for Marine Ecosystems and Resources Studies, Oregon State University*

[5]*Ocean Associates, Inc. contracted to NOAA Pacific Islands Fisheries Science Center*

[6]*Embedded Ocean Systems, Inc.*

[7]*NOAA Fisheries, Pacific Islands Fisheries Science Center*

[8]*Cooperative Institute for Marine and Atmospheric Research, University of Hawai'i at Mānoa*

[9]*Saltwater, Inc. contracted to NOAA Pacific Islands Fisheries Science Center*

[10]*Living Ocean Systems, LLC, Newport, OR, USA*

**David.Mellinger@oregonstate.edu*

Ocean gliders equipped with acoustic sensors are used increasingly often for marine mammal research. They offer a complement to other common acoustic platforms such fixed hydrophones, which survey from one place for up to 2 years, and towed arrays, which cover a large area but often last only weeks. Gliders can survey tracklines of 1000-2000 km on missions lasting 2-3 months, and as such occupy a performance point between those other two. They are also intermediate in cost and low in carbon footprint. Acoustic gliders have mostly been used as recorders, with the recorded data processed after instrument retrieval, although some have been used for real-time detection of baleen whales, with the detection system running on-board the glider. Energy storage is a significant constraint on glider operation, and the high-frequency nature of echolocation clicks (tens to hundreds of kilohertz) implies they require significant processing power, and hence energy, on a glider. Here we describe a system for detecting echolocation clicks of odontocetes that's sufficiently low-power to operate for months on gliders. The hardware consists of a hydrophone and pre-amp connected to a WISPR audio capture and storage system. The WISPR additionally controls a Raspberry Pi (RPi) single-board processor; they share two large-capacity storage cards such that WISPR records on one while the RPi processes data on the other. The RPi uses significant power and is turned on for only a short duration during each glider dive. Detection software is based on click detection followed by the Energy Ratio Maximization

Algorithm (ERMA; Klinck and Mellinger 2011), which detects specific odontocete species based on differences in click energy present in two spectral bands. Click detections must additionally satisfy tests based on click timing and overall number of clicks in a set timespan. When a positive detection is triggered, a detection report is prepared, which upon the next glider surfacing is sent to shore for human confirmation. This detection system was implemented in a Seaglider™ and tested in spring 2023 on sperm whales in the Main Hawaiian Islands. Sperm whales were successfully detected in near-real time by the system, though there were some false positives and negatives. The recorder and detection system power usage averaged ~600 mW over the duration of a dive, low enough to last for a 3-month mission. [Funding from NOAA Uncrewed Systems Office, NOAA PMEL, and ONR.]

Classification updates for Kogia spp., compared to Narrow-Band High Frequency (NBHF) clicks from other species, and across multiple instrument types.

K. Merkens[1*], A. Holdman[2], N. Tregenza[3], E. Oleson[4]

[1]*Saltwater Inc. in support of NOAA/NMFS/PIFSC, Portland, OR, USA*

[2]*NOAA/NMFS/NEFSC, Woods Hole, MA, USA*

[3]*Chelonia Ltd., Mousehole, Cornwall, UK*

[4]*NOAA/NMFS/PIFSC, Honolulu, HI, USA*

**karlina.merkens@noaa.gov*

The classification of Narrow-Band High-Frequency (NBHF) echolocation clicks to species continues to be a challenge globally, with multiple species' distribution ranges overlapping at many locations worldwide and the clicks being similar enough that separating species is essentially impossible. In the northern hemisphere the ranges of species that generate NBHF clicks overlap across wide stretches of ocean, and based on what is known currently about their clicks they cannot be distinguished based purely on passive acoustic recordings; in most cases water depth is used as the primary characteristic for classifying species. In the North Atlantic and North Pacific there are multiple NBHF-clicking porpoise species (harbor porpoise, Phocoena phocoena, and Dall's porpoise, Phocoenoides dalli) whose ranges overlap with both species of the genus Kogia (dwarf sperm whale, K. sima, and pygmy sperm whale, K. breviceps). Previous research indicated a possible frequency sweep in NBHF clicks based on detections from F-POD instruments (Chelonia Ltd.) sampling at 1 MHz, and here we present further analysis of a larger sample size of clicks recorded from the Atlantic and the Pacific at locations where only a single NBHF species is present (based on water depth and long-term visual sightings records). We confirm a downward sweep in many porpoise clicks, while K. sima clicks do not have a clear sweep, but instead the frequencies fluctuate up and down throughout the duration of each click. Additionally we compare clicks recorded on an F-POD and on a High-frequency Acoustic Recording Package (HARP) sampling at 200 kHz, with both instruments on a single mooring deployed below 600m. While our ability to identify individual clicks simultaneously on both instruments is limited by discrepancies in time signatures, we can compare clicks in the same encounter to identify how the different hardware systems compare. Because of the higher sampling rate the F-POD captures more Kogia encounters than the HARP, however the higher amplitude events that are captured on both instruments include approximately the same number of clicks on both instruments. Further progress on distinguishing between species that generate NBHF clicks, as well as comparing performance between instruments that are capable of detecting these higher frequency signals, is essential for effective monitoring of these species worldwide.

Analysis of the DCLDE 2024 North Atlantic right whale sonobuoy dataset using PAMGuard

B.S. Miller[1*], J. Macaulay[2], D. Gillespie[2]

[1]*Australian Antarctic Division*

[2]*University of St Andrews*

**Brian.Miller@aad.gov.au*

Datasets and workshops on Detection, Classification, Localisation, and Density Estimation of marine mammals (AKA DCLDE Workshops) have been conducted more-or-less biannually for nearly two decades in order to benchmark and elicit novel solutions to real-world and forthcoming challenges in the field of marine mammal passive acoustics. The 2024 DCLDE included a publicly available dataset of recordings of critically endangered North Atlantic right whales made from arrays of up to 32 directional (DIFAR) sonobuoys deployed off Canada on two different days. DCLDE datasets like this serve as both a challenge and benchmark for acoustic detection, classification, and localisation methods. Here we present a solution to detection, classification, and localisation of North Atlantic right whale (NARW) upcalls recorded on directional (DIFAR) sonobuoys as part of the 2024 DCLDE dataset. The core of our solution is based on open-source software for passive acoustic monitoring, PAMGuard. We used the PAMGuard Deep Learning Module for detection and classification of right whale upcalls. The PAMGuard DIFAR module was used to obtain bearings for 194 detections, and for triangulation of 58 calls that were received on multiple DIFAR sonobuoys from approximately 3.5 h of recordings from 31 July 2018 (i.e. Day 2 of the DCLDE dataset). Qualitative comparison of acoustic results with visual sightings also included in the 2024 DCLDE dataset indicate that our acoustic triangulations were in broad accord with the visual sightings. We discuss the benefits and limitations of using ‘vanilla, off-the-shelf’ PAMGuard for detection, classification, and localisation in regards to this dataset. We also discuss opportunities for improvement of the 2024 DCLDE datasets that could facilitate density estimation (the DE part of DCLDE).

Using deep learning and dynamic visualizations to efficiently detect minke whales across the Atlantic Ocean

X. Mouy[1*], T. Awbery[2], G. Davis[3], O. Fernandez-Betelu[4], V. Iorio-Merlo[4], J. Macaulay[5], M. Ryan[6], L. Transue[7], S. Van Parijs[3], D. Risch[2]

[1]*Integrated Statistics, Inc., under contract to National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA 02543, United States of America*

[2]*Scottish Association for Marine Science, Scottish Marine Institute, Oban, Argyll PA37 1QA, United Kingdom*

[3]*National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA 02543, United States of America*

[4]*School of Biological Sciences, Lighthouse Field Station, University of Aberdeen, Cromarty, Ross-shire IV11 8YJ, United Kingdom*

[5]*University of St Andrews, St Andrews KY16 8LB, Scotland, United Kingdom*

[6]*IBSS under contract to National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA 02543, United States of America*

[7]*Under Contract to the Northeast Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Woods Hole, MA, USA*

**xavier.mouy@noaa.gov*

Minke whales (Balaenoptera acutorostrata) are distributed across all oceans with the exception of the Southern Ocean. Despite being an abundant species worldwide, minke whales face several conservation challenges: they are still commercially hunted in Norway, Iceland and Japan, and are considered at risk from entanglement, vessel strike, underwater noise, pollution, and climate change. While the summer presence of minke whales in the Atlantic Ocean is relatively well documented, their winter distribution is still poorly understood. Because minke whales are soniferous, passive acoustics can contribute to a better understanding of their distribution and help their conservation. The objective of this study is to develop a methodology to efficiently detect minke whale pulse-trains across the entire Atlantic basin. To do so, we trained a binary ResNet18 deep neural network (DNN) to recognize minke whale pulse trains from spectrogram representations of the acoustic signal. The network was trained using more than 81,000 manually annotated sounds from minke whale (4708) and non-minke whale sounds (76975, that included ambient, haddock, humpback, seismic survey, and pile driving sounds) collected along the north-east coast of the United States, in the Caribbean, and off the east and west coast of Scotland. To capture the variability of acoustic soundscapes in the model, we trained and evaluated the DNN

using data from six different types of acoustic recorders deployed between 2006 and 2022 at 47 different locations and in water depths ranging from 20 m to 180 m. Once trained, the DNN was tested on several months of continuous data collected in the Stellwagen Bank National Marine Sanctuary, in the Massachusetts-Rhode Island offshore windfarm area and off Scotland. To expedite the manual verification of the detections from the DNN, we created a dynamic graphical interface to visualize and navigate through the detections time series, filter and sort by classification confidence, and dynamically produce spectrograms of the detections. We will present the performance of the detector at the different locations and discuss its ability to report on daily presence, hourly presence, and pulse-train count. We will also demonstrate how the detector is being integrated to the software PAMGuard to facilitate its use.

Four years of daily acoustic abundance estimates of fish-eating and mammal-eating killer whales in the Gulf of Alaska

H.J. Myers[1,2*], B.H. Konar[2], D.W. Olsen[3], T. Margolina[4], F.J. Mueter[2], R.A. Guazzo[5], L.A. Horstmann[2], C.O. Matkin[3]

[1]*Marine Mammal Institute, Oregon State University, Newport, OR, USA*

[2]*College of Fisheries and Ocean Sciences, University of Alaska Fairbanks, Fairbanks, AK, USA*

[3]*North Gulf Oceanic Society, Homer, AK, USA*

[4]*Naval Postgraduate School, Monterey, CA, USA*

[5]*Naval Information Warfare Center Pacific, San Diego, CA, USA*

**hannah.myers@oregonstate.edu*

Killer whales (Orcinus orca) are most abundant in high-latitude regions, where their year-round distribution and abundance is difficult to assess using visual methods. In the northwestern Gulf of Alaska, there are four populations of killer whales: southern Alaska residents (fish-eating), Gulf of Alaska transients (mammal-eating), AT1 transients (mammal-eating), and offshores (shark specialists). In this study, we applied call detection rates established using autonomous recording hydrophone data to provide four years of daily acoustic abundance estimates for southern Alaska resident and Gulf of Alaska transient killer whales across five locations. Acoustic abundance estimates matched expectations from visual studies and established distinct seasonal patterns for each population and location. On days fish-eating killer whales were detected, we estimated the mean number of animals for each location was nine to 18 animals, with maximum aggregations of up to 116 animals. For mammal-eating killer whales, we estimated a mean daily acoustic abundance of three animals at all locations with a maximum of 11 animals. The best Bayesian time series models explained 17% of deviance in mean daily abundance estimates for residents and 5% for transients, and included an ARMA (1,1) autocorrelation structure. Over a longer time series, adding additional covariates such as prey abundance or environmental factors may improve model performance. The abundance patterns we observed did not match seasonal or location-specific differences in modeled detection area. We focused only on estimating the number of animals present on days when killer whales are detected, and did not normalize by the detection area as killer whales exhibit specific local movement patterns. The results of this study may inform management policy for this federally protected species and advances passive acoustic monitoring capabilities for killer whales.

Domain Shift in Passive Acoustic Monitoring

A. Napoli[1*], P.R. White[1]

[1]*ISVR, University of Southampton, UK*

**an1g18@soton.ac.uk*

Machine learning models often underperform when deployed to new environments unseen during training. Changes in dataset characteristics, collectively known as “domain shift”, are particularly prevalent in passive acoustic monitoring (PAM), although the effects of this are often not properly considered. In this talk, we discuss domain shift from a PAM perspective, and present a benchmark, based on the detection of humpback whales, intended to emulate challenging domain shift scenarios. The dataset comprises audio from 9 distinct “domains”, each with differing locations, recording methods, environmental conditions, and nontarget sound events. We use this benchmark to train and evaluate a simple CNN-based detector, along with 3 existing baselines. Our results highlight the importance of training on multi-domain data to allow models to generalise to new domains: our simple detector significantly outperforms all baselines, despite having 10,000 times fewer parameters and 100,000 times less training data than the next-best model. When this is not possible, we demonstrate careful use of unsupervised domain adaptation (UDA) can boost performance without requiring additional annotations.

Andrea Napoli and Paul R. White, ‘Unsupervised Domain Adaptation for the Cross-Dataset Detection of Humpback Whale Calls’, in DCASE, 2023. Available: https://dcase.community/documents/workshop2023/proceedings/DCASE2023Workshop_Napoli_21.pdf

ATDOA (asynchronous time difference of arrival): TDOA-based method to localize multiple sound sources using autonomous receivers

E.-M. Nosal[1*], T. Fedenczuk[2]

[1]*University of Hawai'i at Mānoa, Honolulu, Hawaii, USA*

[2]*Abakai International LLC, Waianae, Hawaii, USA*

**nosal@hawaii.edu*

A practical and cost-effective approach to passive acoustic monitoring is to create an array using autonomous receivers. Unfortunately, most passive acoustic localization methods rely on precise time-synchronization between the receivers of an acoustic array, which is not automatic for autonomous receivers with separate clocks. Time synchronization can be achieved in various ways, including by connecting the receivers to a common data acquisition device or by using control/calibration sources of known location to solve for timing offsets. Other approaches that can localize sources with asynchronous receiver leverage propagation effects (such as multipath arrivals or dispersion); these are powerful and effective but can be difficult to implement and only apply in certain cases (with appropriate source and environment conditions). This presentation introduces and explores another approach: Inverting globally for time synchronization offsets in addition to multiple animal positions. To keep the approach tangible and accessible to PAM practitioners, our ATDOA (asynchronous time difference of arrival) framework expands on the linearized time-difference-of-arrival (TDOA) method commonly used in the bioacoustics community. Simulation results are presented to illustrate ATDOA, and results from a controlled experiment using asynchronous microphones are presented. Examples of application to real-world marine mammal datasets are given. [Work supported in part by the ONR Marine Mammals and Biology program]

Deep Learning-Based Underwater Sound Classification Using Stacked Cepstral Features

A. Olcay[1*], P. White[1], J. Bull[2], D. Risch[3]

[1]*Institute of Sound and Vibration Research, University of Southampton, Southampton, UK*

[2]*School of Ocean and Earth Science, University of Southampton, Southampton, UK*

[3]*Marine Science Department, Scottish Association of Marine Science, Oban, UK*

**aao1n21@soton.ac.uk*

This study evaluates the impact of Mel-frequency cepstral coefficients (MFCCs) stacked with their dynamic features, in contrast to the use of linear frequency spectrograms, as input to deep learning models. While this approach has seen success in music genre classification, speech recognition, and vessel noise tasks, its potential in classifying diverse underwater sound sources remains underexplored. Using data from the COMPASS project's PAM moorings in Western Scotland, this study scrutinizes recordings from three different locations—Tolsta, Hyskier, and StoerHead—across various seasons, capturing diverse underwater sounds reflective of regional dynamics. Employing some state-of-the-art Convolutional Neural Network (CNN) models such as ResNet50, MobileNetV1, and EfficientNetB0 with stacked MFCCs, preliminary results show that acoustic data representation in the network front-end has the potential to mitigate the effect of differences in back-end models. Early finding demonstrates that stacked MFCCs approach provides a more generalizable classification across acoustic sites and relative insensitivity to DL model variations.

Validation of Passive Acoustic Density Estimation Approaches for Southern Right Whales (Eubalaena australis)

S.E. Parks[1*], L. Thomas[2], R. Finley[1], A. Eschmann[1], J. R. G. Dombroski[1]

[1]*Syracuse University, Syracuse, NY, USA*

[2]*University of St Andrews, St Andrews, Scotland, UK*

**sparks@syr.edu*

Passive acoustic monitoring (PAM) has become a standard method for detection of signals produced by species of interest with emerging methods to use PAM for density estimation (DE). One major limitation in testing the performance of PAM-DE approaches is having a verification dataset, where true density is known in a realistic field situation – for example through visual observations under conditions where animals within the study area are certain to be detected. In this study, PAM data were collected simultaneously with visual shore-based counts of Southern right whales (Eubalaena australis) in Brazil in 2021 and 2022 to provide a dataset with verified true density for assessing the performance of PAM density estimate methods. Concurrent visual and acoustic data from 18 days, divided into 32 sessions, were used to test Spatial Capture Recapture (SCR) models, including standard SCR which uses data on which sensors detected each sound and extended SCR which uses additional information such as receive level, noise level and time of arrival. Cue rates for southern right whales were estimated from the Dtags (22 tag deployments comprising 84 h of focal tag recordings from 14 previously collected datasets and 8 tags deployed in 2022) and visual focal follows of whales acoustically tracked inside the acoustic array in 2021 and 2022 (19 individual groups followed for 7 total hours of data). We present results comparing the visual and acoustic density estimations for right whales in this location. Several important insights were gained from this study to inform future PAM-DE projects. The combination of spatially varying animal density and propagation, together with a relatively low number of active sensors, meant that standard SCR did not perform well. Extended methods led to more reliable estimates of call detectability. Average cue rates for right whales were highly variable and dependent on the behavioral state. Improved estimates of time spent in different behavioral states by individual whales will allow for more accurate cue rate estimation. These data may be easier to collect than detailed individual acoustic call rate datasets, as visual observations can be used to estimate time in behavioral states.

Latent Diffusion Model Based Spectrogram Augmentation for Improved Baleen Whale Call Detector Robustness

D.M. Parry[1,2*], R.P. O'Shea[3], L Bouffaut[1], H. Klinck[1,4]

[1]*K. Lisa Yang Center for Conservation Bioacoustics, Cornell Lab of Ornithology, Cornell University*

[2]*Department of Natural Resources and the Environment, Cornell University*

[3]*Independent Researcher*

[4]*Department of Fisheries, Wildlife, and Conservation Sciences, Marine Mammal Institute, Hatfield Marine Science Center, Oregon State University*

**dmp295@cornell.edu*

Machine learning (ML) based automated detection methods increase the speed of acoustic analysis but are often overfitted to one region and require large amounts of labeled data for model training. Adding augmented data to training data can allow for generalization to sites with differing noise profiles and decrease the amount of manually labeled training data needed. Latent diffusion models have been increasingly used in recent years for image generation and translation. Paired with the additional diffusion process guidance provided by conditional controls, latent diffusion models provide a powerful tool for image feature translation. This technique can be used to augment image data used to train and evaluate neural networks. This work applied domain transfer techniques to spectrograms of 20 Hz fin whale (Balaenoptera physalus) calls to produce augmented spectrogram image data. Augmentations were performed as a two-step process to generate an augmented dataset using a publicly available dataset of calls from three sites around Antarctica (low levels of anthropogenic noise; training site) and background noise from Long Island, New York (high levels of anthropogenic noise; target location). First, calls from the training dataset were inserted into the target location background spectrogram using image masking and logical operators. Next, the inserted call spectrograms, image masks, and a background noise spectrogram from the target location are passed through a diffusion model to generate augmented spectrogram images. The image masks are used to ensure the diffusion process only occurs around the inserted call so localized style transfer can take place and the rest of the image is unchanged. An adaptive instance normalization conditional control uses the target background noise spectrogram to drive the diffusion process toward generating image content like the target spectrogram.

To evaluate the effectiveness of the augmented data, a convolutional neural network (CNN) was trained to classify spectrogram data as either containing a fin whale call or as background noise. The model was trained using mixtures of hand labeled and augmented spectrograms and subsequently tested on a holdout dataset of fin whale calls from the target recording location. Initial results confirmed the augmented data increased the ability of the fin whale call classification model to generalize in new domains and that equivalent

levels of classification performance can be achieved with less hand labeled data when augmented data is used to supplement the training set.

A comparison of methods to estimate ranges of fin whale calls using seismic data

A. Pereira[1*], R. Hilmo[2], W. Wilcock[2], D. Harris[3], L. Matias[1]

[1]*Instituto Dom Luiz, University of Lisbon, Lisbon, Portugal*

[2]*School of Oceanography, University of Washington, Washington, USA*

[3]*Centre for Research into Ecological and Environmental Modelling, University of St. Andrews, St. Andrews, Scotland, UK*

**afpereira@fc.ul.pt*

Ranging i.e., determining the range (distance) to a calling animal, is a fundamental element of many density estimation methods. Several range estimation approaches have been applied to ocean-bottom seismometer (OBS) data, including multipath arrivals at a single OBS, using multiple OBSs with time difference of arrival methods and using particle velocities at a single OBS. While prior studies have individually demonstrated the use of these methods, a comparative analysis is still needed to understand their strengths and limitations across different contexts of instrument configurations and acoustic propagation conditions. Adressing this gap of knowledge, the “Combining global OBS and CTBTO recordings to estimate abundance and density of fin and blue whales”, or CORTADO project, undertakes the task of performing such comparison using a suite of case studies from OBS deployments around the world. In this study, we specifically focus on comparing the ranges of twelve fin whale tracks recorded in six deployment areas - three in the Pacific Ocean (Marianas trench, Hawaii, and Oregon OOI) and three in the North Atlantic Ocean (ENAM, Azores and Gulf of Cadiz). The ranges of the 20-Hz fin whale calls that composed the tracks were estimated using single station multipath arrivals and particle velocities and were averaged into 1-min bins. In general, the multipath method ranged further than the particle velocity method, reaching distances exceeding 15 km. The particle velocities method was restricted by a validity range, linked to OBS depth and propagation properties. Furthermore, validation of the obtained ranges for this method required a classification process (see Pereira et al. 2024 abstract). The multipath method worked best with seafloor properties where multipaths were strong and arrivals could be identified. The particle velocities method was most effective in deep and softly sedimented areas. Both methods were affected by large bathymetric relief and complex calling behaviour and chorusing. The average difference between the two ranging methods was between 220 – 250 m [standard error: 146 m]. The results of CORTADO contribute significantly in making informed decisions regarding the selection and application of ranging methods in future marine mammal density estimation studies.

Using acoustic spatial capture recapture to estimate call density of Bowhead whales when many detections are false positives

F.T. Petersma[1*], L. Thomas[1], A.M. Thode[2], D. Harris[1], G.V. Cheoo[3], K.H. Kim[4]

[1]*University of St Andrews, St Andrews, Scotland*

[2]*Scripps Institution of Oceanography, San Diego, California, USA*

[3]*Centro de Estatística e Aplicações, Departamento de Biologia Animal, Faculdade de Ciências da Universidade de Lisboa, Lisboa, Portugal*

[4]*Greeneridge Sciences, Santa Barbara, California, USA*

**ftp@st-andrews.ac.uk*

Acoustic spatial capture recapture (ASCR) is a method for estimating the spatial density of animal calls (or other vocalizations) across an array of sensors, accounting for the fact that some calls may be missed (false negatives). The basic method requires only that we can distinguish which calls were detected on which sensors, and that at least some of the calls are detected on multiple sensors. Extended versions use information about received level (RL, or signal-to-noise ratio, SNR), time of arrival, and/or bearing to provide more precise inference. One common feature of acoustic data is false positive detections—i.e., non-target sounds that are classified as being calls from the target species. These can strongly bias ASCR estimates—both because they artificially inflate the count of detections and because, if they are quieter than true animal calls, they make it seem like calls are harder to detect than they really are, inflating the estimate of missed calls. We demonstrate this with a case study on the Bering–Chukchi–Beaufort population of bowhead whales (Balaena mysticetus) detected from an array of 6 Directional Autonomous Seafloor Acoustic Recorders (DASARs) deployed off the north coast of Alaska in two lines at 7km spacing. The automated detection and classification system developed to handle the large dataset resulted in a high proportion of calls registered at only a single recorder ("singletons"), which is biologically implausible. We developed a novel extension to ASCR that is robust to quiet false positives by truncating singletons and conditioning on calls being detected by at least two sensors. In addition, we allow for variable sound source level, non-uniform spatial call density, and bearing measurement error. We verify performance through a simulation study and demonstrate the method using a single day of data, 31 August 2010. We estimate that 40% of singletons were false positives. The estimated number of true calls within a 50 km buffer of the sensor array was 5741 (coefficient of variation 11%); these were strongly non-uniformly distributed, with high density in a band approximately 45 km from the coast, consistent with the whales' known migratory behaviour at that time of year. Our model was based on discarding calls below a specified RL; we introduce an alternative based on SNR that potentially allows quieter calls to be included at the cost of requiring measurements of noise on sensors at times when calls were missed.

Data Management for Detection, Classification, and Localization

M.A. Roch[1*], S. Baumann-Pickering[2], K. Frasier[2], D. Gillespie[3], R. Whitney[1]

[1]*Department of Computer Science, San Diego State University*

[2]*Scripps Institution of Oceanography, University of California San Diego*

[3]*Sea Mammal Research Unit, University of Saint Andrews*

**marie.roch@sdsu.edu*

Over the last few decades, passive acoustic monitoring (PAM) has become a common tool for monitoring wildlife. In this talk, we discuss the components of Tethys 3, a scientific workbench for standardizing and manipulating information related to detection, classification, and localization (DCL) efforts within PAM studies. Tethys has been used by academic, government, and non-governmental organizations to archive millions of acoustic events, enabling research covering decades of acoustic effort on a wide variety of marine mammals, and its characterization of deployment data has been adopted by the US National Center for Environmental Information PAM archive. A key component of this work is the ability to capture detailed information about DCL processes and instrumentation in a standardized manner while also supporting novel measurements when needed. Documenting these processes lets researchers retain detailed information about what has been done, how it was done, which analyses can be combined in a meta study, and as importantly, which analyses should not be combined.

Tethys consists of a database, web server, and client libraries that can access and manipulate DCL and instrument data. Tethys is distributed with a sample database of hundreds of thousands of detections contributed by academic and government labs, including call annotations derived from the DCLDE 2022 data set.

Tethys can ingest data from various sources. Graphical web-based facilities establish correspondence between fields of spreadsheets, databases, or comma-separated value files. This is linked to a standardized representation with provisions for incorporating ad-hoc measurements that fall outside the set of commonly collected data types. Libraries also permit programmatic generation of Tethys data, enabling programs such as PAMGuard to produce Tethys-ready outputs of DCL data.

Users access Tethys data either from a web browser or programmatically. Drag-and-drop web interfaces permit the construction of detailed queries without knowledge of the query language. A point-and-click map interface enables spatial and temporal data exploration and examination of diel, lunar, and crepuscular cycles. Programmatic interfaces are provided for users performing large-scale statistical analyses, enabling data import into various programming languages commonly used for analysis.

Beyond performance - advanced techniques and lessons learned from training a neural network to classify visual representations of beaked whale echolocation clicks

T. Sakai[1*]

[1]*IBSS Corp, contracted to Northeast Fisheries Science Center, NOAA*

**taiki.sakai@noaa.gov*

Beaked whales are a family of deep diving cetaceans that are difficult to survey with traditional visual survey methods, but are ideal candidates for studying with passive acoustics. The echolocation clicks made during their foraging dives are often highly stereotyped and can be classified to species by trained acoustic analysts, and the Wigner-Ville transform of the click is often used as a visual tool to differentiate between different species.

We use a dataset of 200,000 clicks from over 800 distinct beaked whale encounters spanning five species to train a convolutional neural network (CNN) classifier based on the ResNet architecture, using transfer learning to incorporate knowledge from pre-training on the vast ImageNet dataset. The model achieves 80% overall accuracy in classifying individual click detections, which results in over 99% accuracy in classifying the beaked whale encounters (encounter predictions are made by averaging the predictions of all clicks in an encounter). Results for each individual species match or exceed the performance of existing random forest based models trained on this same dataset.

Neural networks are often heralded for their strong performance, but there are other benefits to this style of model that are not often discussed. CNNs offer a high level of flexibility for incorporating new features that may not be possible with other model architectures. In this talk we will show how ancillary measures like inter-click interval can be incorporated into the model, and how they improved results. We also discuss how we simultaneously trained a selection model based on SelectiveNet that allows the model to avoid making predictions for images it is not confident about. Wigner images are frequently quite noisy, and many images are of a low quality that would not be used by a human analyst for classification. The selection model allows the CNN to “pass” on predicting some percentage of inputs rather than forcing it to make predictions on these noisy, low-quality images. Finally, one downside of CNNs is that they lack interpretability compared to other models, so we show how saliency maps can be used to better interpret our results and glean useful insights into how the model is making its decisions.

Bearing method for density estimation: a comparison of performance using CTBTO data at Wake Island and Diego Garcia

K.D. Seger[1*], D.K. Mellinger[2,3], J.L. Miksis-Olds[4], J.K. Boyle[1], D. Harris[5]

[1]*Applied Ocean Sciences*

[2]*Cooperative Institute for Marine Ecosystem and Resources Studies and Marine Mammal Institute, Oregon State University*

[3]*NOAA Pacific Marine Environmental Laboratory*

[4]*CARE, University of New Hampshire*

[5]*Centre for Research into Ecological and Environmental Modelling, University of St. Andrews*

**kerri.seger@appliedoceansciences.com*

The "COmbining global OBS and CTBTO Recordings To estimate Abundance and Density Of fin and blue whales" (CORTADO) project is using data from two types of seafloor-deployed acoustic sensors to implement and compare performance of a suite of methods for estimating density of fin and blue whales. Part of the project is focused on processing Ocean Bottom Seismometer (OBS) data while another part is focused on processing Comprehensive Test Ban Treaty Organization (CTBTO) hydroacoustic data. The primary CORTADO goal is to provide a set of software tools and training materials to expand access and usability of large, historic datasets. It is important, though, to first determine which density estimation methods are most functional at certain locations, on certain instruments, and for which species, so users can make informed choices about which method is most appropriate for their data. As part of this large-scale evaluation, the "bearing method" is being applied to the CTBTO hydroacoustic data for fin whales. This technique uses the bearing estimate and signal-to-noise ratio for each detected fin whale downsweep to estimate how many animals are producing downsweeps at any given time over the detection range of the sensor. Data from the Wake Island and Diego Garcia CTBTO stations from 2008 and 2011 were evaluated to compare the bearing method's utility in three different ways: (1) north vs south triads at a CTBTO station, (2) month-to-month around peak acoustical presence times, and (3) propagation environments in different ocean basins. This talk focuses on the application of the bearing method, addresses its application to CTBTO sensors, and discusses its relative utility across space and time in CTBTO datasets.

Neural Network-Based Detection and Classification for North Atlantic Right Whale Upcalls: Performance, Deployment, and Generalization to Unseen Environments

J. Tatarowicz[1*], B. Yue[1], W. Dong[1]

[1]*ThayerMahan, Inc.*

**jtatarowicz@thayermahan.com*

North Atlantic right whale (NARW) populations have been in decline since the early 2010s, highlighting the need for robust detection models for conservation. Here, ThayerMahan presents the results of our neural network-based classifiers for NARW upcalls, generalization of those models to new environments, and deployment methodologies to edge computing devices on our passive acoustic monitoring (PAM) systems. We utilized transfer learning on convolutional neural networks with the data from the both the DCLDE St. Andrew's and St. Lawrence's NARW datasets. After cleaning label noise in each dataset, we present our detection results at a targeted rate of 1.3 false alarms per day. This model was deployed in a monitoring effort in regions around Martha's Vineyard, implemented on single board computers as a part of the payload on our PAM systems. Results of those initial deployments in noisier, higher vessel traffic areas are presented here, as well as our methodology for fine tuning our models based on the data collected during deployment. Finally, we show the results of these environmentally fine-tuned models on the DCLDE St. Andrew's and St. Lawrence's NARW datasets and the re-analysis of the deployment data from Martha's Vineyard.

Passive acoustic tracking of Rice's whales in the northeastern Gulf of Mexico using a wide-baseline array

L. Tenorio-Hallé[1*], A. Cook[2], A. Debich[3], K. Frasier[4], H. Frouin-Mouy[3], P. Gruden[6], J. Hildebrand[7], E. Nosal[6], M. Soldevilla[3]

[1]*NOAA Southeast Fisheries Science Center, US*

[2]*Rosenstiel School of Marine, Atmospheric, and Earth Science, University of Miami, US*

[3]*NOAA Southeast Fisheries Science Center*

[4]*Scripps Institution of Oceanography, University of California San Diego*

[5]*Cooperative Institute For Marine And Atmospheric Studies (CIMAS), University of Miami*

[6]*Ocean and Resources Engineering, University of Hawai`i at Manoa*

[7]*Scripps Institution of Oceanography, UCSD*

**ludovic.tenorio@noaa.gov*

The endangered Rice's whale (Balaenoptera ricei) is endemic to the Gulf of Mexico where it is also the only resident baleen whale. Its low abundance, currently estimated to be less than 100 individuals, combined with the many potential threats in this highly industrialized region, raises conservation concerns. Furthermore, our limited knowledge of this rare species urges for more extensive monitoring and the development of new methods for characterizing its ecology, behavior and trends. As part of a project to assess the spatiotemporal distribution of Rice's whales along the continental shelf in the northeastern Gulf of Mexico, a wide-baseline array of 18 passive acoustic moorings was deployed from 2021 to 2023. Using this dataset, a semi-automated "track-before-detect" framework was developed to track individual whales. This method uses long-term cross-correlations between sensor pairs (referred to here as "cross-correlograms") to identify tracks in the time-difference-of-arrival (TDOA) domain before associating those from the same whale across sensor pairs. Standard TDOA localization can then be used to obtain tracks in the spatial domain (latitude and longitude) from the TDOA tracks. This framework is particularly well-suited to wide-baseline array datasets in which (1) the detection range of the source relative to the array spacing poses detectability issues and often limits the localization to three sensors, and (2) multiple acoustically-active whales are producing similar calls, making the association problem easier to perform using TDOA tracks than with individual calls. The approach developed here facilitates analyzing large datasets to better characterize the behavior of this endangered species and makes progress towards density estimation of baleen whales.

Localizing North Atlantic Right Whales Using a Deformable Sonobuoy Grid

K.P.D. Thebeau[1*], D. Barclay[1], C. Binder[2]

[1]*Dalhousie University, Halifax, Canada*

[2]*Defence Research and Development Canada Atlantic, Dartmouth, Canada*

**kamden.thebeau@dal.ca*

In 2018, a large-scale data collection effort was conducted in the Gulf of St. Lawrence, a known feeding ground for North Atlantic Right Whales (NARW), over two days. On each day, visual surveys were conducted, 32 sonobuoys were deployed to gather directional acoustic time series, and a Slocum glider operated in the area to collect oceanographic data. Following the collection phase, the acoustic data were manually annotated with a focus on NARW vocalizations. This project uses the multi-modal dataset to test and compare the results of three localization algorithms for NARW calls. The NARW calls were correlated with the acoustic files to determine the same occurrence across several sonobuoys. These matched filter results were manually verified using the properties of the call and surrounding signals. The first method of localization used the directionality of the calls, where probability density maps were created by overlapping the bearing statistics across each of the relevant sonobuoys, commonly referred to as cross-fixing. To improve the bearing distributions, the signals were isolated with thresholding after using a conditional whitener and power-law statistic on its short-time Fourier transform. The direction itself was calculated using an arctangent bearing estimator. Range dependence was also incorporated with the bearing maps by using probabilistic transmission loss, where the transmission loss was calculated using in situ data for sound speed profiles and a range of bottom parameters for the Gulf of St. Lawrence. This transmission loss was used with the received levels on the hydrophones to get a source level range which could then be compared to values from literature. In the next approach, localization was performed using a spherical interpolation method to initialize a maximum likelihood time difference of arrival algorithm. Finally, matched-field processing was used to model replica fields at potential source locations and correlate the replicas with the received pressure on the hydrophones.

Not just research: operational use of deep learning models for PAM

M. Thomas[1*], B. Martin[1], S. Matwin[2]

[1]*JASCO Applied Sciences*

[2]*Dalhousie University, Faculty of Computer Science*

**mark.thomas@jasco.com*

Over the past several years there has been a dramatic increase in reporting of deep learning models, particularly convolutional neural networks (CNNs), to classify spectrograms possibly containing marine mammal vocalizations. While the models presented in the literature appear to be quite successful, we suggest that the operational use of some of these models may be slightly overstated. For one, generalizability may suffer due to limited diversity in the datasets used for training and evaluation. And second, framing the process of classifying spectrograms as “detection” yields less interpretable results that may be ineffective for downstream tasks.

This research explores the application of single-stage detection algorithms, alongside CNNs, to examine their effectiveness in detecting and classifying marine mammal vocalizations across various acoustic environments. By conducting a comparative analysis under different noise characteristics and environmental conditions, this work aims to highlight the significance of dataset composition and model choice in achieving more meaningful outcomes. We discuss the potential benefits of detection-based models in enhancing the interpretability of results, such as enabling vocalization counting and localization, as well as performing multi-species detection (at a vocalization level) using a single model. Through our findings, we propose recommendations for best practices in dataset creation to improve model generalizability and utility.

Automatic detection of humpback whale calls: a comparison between a machine-learning convolutional neural network (CNN) detector and the Low-Frequency Detection Classification System (LFDCS)

L.M. Transue[1*], V. Kather[2], G.E. Davis[3], S.M. Van Parijs[3], M. Baumgartner[4]

[1]*Under Contract to the Northeast Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Woods Hole, MA, USA*

[2]*Technical University Berlin, Audio Communication and Technology, Berlin, Germany*

[3]*Northeast Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Woods Hole, MA, USA*

[4]*Woods Hole Oceanographic Institution*

**lindsey.transue@noaa.gov*

Long-term passive acoustic monitoring of baleen whales has become increasingly possible with the development of automatic detection systems that reduce manual analysis requirements. Detection technologies continue to emerge and improve our capabilities to reduce time-intensive analysis. Humpback whales present a challenge to any automated detector due to the large variety of call types they produce, which change within and across years. Two different types of detectors (non-machine learning and machine learning based) were compared for performance and analysis speed of North Atlantic humpback whale acoustic presence. The non-machine learning based Low-Frequency Detection Classification System (LFDCS) uses pitch tracks to identify and classify possible baleen whale call types by calculating a spectrogram using the short-time Fourier transform (STFT) and tracing contour lines through the dominant frequencies of tonal sounds over time. The attributes of each resulting pitch track are then compared to those in a custom-created call library and classified into call types using quadratic discriminant function analysis. The second detector, a modified ResNet-50 convolutional neural network (CNN, machine learning based), also works with spectrogram data and conducts binary image classification trained from manually annotated spectrograms of humpback whale vocalizations. The detector was originally trained on humpback whale song data from the North Pacific and later fine-tuned on North Atlantic data using the “AcoDet” (acoustic detector) framework. This study compared the daily detection and missed detection rates of humpback whale presence between LFDCS and the CNN detector at four bottom-mounted recording sites from June 2022 to May 2023. One site was selected from each of the following regions: offshore Gulf of Maine, Stellwagen Bank National Marine Sanctuary, Southern New England, and offshore Chesapeake Bay. This is the first study to evaluate the performance of the CNN detector against LFDCS, and the results will help compare the effectiveness of the two systems for detecting humpback whale vocalizations in the western North Atlantic.

Estimating the number of animals, animal tracks and the motion parameters of vocalizing marine mammals using compact hydrophone arrays

I. Urazghildiiev[1*], B. Martin[2], J. Miksis-Olds[3]

[1]*JASCO Applied Sciences (USA) Inc., Silver Spring, MD, USA*

[2]*JASCO Applied Sciences (Canada) Ltd., Dartmouth, Canada*

[3]*Center for Acoustics Research and Education, University of New Hampshire, USA*

**ildar.urazghildiiev@jasco.com*

Estimating the number of vocalizing animals, as well as tracking and estimating their motion parameters is a challenge. The proposed solution uses a fixed underwater compact array providing the ability to estimate the angle-of-arrival (AOA), sound pressure level (SPL), and signal-to-noise ratio (SNR) of each detected signal. The processing technique consists of several stages: estimating the distances to the sources; estimating the number of sources; estimating the source positions and motion parameters. The distance is estimated from the propagation loss calculated using the Range-dependent Acoustic Model (RAM). The number of sources and minimum animal density are assessed using clustering of measured AOAs and SPLs. Source positions and motion parameters are evaluated using a polynomial approximation of the measured AOA and distances.

The technique was evaluated using data collected in the Gulf of Maine using four-channel cartesian arrays deployed by the Atlantic Ecosystem Observation Network (AEON) project. The proposed technique was tested on calls from sei (Balaenoptera borealis), fin (Balaenoptera physalus), and humpback whales (Megaptera novaeangliae). The advantages and disadvantages of the technique are demonstrated. The value of data harvested with this technique will support new research related to inferred migration dynamics and behavior.

Bayesian localisation of NARW using an autonomous field of sonobuoys

A.M. von Benda-Beckmann[1*]

[1]*TNO, Acoustic Sensor and Sonar Systems, The Hague, The Netherlands*

**sander.vonbendabeckmann@tno.nl*

Large acoustic networks allow for improved localisation of marine mammals. Autonomous acoustic networks are readily increasing in size and use, and allow to monitor presence of marine mammals over larger areas. The DCLDE2024 consists of a network of drifting sonobuoys aimed to detect and localise North Atlantic Right Whale (NARW) calls. The drift of the autonomous recorders can lead to substantial uncertainty in receiver locations, which translates into uncertainty in estimated location of the whales. The DCLDE2024 dataset was used to investigate the effect of position uncertainty on localisation accuracy using a Bayesian framework. Location estimates were obtained from both bearing estimation and time-delay-of arrival (TDOA) between sonobuoys. A Hamiltonian Monte Carlo (HMC) solver was used to estimate source location and receiver location. We investigated whether targets of opportunity can be used to improve localisation of NARW calls.

Diving into Deep Learning to find Risso's dolphins Echolocation Clicks in Scottish Waters.

T. Webber[1*], N.C.F van Geel[1], D. Risch[1], G. Hastie[2], R. Sinclair[3], L. Hartny-Mills[4], C. Lacey[5], S. Quer[6], S. Beck[7]

[1]*Scottish Association of Marine Science, Oban, PA37 1QA, Scotland*

[2]*Sea Mammal Research Unit, University of St Andrews, St Andrews, KY16 8LB, Scotland*

[3]*NatureScot, Inverness, IV3 8NW, Scotland*

[4]*Hebridean Whale and Dolphin Trust, Tobermory, Isle of Mull, PA75 6NQ, Scotland*

[5]*Marine Mammal Research Program, Hawaii Institute of Marine Biology, HI 96744, USA*

[6]*Marine Directorate, Scottish Government, Edinburgh, EH1 3DG, Scotland*

[7]*Agri-Food and Biosciences Institute, Government of Northern Ireland, Belfast, BT9 5PX, Northern Ireland*

**thomas.webber@sams.ac.uk*

Reliable species classification is important to be able to conduct species-specific monitoring, particularly for the purposes of long-term conservation management. Risso's dolphins are a designated feature of the Scottish Marine Protected Area (MPA)network, but currently there is limited knowledge regarding their year-round occurrence along the Scottish west coast. Designing effective monitoring and management measures is therefore made even more difficult. Data exists from a network of static acoustic recorders in coastal waters off western Scotland (COMPASS; https://compass-oceanscience.eu/) which could help to improve our understanding of Risso's dolphin occurrence in the region, including the MPA specifically designated for the species. However, manual annotation of these large datasets is extremely labour intensive. Deep learning offers the possibility of decreasing processing time while still achieving accurate classification. A long short term memory (LSTM) deep learning classifier, designed to identify Risso's dolphin echolocation clicks, was trained from visually verified recordings of dolphin species in UK waters collected using towed hydrophones (data provided by HWDT and SCANS III). Manually annotated click trains provided two datasets which were used to train the model (Train and Validation). An additional dataset (Test) was used to test the ability of the model to generalise to previously unseen data. Results show good performance of the classifier to the validation data at both the individual click level (89.5% accuracy: 89.8% precision; 88.8% recall) and the event level (98.1% accuracy; 97.7% precision, 95.5% recall). Preliminary results of the test dataset also show promising results at both the individual click level and the wider event level. Whilst still being optimised, the final model will be applied to the COMPASS network to provide outputs which can be

incorporated into distribution modelling frameworks, eventually providing regulators with evidence to aid management decisions.

Can one model do it all? Exploring the application of multi-sound source detection algorithms to new marine soundscapes.

E.L. White[1*], P.R. White[2], J.M. Bull[3], R. Tandon[4], D. Risch[5]

[1]*School of Ocean and Earth Science, University of Southampton*

[2]*Institute of Sound and Vibration, University of Southampton*

[3]*School of Ocean and Earth Sciences, University of Southampton*

[4]*RS Aqua*

[5]*Scottish Association of Marine Science*

**elw1d23@soton.ac.uk*

Effective analysis of marine acoustic ‘Big Data’ could permit the analysis of ecosystem health and species presence, if automated detection and classification algorithms are capable of discrimination between species of interest and the presence of anthropogenic and environmental noise. The rapid expansion of the field of deep learning has the potential to enable a revolution in bioacoustics, allowing for the robust detection and classification when we want to account for multiple sound sources in the marine environment.

In this work, algorithms from the field of self-supervised learning are explored for multi-class marine sound source detection, spanning anthropogenic, geophonic and biological signals of interest. We demonstrate this approach with field data collected in the West of Scotland. This presentation will update the community with ongoing experiments of using a new approach to multi-class signal detection from broadband acoustic data. The approach is compared to existing methods, including a small scale convolutional neural network (CNN) developed for a similar task, but limited to broad classes with high levels of signal variation between them, and traditional signal processing pipelines.

As global Passive Acoustic Monitoring (PAM) datasets continue to expand it is critical we improve our confidence in the performance of models across different marine environments, if we are to exploit the full ecological value of information within the data. We explore the transferability of developed models to new acoustic environments by using a pre-trained model developed for one location (West of Scotland, UK) and deploying it in a distinctly different soundscape (Gulf of Mexico, USA). The model is fine-tuned on training sets of differing sizes, from the unseen site, to understand the adaptability of a network to new marine acoustic environments. Fine-tuning with a small sample of site-specific data significantly improves the performance of a CNN in the new environment, across all classes. We demonstrate an improved performance in area-under-curve (AUC) score of 0.30, across four classes by fine-training with only 50 spectrograms per class, with a 5% improvement in accuracy between 50 frames and 500 frames. We find that only a small amount of site-specific data is needed to retrain a CNN, and researchers can use only ambient data to

achieve this, minimising the amount of manual labelling required to disseminate the soundscape of a new environment.

A generalized deep-learning approach for difficult signal detection challenges in passive acoustic datasets

D.F. Woodrich[1*], A.N. Allen[2], H. Frouin-Mouy[3,4], C.L. Berchok[5]

[1]*Cooperative Institute for Climate, Ocean, & Ecosystem Studies (CICOES), University of Washington, USA*

[2]*Pacific Islands Fisheries Science Center, USA*

[3]*Cooperative Institute For Marine And Atmospheric Studies (CIMAS), University of Miami*

[4]*(previous) JASCO Applied Sciences, Victoria, Canada*

[5]*Alaska Fisheries Science Center, USA*

**daniel.woodrich@noaa.gov*

The Alaskan Arctic is one of the most complex marine soundscapes in the world. Signal detection for purposes of monitoring species occurrence requires distinguishing signals which can both mask and share characteristics with one other (e.g., humpback and bowhead whales, bearded seals and bowhead whales). In addition, identifying specific patterns within repertoire from complex signals such as humpback song require robust techniques. To achieve signal detection in these challenging problems, a generalized approach was developed in TensorFlow. This featured a novel, specialized augmentation policy, dynamic on-the-fly spectrogram slicing during training, and a deep-learning architecture (EfficientNetB1) to perform sliding-window binary classification. This model training routine was integrated into the INSTINCT framework, which allowed for model parameterization, customized sampling strategies for training and evaluation, and integrated human-in-the-loop verification during rollout. This approach will be demonstrated in two case studies: one distinguishing bearded seal calls from bowhead calling and ice noise, and the other distinguishing among humpback song phrases and those from bowhead whales. (1) Starting with a ~500 hour hand-labeled dataset provided by JASCO Applied Sciences, we trained a bearded seal detection model, and applied it to three year-long mooring deployments in our dataset in an active learning strategy. After reviewing and incorporating high likelihood positive and negative label outputs (including instances of co-occurrence with bowhead song and ice noise), we applied the model to two year-long mooring deployments in our dataset, achieving an accuracy of 96% and 98%, respectively. This result was achieved in a blind test with no sensitivity calibration to the manual analysis results. Further improvements and continued application of this dataset from both the technical and species monitoring perspective will be presented and discussed. (2) Humpback song from 2015-2016 was sampled on recorders in Hawaii and Alaska, analyzed for reoccurring phrase types, and annotated for presence of eight yearly humpback phrases. These phrases were used to train phrase-specific detection models. We will demonstrate the use of active learning to improve label quality in an initial round, distinguish the phrase from similar bowhead and

humpback calls in a second round, and present initial results of model inference on the larger Alaska dataset.

Signal Processing Considerations for use of compact Volumetric Acoustic Sensors

W.M.X. Zimmer[1*]

[1]*NATO-STO CMRE, I-19126 La Spezia, Italy*

**wzimmer@wmxz.eu*

Three-dimensional localization of sound sources requires large scale volumetric distribution of acoustic sensors. If three-dimensional direction finding is satisfactory, the scale of the volumetric sensor distribution can be reduced significantly. Typical implementations feature 4 sensors in tetrahedral configurations with sensor-to-sensor spacing in the order of 1 m. Direction finding of such systems is accomplished by estimating the time delay of arriving sound between the different pairs of sensors in the tetrahedron. As 1 m sensor spacing results still in bulky sensor array dimensions, a compact Volumetric Acoustic Sensor has been developed characterized by sensor spacing in the order of 7 cm. This presentation describes the resulting system and the rational of its design and discusses the advantages such compact system in terms of signal processing and data analysis. Different use cases and results obtained are presented and discussed. Emphasis is paid to the individual Detection-Classification-Direction finding components of Real Time, i.e. in-situ, signal processing.

Abstracts for Poster/ Speed Talks

Geographical and seasonal occurrence of minke whale boings in the central and western tropical North Pacific

A.N. Allen[1*], M. Harvey[2], E.M. Oleson[1]

[1]*NOAA Pacific Islands Fisheries Science Center, Honolulu, HI, US*

[2]*Google, Inc., Mountain view, CA, US*

**ann.allen@noaa.gov*

Minke whales in the North Pacific are widely distributed, ranging from tropical to polar waters, and are generally thought to engage in an annual migration between high and low latitudes. They typically occur in small groups, with short surfacings and inconspicuous blows, making them difficult to study using visual survey methods. North Pacific minke whales produce a unique vocalization called a 'boing', which has been used to track their presence, behavior, and even estimate calling density. Most studies thus far have focused on short term detection and tracking of boings. In the tropical North Pacific, the Pacific Islands Passive Acoustic Network (PIPAN) is a network of bottom mounted High-frequency Acoustic Recording Packages (HARPs) spread over a variety locations, including Hawaii, Wake, Palmyra, and the Mariana Archipelago. The HARPs are deployed for a period of several months to a year, often with duty cycled recording periods. Several sites have had recorders deployed almost continuously for the last ten to fifteen years. We used a multi-class EfficientNetB0 machine learning model to identify boings in the entirety of our PIPAN dataset, and quantified the model performance at each of the recording sites. We then used the output of the model to establish spatial occurrence of boings across the PIPAN and identify seasonal patterns at sites with regular presence. Our results improve our understanding of the acoustic ecology and behavior of minke whales in the central and western tropical North Pacific.

Calling Behavior and Localization of Blue Whales in Southern California

G.M. Arrieta[1*], S. Wiggins[1], K.K. Cohen[1], B.J. Thayre[1], A. Širović[2], S. Baumann-Pickering[1]

[1]*Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA 92037, USA*

[2]*Department of Biology, Norwegian University of Science and Technology, 7491 Trondheim, Norway*

**garrieta@ucsd.edu*

Blue whales, Balaenoptera musculus, are a protected species and with a changing ocean the continual monitoring of these whales has become increasingly important. Blue whales produce low frequency sounds including short (1-4 s) down-sweeping D calls, commonly associated with foraging. This study used long term passive acoustic monitoring to investigate a new sequence of D calls observed in data collected between 2007 and 2020 in Southern California. D calls were detected using a combination of automated and manual methods in which a detector was run on the data and then verified by a skilled analyst. Blue whale D calls are often variable in time and in frequency, therefore the presence of sequential D calls has not been documented. Sequences were present in all years of this time series, but the highest number of sequences per year and per day was in 2019. Yearly, seasonal, and diel trends of all D calls and D call sequences were investigated to understand the occurrence of this newly reported call type. In addition, 28 D call sequences were characterized in detail and localized from two hours of data to understand their patterning and spatial distribution. The D calls during this time had an average 7.3 s +/- 1.7 s inter-call interval within a sequence and 7 calls +/- 1 call per sequence. The localization resulted in 194 locations for all D calls and an average location for each sequence. Average locations were scattered within a 250 m by 350 m area, indicating that an individual whale may be producing the sequences. Using the location for each call, source level was computed using received level and transmission loss. The average root mean square source level calculated over 30 to 80 Hz was 162.3 +/- 2.2 dB re 1µPa at 1 m. By combining both observations of calling behavior and localizations, we can gather insight into the temporal and spatial behavior of blue whales when emitting D call sequences. Understanding the occurrence of D calls is integral for getting accurate density estimation numbers for blue whale populations all over the world.

The Application of a North Atlantic-wide Minke Whale Detector to a Large-Scale Recording Array on the West Coast of Scotland

T. Awbery[1*], X. Mouy[2], G. Davis[3], O. Fernandez-Betelu[4], V. Iorio-Merlo[4], L. Transue[2], S. Van Parijs[3], D. Risch[1], B. Wilson[1], J. Onoufriou[5], L. Hartny-Mills[6], L. McWhinnie[7]

[1]*Scottish Association for Marine Science, Oban, Argyll PA37 1QA, Scotland*

[2]*Integrated Statistics, Inc., under contract to National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA 02543, United States of America*

[3]*National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Northeast Fisheries Science Center, Woods Hole, MA 02543, United States of America*

[4]*School of Biological Sciences, Lighthouse Field Station, University of Aberdeen, Cromarty, Ross-shire IV11 8YJ, Scotland*

[5]*Scottish Government, Marine Lab, 275 Victoria Road, Aberdeen, Scotland*

[6]*Hebridean Whale and Dolphin Trust, Isle of Mull, Scotland*

[7]*Institute of Life and Earth Sciences, Heriot-Watt University, Scotland*

**tim.awbery@sams.ac.uk*

Increasing capacity of acoustic recorders has resulted in a rapid rise in acoustic data being collected, making it impractical to manually review them. Automated detection and classification approaches can provide a solution to this but require extensive testing on unseen data (i.e. those not involved in training) to provide robust results. This is important when studying long-ranging species such as minke whales which vocalise across different soundscapes during migration and the generalisation of a detector across different habitats within a species range is an important consideration.

In this study, a deep learning detector was developed using 4,708 calls from 47 recorders across the North Atlantic spanning a 16-year period (details in Mouy et al., this conference). The detector was tested on data from the SAMOSAS array, a network of 12 RT-SYS broadband recorders to the west of Scotland, UK with recorders at varying depths and in different soundscapes. Data were collected between September 2020 and August 2021 with fifty-nine days of data across four recorders used in the training of the detector and the remaining 2630 days (97.8%) remained unseen. The entire acoustic dataset was run through the detector with potential detections validated manually to assess its performance, and to investigate minke whale seasonal patterns in the region.

The detector returned 4301 detections of which 777 and 3356 were validated as true and false positives respectively, with processing time reduced approximately 95%. False positives were generally found between October and March (83.3%) with preliminary quanlitative

analysis suggesting they were predominantly triggered by environmental, mooring, and vessel noise. Minke whale calls were detected throughout the year, excluding January and February with a considerable peak in October and November. The autumn peak in minke whale vocalisations and a clear diurnal pattern in calls corresponded with results found in an earlier manual validation of a subset of the array. It was less successful in detecting a second peak in spring found during manual validation, the reason for which is currently under investigation.

Nevertheless, using the detector over such a large scale developed a greater understanding of how vocalising minke whales use western Scottish waters. This study has demonstrated the detector's ability to reveal previously unknown vocalisation patterns and to replicate to an extent, findings from manual validation. Inclusion of a greater variety of calls and background noise during retraining is likely to increase the true positive rate in the future.

Comparing F-POD delphinid and porpoise click detections with ground truth manual annotation

J. Beesau[1*], M. Torterotot[1], C. Gicquel[2], F. Samaran[1]

[1]*Lab-STICC UMR 6285, ENSTA Bretagne, France*

[2]*French biodiversity agency, Iroise marine park, France*

**julie.beesau@ensta-bretagne.fr*

Long-term passive acoustic monitoring (PAM) of cetacean populations is often challenged by the limited power autonomy or storage capacity of broadband recorders. This is particularly the case with small delphinids and / or porpoises monitoring, as their clicks can reach very high frequencies (up to 200 kHz for porpoises). The F-POD is an echolocation-click logging device developed by Chelonia Limited and commonly used to monitor dolphins and porpoises in PAM projects. It overcomes the broadband recorders' autonomy restrictions by not storing the raw data and can therefore be deployed for longer periods than a broadband recorder programmed to monitor odontocete clicks. However, the downside is that once the click detections are extracted from the F-POD, there is no way of double-checking the detector performances.

For one year (May 2022 - May 2023), seven F-PODs were paired with broadband recorders (Sylence, RTSys) on moorings deployed in the Iroise sea for the CETIROISE project (funded under the French economic recovery plan), which aimed at monitoring marine mammal presence within the Iroise Marine Natural Park. The broadband recorders were sampling at Fs = 128kHz, allowing dolphin clicks to be recorded. F-POD delphinid clicks detections were manually compared with broadband spectrograms, using APLOSE, a web-based annotation platform. For one month, a F-POD was also paired with a Sylence programmed to sample at Fs = 512 kHz, to be able to compute the performances of the F-POD porpoise click detector.

We found that for delphinid clicks, the F-POD has a conservative behaviour, i.e. it misses most of the clicks, but almost all detections are true positives. We will also present the results on this comparison on harbour porpoise clicks and discuss the implications on the interpretation of the F-POD detection results to monitor marine mammals.

Detection and localisation of Sperm Whales using multiple hydrophones on several gliders.

G.A.S. Bekki[1*], J.M. Bull[1], P.R. White[2]

[1]*School of Ocean and Earth Science, University of Southampton, National Oceanography Centre Southampton, SO14 3ZH, UK*

[2]*Institute of Sound and Vibration Research, University of Southampton, Southampton, SO17 1BJ, UK*

**g.bekki@soton.ac.uk*

Most Sperm whale (Physeter macrocephalus) acoustics signals have been recorded using static or towed hydrophone arrays. This project aims to detect and locate Sperm whales using data from underwater gliders that can survey more extensive geographical and depth ranges. Four SeaExplorer gliders, developed by Alseamar, were deployed in the PELAGOS protected sanctuary in the Mediterranean. Firstly, a low computational Sperm whale echolocation click detector will be developed to deploy it directly on the gliders. Then, information about a whale's position will be determined by combining the bearing given by several techniques, such as beamforming and Time Difference Of Arrival (TDOA). By collating this information from each glider, a better estimate of the cetacean's location can be calculated, and insight into Sperm whale distribution and behaviour close to several major canyon systems can be gained.

Phase Locked Loops to track harmonic calls in frequency and space

A. Berg[1*], J. Rychen[1], L. Gies[2]

[1]*ETH Zurich*

[2]*Universität Wien*

**bergar@student.ethz.ch*

Extracting the individual cetacean calls from hydrophone recordings is a crucial step needed to analyze their communication. Using hydrophone arrays facilitates the separation of overlapping calls as they have different arrival times on the hydrophones. We show that phase locked loops allow to track the frequency contour and the spatial location in the same processing step. The harmonic structure is revealed in fine detail by yielding the amplitudes and phases of the higher harmonics. By monitoring the phases at the different hydrophones, we obtain the track of the animal during the call emission. As the call is emitted, the animals typically glide several meters through the water. This movement corresponds to several wavelengths of the fundamental frequency, therefore, the phase on different hydrophones can turn by several rounds. Thus, we can obtain the swim direction for every call. We show some first preliminary results as a proof of concept. We also discuss the challenges arising from echoes or low signal-to-noise ratios, and outline strategies for overcoming the issue of spatial aliasing. We discuss how such signal processing can be applied to infer about emission beam characteristics. Additionally, we show that this method facilitates the separation of overlapping calls, thereby paves the way for an in-depth analysis of cetacean communication.

Analysis of cachalot dialogues, click by click, an ethoacoustical approach

L. Berkenbaum[1,2*], F. Sarano[2], O. Adam[4,2], H. Glotin[3,2]

[1]*Université de Toulon, AMU Univ, CNRS, LIS, DYNI team, Marseille, France*

[2]*CIAN, Centre d'Intelligence Artificielle en Acoustique Naturelle, Toulon, France*

[3]*Longitude 181, Valence, France*

[4]*Sorbonne Université, CNRS, Institut d'Alembert, UMR 7190, Paris, France*

[5]*Université Paris-Saclay, CNRS, Institut des Neurosciences Paris-Saclay, 91190, Gif-sur-Yvette, France*

**lara.berkenbaum@lis-lab.fr*

The vulnerable status of sperm whales (Physeter macrocephalus), particularly threatened by anthropogenic activities, highlights the need for in-depth research. Currently, knowledge of their acoustic emissions during social interactions is incomplete. Our interdisciplinary team, comprising the Centre for Artificial Intelligence in Natural Acoustics with LIS UTLN, the Alembert Institute, Longitude 181, and the Indian Ocean Marine Life Foundation, conducts expeditions to Mauritius in the Indian Ocean to observe the clan of sperm whales known as 'Irène Gueule Tordue' [3].

In 2023, utilizing the innovative OPALE mobile high definition AV antenna, from SMIot and LIS DYNI [2], equipped with five hydrophones and two GoPro cameras, we collected high-definition acoustic datasets accompanied by underwater videos.

Whitehead [1] emphasizes that the social structure of sperm whales is organized into clans, encompassing behaviors and coda dialects, under the assumption of their alignment. However, variations in non-vocal behaviors and other characteristic traits, often undocumented, become apparent.

OPALE enables us to fill in certain missing information. By recording interactions between individuals at close range, it offers unique observations. Through in-depth analysis of audio data, we are able to identify the speaker by establishing a video pointing system to locate the source of each click emitted by the sperm whales.

We are eager to present our preliminary results, exploring interactions among sperm whales, integrating various types of vocalizations emitted beyond codas, while linking them to their social dynamics. For example, we recorded exchanges between two immature males emitting staccatos of clicks. In exploring vocal and non-vocal behaviors within the matrilineal social units of the clan, our subsequent objective is to deepen the understanding of sperm whale society, in line with the observations of Whitehead [1].

This ethoacoustical research marks the beginning of the exploration of interactions between individuals, contributing to the understanding of vocal communications and social behaviors of sperm whales, thereby reinforcing global conservation efforts for this vulnerable species.

[1] Whitehead, H. (2024). Sperm whale clans and human societies. Royal Society Open Science, 11 (231353). https://doi.org/10.1098/rsos.231353

[2] OPALE : a super resolution audiovisual mobile antenna for cetacean ethoacoustical researches, Glotin et al. DCLDE2024

[3] La Voix des Cachalots, Longitude 181. https://www.longitude181.org/programme-cetaces-cachalots/

SERCEL QuietSea, Development and testing of a harbour porpoise detection buoy in Ramsay Sound, Wales

T. Bertet[1*], S. Coatelan[1], A. Vogler[2]

[1]*Sercel, Brest, France*

[2]*ORE Catapult, Glasgow, United Kingdom*

**theo.bertet@sercel.com*

Work presented here will describe the development and real-world testing of a novel passive acoustic monitoring system for the detection of harbour porpoises and other marine mammals. A surface buoy was fitted out with suspended hydrophones, a power management system, and satellite communications to monitor the presence of the target species at the deployment site in Ramsay Sound, Wales. An F-Pod and a high frequency audio recorder device were also integrated into the system to provide comparison of detections against an off-the-shelf system. Data from a second F-Pod device deployed in close proximity was also available for further benchmarking. The test buoy was developed by Sercel and deployed in partnership with ORE Catapult's Marine Energy Engineering Centre of Excellence (MEECE) and CGG for a two-month period at a seaweed aquaculture site in late spring 2023. An overview will be given of the technology and site finding exercise, deployment, in-situ data downloads and recovery. An important aspect of this project was the highly positive interaction with local stakeholders and the community, which has already resulted in the conception and delivery of additional marine energy and green economy related projects. Furthermore, an overview of the capabilities of the system will be presented, including multiple marine mammals detection throughout a large frequency band (5 Hz to 192 kHz) covering the whole cetacean range.

A transfer learning approach for unsupervised whistle categorisation

C. Biermann[1*], J.N. Oswald[2], H. Worthington[3]

[1]*University of St Andrews*

[2]*University of St Andrews, Sea Mammal Research Unit*

[3]*University of St Andrews, Centre for Research into Ecological and Environmental Modelling*

**cewb1@st-andrews.ac.uk*

Classifying and grouping whistles, or click trains, is an important step for exploring delphinid vocal repertoires, for comparing among species and populations, and developing classifiers. Whilst labelled datasets can be utilised for supervised training, unsupervised approaches prove valuable in dividing unlabelled datasets into vocalisation categories. Approaches for categorising whistles by tonal qualities, or click trains by temporal patterns and frequency content, are available but require intensive data processing prior to categorisation. Here we explore unsupervised approaches based on spectrogram images which could act as a preliminary data processing step to partition large datasets into clicks and whistles (if one is of particular interest), to separate potentially complete versus fragmented calls identified from automated whistle detectors/algorithms, or to identify those recording segments containing signals of interest with high signal-to-noise ratios. The aim is to group the spectrogram images based on features of the image, thus classifying based on combinations of tonal qualities of whistles or click patterns, but also the presence of background noise, multiple simultaneous whistles or the presence of harmonics.

To achieve this, we adopt a transfer learning approach applying a pre-trained Convolutional Neural Network (CNN) to a dataset on approximately 2000 spectrogram images depicting whistles of the bottlenose dolphin (Tursiops truncatus). Many public CNNs have demonstrated strong performance across various image classification tasks, the VGG16 (CNN) stands out as one of the top performing. We look to explore how this broadly trained CNN, which has not been trained specifically on spectrogram images, can be adopted for this purpose. Having applied the CNN for feature extraction a Principal Component Analysis (PCA) is used for dimension reduction, transforming the high-dimensional feature space into a lower orthogonal set. K-means clustering on the principal components assigns a cluster label to each image.

The results show some early promising outcomes. Clusters appear to exhibit sensible groupings with a high degree of similarity within clusters. In particular, the clusters appear well suited to grouping images with similar levels of background noise, similar tonal patterns and fragmented whistles.

A preliminary description of Atlantic white-sided dolphin (Lagenorhynchus acutus) vocalisations

S. Calderan[1*], O. Boisseau[2], C. Lacey[3], R. Leaper[4], N. van Geel[1], D. Risch[1]

[1]*Scottish Association for Marine Science (SAMS), Argyll, Scotland, UK*

[2]*Song of the Whale Research Team, Marine Conservation Research, UK*

[3]*Sea Mammal Research Unit, Scottish Oceans Institute, University of St Andrews, UK*

[4]*International Fund for Animal Welfare London, UK*

**Nienke.vanGeel@sams.ac.uk*

Passive acoustic monitoring of marine mammal vocalisations is critical for studying and thus contributing to the conservation of many offshore cetacean populations. Of the nine dolphin species which regularly occur in the North Atlantic, least is known about the acoustic behaviour of the Atlantic white-sided dolphin (Lagenorhynchus acutus). This study aimed to identify, access, and analyse acoustic recordings with visually-confirmed species identification information from different study sites to describe this species' vocal repertoire in detail, including whistles, clicks, and burst pulses.

A total of 69 click events from the eastern (16) and western (53) North Atlantic were selected for analysis of click characteristics. Atlantic white-sided dolphin echolocation clicks were rather narrowband, with a simple waveform structure. More than 72% of events showed more than 50% of the total energy averaged over all clicks within a 16 kHz band (16-31 kHz). The median peak frequency for all the templates combined was 25.4 kHz, with 90% of all click templates demonstrating a peak between 21.5 and 33.2 kHz.

There were no significant differences between clicks recorded in the eastern and western North Atlantic. Apart from differences in minimum frequency, which was lower in the western North Atlantic, there were also no significant differences recorded in measured whistle parameters. Though sample size was small, several whistles showed distinctive shapes, including abrupt frequency steps and breaks in whistle contours. The median fundamental whistle frequencies were found to be between 5.6 and 19.6 kHz. Median whistle duration was between 0.7 and 0.9 seconds.

Only a few burst pulses were present in the existing datasets, preventing robust measurements. Thus, more work is needed to describe this part of the species' vocal repertoire.

The distinctive characteristics of both the clicks and whistles of Atlantic white-sided dolphins suggests that the development of a classification routine based on a combination of all call types would be highly beneficial in distinguishing the species from other species that overlap in time and space using passive acoustics. Atlantic white-sided dolphins have been identified as a suitable indicator species for monitoring climate change in the North Atlantic, due to their behaviour, ecology, and temperature preferences. A better understanding of their current seasonal distribution will thus allow an assessment of future distributional changes and threats to the species due to changing oceanic conditions and human marine expansion and allow the development of effective conservation measures.

Vocal Behavior of Visually-Verified North Atlantic Minke Whales off Jacksonville, Florida, in Winter

A. Carroll[1*]

[1]*Naval Undersea Warfare Center, Division Newport, US*

**alexandra.e.carroll.civ@us.navy.mil*

Performance of automated passive acoustic monitoring (PAM) classifiers rely on the training data utilized. Classifiers can be optimized for a specific location with the addition of locally sourced data that can provide regional variation of the signal(s) to be targeted. On 06 March 2023, two North Atlantic minke whales (Balaenoptera acutorostrata) were acoustically detected and localized in real-time off Jacksonville, Florida, and visually monitored by an on-water field team for 2 hours. Manual Review of the resulting 6.83 hour acoustic data set revealed two animals producing near continuous low-frequency pulse trains with energy within the 50 –200 Hz band. Not all pulse train varieties observed during this encounter have been documented previously in the Atlantic, with notable variation in frequency, duration, and the combination of multiple varieties in a single pulse train. These differences likely represent regional and/or seasonal variation for this species, which are invaluable for training classifiers. These data also provide insights on North Atlantic minke whale winter habitat and vocal behavior, for which we currently lack data.

Time-Frequency Exploration of the Repertoire and Evolution of Humpback Whale Songs in the Caribbean Sea

S. Chavin[1,2*], J. Couvat[3], H. Glotin[2]

[1]*Université de Toulon, Aix Marseille Univ, CNRS, LIS, Marseille, France*

[2]*CIAN, Centre d'Intelligence Artificielle en Acoustique Naturelle, Toulon, France https://bioacoustics.lis-lab.fr/*

[3]*Agoa Sanctuary, French Agency of Biodiversity,Trois-Ilets, Martinique, French West Indies*

**stephanechvn@gmail.com*

Humpback whales (Megaptera novaeangliae) around the world exhibit complex and organized vocalizations into a song form [1]. Traditional manual annotation methods for studying whale songs are labor-intensive and time-consuming. In this paper, we worked on the humpback whales in the breeding area of the Caribbean Sea and introduced an innovative time-frequency annotation method to use the YOLOv5 neural network [2, 3] in order to automate the detection and analysis of humpback whale vocalizations.

Our approach addresses the limitations of existing methods, providing a streamlined and accurate solution for long-term surveys. We present the methodology, including dataset details, YOLOv5 configuration, data augmentation and adaptations for Caribbean Sea humpback whales. The results showcase the model's performance metrics and visualizations of annotated time-frequency spectrograms. Additionally, we unveil a part of the vocalization repertoire of Caribbean Sea humpback whales, classifying and describing 28 song unit types observed during the survey and their proportion for each recording station.

Then we interpret the findings in the context of existing research, highlighting the implications of YOLOv5 for advancing humpback whale vocalization studies. We conclude by summarizing key contributions and suggesting avenues for future research in marine mammal acoustics. This study presents a novel and accessible approach, leveraging artificial intelligence to enhance our understanding of humpback whale communication in the Caribbean Sea.

We thanks AGOA and all the precious and numerous partners of CARIMAM https://interreg-caraibes.eu/download/file/fid/5867.

We thank EUROPAM Biodiversa project, ANR grants ADSIL ANR-20-CHIA-0014 and ULPCochlea ANR-21-CE04-0020 received by H. Glotin.

[1] Payne, Roger and Scott McVay. “Songs of Humpback Whales.” Science 173 (1971): 585 - 597.

[2] Glenn Jocher et al. (2022) “ultralytics/yolov5: v7.0 - YOLOv5 SOTA Realtime Instance Segmentation”.

[3] Deloustal Nicolas, Chavin Stéphane, Glotin Hervé (2022), “Suivi et identification temps-fréquence bioacoustique par transfert deep learning sur YOLO : gestion des chorus” http://sabiod.lis-lab.fr/pub/YOLO_DYNI.pdf

Classifying vessels and co-occurance with mammals using CNNs based on underwater acoustics

W. Decrop[1*], C. Parcerisas[1,2], E. Schall[3], E. Debusschere[1], K. Deneudt[1]

[1]*Flanders Marine Institute (VLIZ), InnovOcean site, Jacobsentstraat 1, 8400 Oostende, Belgium*

[2]*WAVES Research Group, Department of Information Technology, Ghent University, Technologiepark-Zwijnaarde 126, 9052 Gent, Belgium*

[3]*Alfred Wegener Institute for Polar and Marine Research, 27570 Bremerhaven, Germany*

**wout.decrop@vliz.be*

Marine Protected Areas (MPAs) have been established to safeguard coastal regions against increasing human activities. Within these areas, human interactions with the environment require regulation. However, effectively monitoring these regions presents significant challenges, often leading to illegal activities. An effective way to approach the presence of human activity, is through ocean acoustics. By using machine learning techniques, object detection and classification can be performed to lay the groundwork for monitoring underwater acoustics.

A machine learning model is trained based on AIS data and passive acoustic hydrophone recordings. Although AIS data's main purpose is to avoid collisions (in real time), the historical data contains information about the voyage information (e.g., type of vessel) and the position (e.g., longitude). The acoustic data is based on over 100 days of acoustic data obtained from two North Sea stations. By combining the recordings with their relative distance to by-passing boats, and their type, a database was established to detect and classify marine vessels. From this dataset, a convolutional neural network (CNN) was created. It aims to predict the distance, activity, and type of vessel based on their sound signature.

The DCLDE dataset serves two primary purposes. Firstly, it validates the performance of the created model by predicting vessel distance based on recordings. Secondly, it predicts the co-occurrence of mammals and ships. In future steps, the model could be retrained using the DCLDE dataset to further enhance its performance.

Trajectory estimate of baleen whales using a single OBS in the Indian Ocean

R. Dréo[1*], W. Crawford[1], G. Barruol[2], S. Bazin[3], J.-Y. Royer[3], F. Samaran[4]

[1]*Institut de physique du globe de Paris, Université de Paris Cité, CNRS, F-75005 Paris, France*

[2]*CNRS, IRD, Grenoble INP, IGE, Université Grenoble Alpes, Grenoble, France*

[3]*Université Brest, CNRS, Ifremer, UMR6538 Geo-Ocean, 29280 Plouzané, France"*

[4]*ENSTA Bretagne Lab-STICC – UMR CNRS 6285, 29806 Brest Cedex 09, France*

**dreo@ipgp.fr*

Ocean Bottom Seismometers (OBS) are sensors commonly used to monitor the seismic activity during periods up to several months. They are generally equipped with a combination of a 3-components seismometer (to estimate the direction of seismic events) and a hydrophone. The past years, OBS showed a great potential for passive acoustic monitoring of sources such as ship or baleen whales. The seismometer is a vector sensor designed to measure the ground displacement, but appears also to be sensitive to the sound propagating in the water column. Therefore, using the 3-components and the hydrophone, it is possible to compute the intensity vector which provides information on the azimuth of the incoming waves propagating in the water column.

From the capability of estimating a source azimuth and the characteristics of the multi-paths propagation, we propose an adaptation of a method initially developed for ships, to estimate the trajectory of a singing whale using a unique OBS. The algorithm is made of two independent processing chains. The first one aims at estimating the azimuth of the source, while the second chain provides possible distance solutions by exploiting the multi-paths propagation.

Demonstration of the potential (and limitations) of this method is done on the REVOSIMA (Réseau de surveillance volcanologique et sismologique de Mayotte) dataset, with a particular focus on the SWIO pygmy blue whale. In the Mayotte area, distance estimates on this species can be up to 60 km, while the azimuth can be measured at higher ranges. The main limitation resides in linking the results provided by the two chains when having multiple singing individuals at the same time, which happens quite often in this area. To improve the performance of this method, in the future, the addition of a tracking algorithm could provide a way to solve this problem, and also extend the tracking of an individual at ranges up to 80-100 km when only the azimuth information remains.

This work is co-funded by the French Biodiversity Agency.

First attempt at building a mini DCASE-like data challenge for the DCLDE workshop

D. Cazau[1*], N.H.D. Paul[2], B. Padovese[3], F. Soares[3], G. Dubus[4], N. Farrugia[5], A. Marmoret[5], O. Adam[4]

[1]*ENSTA Bretagne, Lab STICC UMR 6285, Brest, France*

[2]*Centre for Marine Science and Technology, Curtin University*

[3]*School of Environmental Science, Simon Fraser University*

[4]*Sorbonne University, CNRS, Institut d'Alembert UMR 7190, LAM, Paris, France*

[5]*IMT Atlantique, Lab-STICC, UMR CNRS 6285, F-29238 Brest, France*

**gabriel.dubus@dalembert.upmc.fr*

For 20 years, the DCLDE workshop has been gathering the international community of underwater Passive Acoustic Monitoring (PAM) experts to share their research works. Since the very beginning of the workshop, an emphasis has been put on the development of automated tools addressing general tasks such as the detection and classification of marine mammal vocal sounds, especially through the creation of a dedicated data challenge. Without any doubt, the DCLDE community is the place-to-be where the FAIR (Findable, Accessible, Interoperable, Reusable) state-of-the-art of these methods should be developed. This challenge has already favored the emergence of elements upon which we could build up this state-of-the-art, such as the release of publicly available datasets (eg. the DCLDE2015 LF dataset) and the showcase of reference processing tools (eg. the PAMGuard software).

Despite these long-term efforts, our DCLDE community remains behind other communities on the road of building state-of-the-art AI models. A comparison with the DCASE (https://dcase.community/) workshop (10 years younger) has allowed us to highlight the lack of key elements. First, the absence of reference tasks makes it difficult to benchmark methods on the long term. Second, the tasks and the formatting of their associated materials (i.e. dataset, downloading pipeline, code baseline, evaluation protocol..) are not well defined or standardized. In DCLDE, the tasks proposed are often too vague and the datasets too voluminous, while in DCASE the tasks are defined with a machine learning-oriented vocabulary and ontology, ensuring proper data formatting for AI models (eg development/evaluation partition, light volume, long-term open access…).

We propose to fill this gap existing between our DCLDE community and more oriented machine learning communities like DCASE. After analyzing the different elements explaining this gap, we will present our first attempt at building a DCASE-like data challenge for marine bioacoustics. Currently, our challenge has set up three main tasks, one for each of the three main families of machine learning models: supervised, semi-supervised and unsupervised models. We will review all DCASE standards and practices we closely followed, such as making exhaustive and technical task presentations, AI-oriented formatting of data, and centralizing all materials on dedicated webpages. We hope that DCLDE participants will come challenging our current baselines, but more importantly, we hope that this preliminary proposition will be used as a discussion basis in DCLDE to see how this kind of "individual" initiatives could be scaled up to the international community level, and more importantly to prepare altogether the future of AI in DCLDE.

Using passive acoustic to better understand dolphins' behaviour around fishing nets in bycatch context

M. Dupont[1*], V. Lagarde[2], V. Loirat[2], B. Guichard[3], F. Samaran[1]

[1]*ENSTA Bretagne*

[2]*CDPMEM29*

[3]*OFB*

**mathieu.dupont2@ensta-bretagne.fr*

Bycatch due to fishery interactions is considered as the main threat to common dolphins (Delphinus delphis) in European waters. Some solutions are being investigated such as the use of pingers to prevent incidental captures. However, little is known as for the nature of these interactions and more interestingly on the circumstances of captures. This lack of knowledge implies a more challenging implementation of appropriate and effective means for mitigation of small cetaceans. Passive Acoustic Monitoring (PAM) represents a cost-effective and reliable solution to monitor how small cetaceans behave around nets using their acoustic behaviour. Whistles are typically associated with communication behaviour whilst clicks and buzzes are associated with foraging behaviour.

The aim of the APOCADO project is to address this question and to provide an insight on interactions between delphinids and fishing nets in the Iroise Sea (Brittany, France). Acoustic data was collected using SoundTraps (ST300 / ST400) directly deployed on the fishing nets. They were recording continuously at a sampling frequency of 144 kHz. During one year (May 2023 - April 2024) seven recording campaigns involving five different fishing vessels and different type of fishing nets have been conducted cumulating 3700 hours of acoustic data.

Various whistles, clicks and buzzes are reported throughout the recordings. Automatic detection tasks have been assigned to PAMGuard software for whistles and click trains have been detected through a custom detector. The occurrence of common dolphins and their associated acoustic behaviour around the nets were monitored at each step of the fishing activity (setting, soaking and hauling). Results were compared with other criteria such as the type of net used (gillnet, trammel), the location of the fisheries or the season. To draw additional conclusions results were also correlated with relevant auxiliary data such as tidal coefficients, ambient noise or night-time/daytime periods.

An Initiative for whale Detection in the Santos Basin, Brazil: Through Passive Acoustic Methods

S. Espirito Santo[1,2*], W. Soares-Filho[1], F. Lima[1], A. Bazyl[1], F. Xavier[2], L. Calado[2]

[1]*Brazilian Navy Research Institute, Rio de janeiro, Brazil*

[2]*Paulo Moreira Admiral Sea Studies Institute (IEAPM), Arrail do Cabo, Brazil*

**sidneyrafaelgomes@gmail.com*

The Passive Acoustic Mapping (PAM) has proven to be a promising technique in cetacean distribution studies, as conventional cetacean observation techniques are limited to daylight hours and weather conditions favorable for navigation. Due to the volume of acoustic data generated by a PAM, the automation of marine mammal detection has become usual. Some studies have employed automatic cetacean detection on the Brazilian coast using tools like PAMGuard and other automatic detection programs. However, there is a lack of information regarding the application of detectors on larger volumes of data. In this work, the efforts undertaken in the detection of whale vocalizations in the Santos Basin are presented, based on acoustic recordings obtained using gliders between 2015 and 2022, as part of the Santos Basin Underwater Soundscape Monitoring Project (PMPAS-BS). The execution of the PMPAS-BS is a conditioning required by the Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis (IBAMA) under the federal environmental licensing of the oil and gas production activities of Petrobras at Santos Basin. The glider was programmed to acquire acoustic data received by a hydrophone during its descent, operating with a sampling frequency of 125-128 kHz. The detection module of the PAMGuard Whistle Moan Detector was employed, with a frequency band of 50 to 3000 Hz and a SNR of 6 dB. The effort in collecting acoustic recordings was 12,717.23 hours, totaling 15.64 TB of data . A total of 835 segments of whale moan signals were detected. There were more whales in the month of september and during the nighttime period, which is consistent with literature reports. Spatially, there was a concentration of detections in the northeast area of the Santos Basin.

We are still taking detections attributed to the glider's self noise, which mainly occurs at the beginning of the glider's descent. But, manual inspection has been successful in discarding self noises. The detection tool has shown promising results, especially for segments with a higher density of detections, where reliability is high.

Evaluation of an abbreviated, fully unsupervised approach for classification of odontocete echolocation clicks

M.L. Garcia[1,2*], R.E. Cohen[1], L.G. Torres[3], H. Klinck[1,4]

[1]*K. Lisa Yang Center for Conservation Bioacoustics, Cornell Lab of Ornithology, Cornell University*

[2]*Department of Natural Resources and the Environment, Cornell University*

[3]*Marine Mammal Institute, Department of Fisheries, Wildlife, and Conservation Sciences, Oregon State University*

[4] *Marine Mammal Institute, Department of Fisheries, Wildlife, and Conservation Sciences, Oregon State University*

**mg2377@cornell.edu*

While the northern California Current System (nCCS) off the central Oregon coast is a highly productive foraging ground for top predators, odontocete biodiversity in the region has yet to be systematically studied. With increasing regional climatic pressures, such as marine heatwaves and hypoxia, describing local, baseline odontocete phenology provides critical information for quantifying climate change impacts over time.

Odontocetes produce echolocation clicks with distinct spectral and temporal features that can facilitate species-level attribution and identification via passive acoustic monitoring (PAM). Clicks, however, can total many millions within large PAM datasets, making analysis unwieldy without the application of automated methods. Though previous machine learning methods have successfully detected and classified odontocetes, the need to pre-define classes, manually validate detections, and run multiple species-specific detectors sequentially can bottleneck analyses. One notable exception is a workflow developed by Frasier et al. (2017; 2021), which achieves multi-species odontocete identification via a single workflow combining unsupervised and supervised methods.

In this study, we leverage preexisting knowledge of expected species in the nCCS from numerous visual surveys to explore whether an abbreviated, fully unsupervised adaptation of the Frasier workflow offers comparable accuracy to the complete workflow. The neural network classification step is necessary to obtain click-level labels in cases where clicks are so numerous, and the datasets so large, that unsupervised clustering is computationally limited to a random subset of clicks. For datasets with more tractable levels of odontocete species presence, unsupervised clustering of all clicks becomes computationally feasible, obviating a further labeling step.

Central Oregon waters embody an ideal testing ground for a fully unsupervised workflow, with modest regional odontocete species diversity comprising sperm whales, three beaked whales, and three delphinids, all of which have been well acoustically characterized. Using two years of passive acoustic data collected by Rockhopper units across three sites spanning the continental slope and abyssal plain (300 m, 630 m, and 2,860 m water depth), we implement an abbreviated version of Frasier's 2021 workflow including (1) a generic impulse click detector and (2) unsupervised clustering to identify and label known signal types. Comparing these results to both (1) a ground truth data set constructed via manual annotations from long-term spectral averages and (2) results from Frasier's full workflow inclusive of the neural network, we generate performance evaluation curves to evaluate the accuracy of this abbreviated workflow for future applications toward uncovering climate change impacts on odontocete phenology and ecology.

Passive acoustic in Arctic and Mediterranean seas to compare nictemeral rhythms of cetaceans and anthropophony

J. Girardet[1*], V. Sarano[2], A. Marchini[3], S. Chavin[4,5], H. Glotin[2]

[1]*Université de Toulon, AMU Univ, CNRS, LIS, DYNI team, CIAN, Marseille, France*

[2]*Longitude 181, Valence, France*

[3]*Pavia univ., Pavia, Italy*

[4]*Université de Toulon, Aix Marseille Univ, CNRS, LIS, Marseille, France*

[5]*CIAN, Centre d'Intelligence Artificielle en Acoustique Naturelle, Toulon, France*

**justinegirardet@gmail.com*

Cetaceans heavily rely on sound for essential life activities such as communication and navigation. The ever-increasing anthropogenic activities pose considerable threats including vessel collisions and acoustic pollution potentially leading to acoustic masking, behavioral changes and/or physical damages. Using long term passive acoustic monitoring, we aimed to explore the impact of anthropophony on cetaceans species in Northern Norway and in Northwestern Mediterranean Sea as part of the Europam project1.

Acoustic recordings were acquired from three autonomous sonobuoys called BOMBYX. In the Mediterranean Sea, antennas were placed at 25m depth near Port-Cros national park France (2015-2018), and off Monaco (2022-2023). The recording protocol has changed over the years (continuous recording, 5 min recording every 20 min or hour). In Arctic, Norway, the antenna was positioned on the seabed in Kvaenangen fjord (November 2022-January 2023. Neural networks were trained to automatically detect vocalizations of targeted species and outlined their acoustic presence. Ambient noise for two frequency bands (10-400 Hz and 6500-12500 Hz) was estimated by power spectral density (PSD).

In the Mediterranean Sea, our analysis revealed the year-round acoustic presence of sperm whales, with no particular agenda. A discernible daily pattern was identified with higher detections during morning and afternoon hours (7am to 6pm). Fin whales were mainly detected in autumn and winter. Ambient noise calculations revealed strong anthropogenic activity between 10am and 6pm in summer in the low- and medium-frequency bands2. Levels in autumn and winter are lower, except for low-frequency raking noises that warrant further investigation, as they may hinder the detection of fin whales. Analyses of 2015-2018 recordings have shown a significantly lower acoustic activity of sperm whales under high levels of anthropogenic noise3. Ongoing investigations aim to validate this trend in 2022-2023.

In Northern Norway, findings unveil a continuous presence of orcas, fin and humpback whales with distinct diurnal variations4. Fin whale pulses were predominantly detected during daylight hours, while orcas and humpback whales exhibited heightened activity during nighttime. This diurnal activity will be linked with ambient noise calculations results to highlight potential adaptations of cetaceans species to noise disturbances.

Our research will help reveal patterns of presence and behavior of cetaceans species in response to varying levels of anthropogenic noise.

[1] Projet Europam 2023-2026. https://bioacoustics.lis-lab.fr/europam/

[2] Poupard M., et al. Passive acoustic monitoring and interspecific interactions of orcas, fin whales and humpback whales in arctic Kvaenangen Fjord, Norway. Submitted to Marine Mammals 2023-12.

[3] Glotin H., et al. Bilan d'une décennie d'observations de grands cétacés en milieu anthropisé Nord Pelagos : Bombyx-1&2, Km3Env et antennes mobiles Jason. 2024. Rapport scientifique Pelagos.

[4] Poupard M., et al.. Passive acoustic monitoring of sperm whales and anthropogenic noise using stereophonic recordings in the Mediterranean Sea, North West Pelagos sanctuary. 2022. Scientific report.

Does the beam pattern matter? Impacts of different assumptions on off axis source level of echolocating clicks for PAM density estimation of deep diving species

K. Gkikopoulou[1,2*], D. Gillespie[1], M. Johnson[3], P. Tyack[2], N. Aguilar de Soto[4], T. Marques[5,6]

[1]*Sea Mammal Research Unit, Scottish Oceans Institute, University of St Andrews, Fife, KY16 8LB, Scotland*

[2]*Center for Research into Ecological and Environmental Modelling, Buchanan Gardens, University of St Andrews, St Andrews, Fife, KY16 9LZ, UK*

[3]*Aarhus Institute of Advanced Studies, Aarhus University, Aarhus, Denmark*

[4]*BIOECOMAC, Department of Animal Biology, Edaphology and Geology, University of La Laguna, La Laguna, Spain*

[5]*Center for Research into Ecological and Environmental Modelling, Buchanan Gardens, University of*

St Andrews, St Andrews, Fife, KY16 9LZ, UK

[6]*Centro de Estatística e Aplicações, Departamento de Biologia Animal, Faculdade de Ciências da Universidade de Lisboa, Portugal*

**kg366@st-andrews.ac.uk*

Density estimation from PAM using cue counting requires knowledge of parameters such as the cue production rate, the average detection probability, and their associated variance.

While is important to estimate accurate cue production rates, as this translates the number of detected cues to animal density, less information exists about the off-axis source level of echolocation clicks and how this affects the average detection probability and density estimation. Odontocetes produce clicks with energy focused on a forward beam increasing the target efficiency and reducing echoes from the environment. For a PAM detection system this translates to reduced detection probabilities in comparison to omni-directional sources. The effect of how much this affects the detection probability is a result of species-specific characteristics and its echolocation signals such as signal frequency and apparent source level.

We investigated the effect of uncertainty in the source level of off-axis echolocation clicks (i.e. the beam pattern) for two deep diving species, sperm whales (Physeter macrocephalus) and Blainville's beaked whales (Mesoplodon densirostris) and its implication on detection probability when using a simple cue counting density estimation method. We simulated a network of receivers and then used animal orientation and time series of echolocation click production, derived from animal borne tag data, to estimate the sound received level based on the different plausible beam patterns (n=2 for sperm whales, and n=3 for beaked

whales). Proportions of off-axis clicks varied with species and assumed noise levels. Sperm whales are detected from a variety of angles regardless of noise levels and beam patterns, indicating that knowing the off-axis beam pattern is very important for this species. In high noise conditions, only on-axis beaked whale clicks are detected, meaning that detailed knowledge of the of-axis beam pattern may have little importance for this species. However, in quieter conditions, beaked whale clicks are detected over a wider range of angles, and a different assumption about the shape of the off-axis beam pattern can change detection probability for this species by up to 50%. We do not provide new measurements of off-axis source level rather we show that existing uncertainty of beam pattern can affect average detection probability estimates and consequently bias in density estimation. This consideration is especially important when detectability is estimated by models or simulations but can also be relevant in defining optimal survey designs and priorities for future empirical measurements of off-axis beam pattern for different species.

OPALE : a high resolution multistream audiovisual mobile antenna for cetacean ethoacoustics

H. Glotin[1*], V. Barchasz[2], S. Marzetti[2], V. Gies[2]

[1]*Université de Toulon, Aix Marseille Univ, CNRS, LIS, DYNI team, SmioT, CIAN, Toulon, France*

[2]*Université de Toulon, AMU Univ, CNRS, IM2NP, SmioT, CIAN, Toulon, France*

**glotin@univ-tln.fr*

Diarization of high definition acoustic and visual observations (2) is a key point in ethoacoustics (animal behavior and communication studies). The need for multistream is crucial in the case of whale cocktail parties. Hence, since 2017 we have developed a mobile high definition multi-hydrophone audiovisual antenna. It includes the QHBv3 audio recorder developed in Toulon university, allowing precise click detection, tracking of Time Differences of Arrival and computation of Directions of Arrival with audio-video synchronization. This multistream system provides a three-dimensional understanding of the cocktail party of cetaceans (5), and a video pointing system to locate the source of each click and voicing emitted by the cetaceans.

Therefore, we designed and developed at Smiot (4) a novel multi-channel ultra high velocity acoustic device, called Qualilife HighBlue (QHB) (1). It includes shallow neural networks, a mixed signal features extractors and communications devices. Multiple mobile deployments of this monitoring system over the world are presented, demonstrating its robustness, versatility and efficiency.

The recent version of OPALE is equipped with four or five hydrophones and 1 or 2 GoPro cameras. It recorded close-range interactions of Cachalots or Orcas providing unique ethoacoustical observations. It opens novel ethoacoustical research that focuses on exploring interactions between individuals.

Identifying by beam forming each speaker enables in-depth analysis of “dialogues” among individuals. This advancement also allows for excluding clicks and voicings considered non-relevant in the observed interaction, expanding our capacity to analyze socio-acoustic behaviors and individual click trains, with high definition waveform (2,3). Also, obtaining the direction of each click, knowing which click features characterize an individual and which describe the information contained in a click. Thus it can help to improve a cetacean head model from free ranging animals.

We thank Maxence Ferrari who helped in the design of one version of OPALE and F. Sarano, V. Sarano, M. Poupard, H. Glotin P. Giraudet, L. Berkenbaum who participated in recordings with OPALE on Cachalot or Orcas. We thank EUROPAM Biodiversa project, ANR grants ADSIL ANR-20-CHIA-0014 and ULPCochlea ANR-21-CE04-0020 received by H. Glotin.

[1] V Barchasz, V Gies, S Marzetti, H Glotin, A novel low-power high speed accurate and precise DAQ with embedded artificial intelligence for long term biodiversity survey, in Proc. Acustica Symposium, 2020

[2] M Ferrari, H Glotin, M Oger, R Marxer, M Asch, V Gies, F Sarano, 3D diarization of a sperm whale click cocktail party by an ultra high sampling rate portable hydrophone array for assessing individual cetacean growth curves, in Proc Forum Acusticum, 3239-3243, 2020

[3] M Ferrari, H Glotin, R Marxer, V Barchasz, V Sarano, V Gies, M Asch, High-frequency near-field Physeter macrocephalus monitoring by stereo-autoencoder and 3d model of sonar organ, IEEEOCEANS, 1-4 2019

[4] V Gies, V Barchasz, H Glotin, Smiot: Scientific microsystems for the internet of things, technological platform

[5] H Glotin, F Caudal, P Giraudet, Whales cocktail party: a real-time tracking of multiple whales Int. Journal Canadian Acoustics 36 (1), 7p, 2008

Fully Automatic Detection and Classification of Sperm Whale Codas

G. Gubnitsky[1*], R. Diamant[2]

[1]*Department of Marine Technologies, University of Haifa, Aba Khoushy Ave 199, Haifa, 3498838, Israel*

[2]*University of Zagreb Faculty of Electrical Engineering and Computing, Unska 3, Zagreb, Croatia*

**ggubnits@campus.haifa.ac.il*

CETI (Cetacean Translation Initiative) is an interdisciplinary research initiative that aims to understand and translate the communication of sperm whales (Physeter macrocephalus) using advanced machine learning and state-of-the-art robotics. A key technology for this is the detection of sperm whales' unique communication signals, known as codas. Unlike all other cetaceans, which use time-frequency modulation of their sounds to communicate with each other, sperm whale codas consist of a varying number of click-like signals whose rhythm, tempo and number of clicks are used to encode social messages. To date, codas have been recognized by various automatic click detection methods and subsequent manual annotation to distinguish coda clicks from echolocation clicks, separate vocalizers and associate clicks with the same coda. However, the analysis of large quantities of data requires complete automation to detect, cluster and characterize coda types. In this talk, we will present a method for automatic coda identification based on graph-based clustering. Exploiting the high similarity between the clicks that form the sequence of the coda, our detector recognizes codas by searching for highly concentrated clusters of nodes in an undirected graph representation. By restricting the cluster solution to structures of clicks that are known to be valid and to the examination of transient sequences that reflect an echolocation click structure, our detector can further distinguish between codas and echolocation clicks and discriminate vocalizations from simultaneously emitting whales. The detection is formalized as a graph Laplacian optimization problem that can be solved optimally with low complexity, allowing implementation on real-time platforms. The performance of our detection approach is evaluated on three datasets: seven months of recordings from the Mediterranean Sea, containing manually verified ambient noise for false alarm estimation; 30 minutes of manually labeled data from the Dominica Island, containing approximately 730 codas recorded with a single hydrophone on a boat to assess the detection rate of remote sources; and a dataset from Dominica Island containing 843 labeled codas from Dtag recordings to assess detection in the presence of echolocation clicks.

Using passive acoustic monitoring to investigate northern bottlenose whale (Hyperoodon ampullatus) migration theories within the eastern North Atlantic

C.E. Haas[1,2*], P.J.O. Miller[2], S.K. Hooker[2], J. Svavarsson[3], P.J. Wensveen[3]

[1]*Faculty of Life and Environmental Sciences, University of Iceland, Reykjavík, Iceland*

[2]*Sea Mammal Research Unit, Scottish Oceans Institute, School of Biology, University of St Andrews, St Andrews, UK*

[3]*Westman Islands Research Centre, University of Iceland, Vestmannaeyjar, Iceland*

**ceh4@hi.is*

Northern bottlenose whales (Hyperoodon ampullatus) are deep-diving beaked whales distributed in Arctic, sub-Arctic and temperate offshore waters. While the well-studied Scotian Shelf population (Canada) has been found to be resident and genetically distinct, matrilines shared between populations in more northern regions off Canada, Greenland, and Iceland indicate at least historical mixing between these populations. During whaling times, the Icelandic population has been suggested to undergo north-south migrations, with sighting data and recent evidence from satellite-tagged animals supporting this theory. However, visual sighting data also supports seasonal inshore-offshore migrations. Here we investigate occurrence of this species via acoustic detection on duty-cycled bottom-moored hydrophones. Multi-year data were collected in three locations; east and north-east of Iceland (2020-2023) and off Jan Mayen Island (Norway; 2015-2017). We processed recordings with the SPICE detector remora in Triton (MATLAB) using a two-step detection and classification approach: (1) detection of sounds based on a peak-to-peak level threshold, and (2) click classification using known acoustic characteristics of bottlenose whales' regular echolocation clicks. From the raw detection and classification output, the percentage of recorded files containing detections was calculated for weekly time bins. The results revealed seasonal acoustic occurrence patterns that remained stable across years, consistent with annual north-south migrations. Northern bottlenose whales were acoustically present in Icelandic waters throughout autumn to spring, being detected during 11.5% of recorded files averaged across weeks. Detections around Iceland were highest in March and April (weekly maximum: 54.3%) and reached a low between May and early June (mean 2.7%). Around Jan Mayen, detections gradually increased from September and reached their maximum (64.7%) the following May and early June indicating a steady northward migration. Detections in Jan Mayen dropped rapidly in late June, followed by a short detection peak in Icelandic waters, and low numbers of acoustic detections (mean 2.0%) for all three locations throughout late July, August and September indicating a southward migration. During these months, visual sightings of northern bottlenose whales further south in the Faroe Islands, off Ireland and the UK, and around the Azores (Portugal) are common. To validate acoustic occurrence patterns, detector false positive rate is being manually evaluated on a subset of the data representative of different times of the year and

day, and different numbers of detections contained within a recording period. Potential drivers of northern bottlenose whales' seasonal occurrence patterns will then be explored via investigation of environmental covariates at the different recording sites.

Applying distance sampling to estimate densities of fin whale calls recorded by ocean bottom seismometers in the Marianas region

R. Hilmo[1*], D. Harris[2], W.S.D. Wilcock[1]

[1]*University of Washington*

[2]*University of St. Andrews*

**wader@uw.edu*

Ocean bottom seismometers (OBSs) function as opportunistic instruments for passive acoustic monitoring of fin whales. To estimate animal population densities from acoustic data using distance sampling, a standard density estimation method, it is essential to estimate ranges to detected calls. Regional OBS networks are too widely spaced to locate fin whale calls using arrivals on multiple instruments. This motivates the development of ranging methods using single instruments.

We developed and evaluated a single-OBS ranging method for fin whale calls that uses multiple arrivals of calls reflected from the seafloor and sea surface. The timings of multipath arrivals were used to estimate horizontal ranges to calling fin whales using OBSs deployed at various ocean depths (2000-6000m) and on seafloor substrates ranging from basalt to 3 km of sediments. We found that OBSs deployed on the Pacific plate near the Marianas Trench at ~6000m depth were exemplary sites for ranging due to strong recorded multipaths from a highly reflective chert layer about ~150 m below the seafloor. Fin whale ranges were successfully calculated out to ~40 km.

We are applying our multipath ranging method to estimate densities of fin whale calls recorded on 12 OBSs deployed over a 200 x 300 km2 area along the Marianas trench from February 2012-February 2013. We plan to detect and range to calls within 40 km at each OBS, then apply point-transect distance sampling. We will test and evaluate the quality of various detection functions fit to the range distribution to determine which is best for estimating the average probability of detecting a call as a function of distance. We will test the whether the inclusion of detection function covariates including received call amplitudes and signal-to-noise ratios improve the detection function.

Preliminary results from one OBS found that detection of calls was nearly certain out to 12 km. However, beyond this range detection functions must account for local minimums in detection probability due to difficulties in detecting calls of low amplitude associated with the critical angle. Range distributions are expected to vary by instrument due to different deployment depths, so OBS site will be tested as an additional detection function covariate. We will use call detections, associated ranges, and the estimated detection function to investigate spatial and seasonal variations in fin whale call densities throughout the Marianas network. Converting call densities into estimates of animal abundance would require an independent estimation of fin whale call rates.

A Comparison of Two Distributed Acoustic Sensing Systems for Recording Marine Mammal Vocalizations

E.B. Horeh[1*], S. Abadi[1], W.S.D. Wilcock[1]

[1]*School of Oceanography, University of Washington, Seattle, Washington 98195, USA*

**erfanbh@uw.edu*

Studies that seek to detect and localize marine mammal vocalizations within the oceans are challenged by the practical limitations of collecting data with wide spatial coverage, long duration, and good timing precision. This provides a motivation for new approaches that can complement the use of traditional hydrophone networks. Distributed Acoustic Sensing (DAS) is a relatively new technology that is generating considerable interest because it can turn an optical fiber into a sensing element that could be piggybacked on the global network of telecommunication cables.

DAS interrogators work by injecting repeated pulses of light into one end of an optical fiber and measuring changes in strain along the fiber manifested in changes in the relative timing of backscattered returns. DAS systems collect measurements on uniformly spaced channels along a fiber with the reading on each channel averaged over a distance termed the gauge length. For acoustic applications, DAS thus transforms a fiber-optic cable into a dense linear array capable of capturing the strain changes associated with acoustic signals.

We are comparing the performance of two common DAS interrogators, the Optasense QuantX and Silixa iDASv3, for monitoring marine mammal calls and ship noise. The study uses 4 days of DAS data collected in November 2021 on the two cables of the Ocean Observatories Initiative Regional Cabled Array extending offshore central Oregon. DAS measurements extended out to the first optical repeaters at distances of 65 km and 95 km on the north and south cables, respectively. Optasense data was collected on both cables and Silixa data on the south cable. The sample rates vary from 200-1000 Hz and the gauge lengths are 30 m and 50 m for Optasense and 2 m, 10 m, and 30 m for Silixa.

We are initially investigating ship signals because these are broadband signals from vessels with known tracks, but we will also examine fin whale calls which extend from ~15-30 Hz and blue whales calls which extend from ~14-45 Hz. We are comparing the detection range of signals and looking at the signal to noise of filtered time series, spectra, and spectrograms. Initial comparisons suggest that the Silixa measurements are more sensitive to higher frequency ship signals but further analysis is required to determine whether this reflects the shorter gauge lengths or differences in the interrogators.

Intraspecific geographic variation of rough-toothed dolphin whistles and its influence on acoustic classification

V.M. Janik[1*], C.R. Day[1], E. McCloskey[1], A. Snyder[1], A.D. Paro[2], G. Fogarty[1], E. Hayward[1], J. Lees[1], O. Konovalov[3], J.N. Oswald[1]

[1]*Scottish Oceans Institute, University of St Andrews, Fife KY16 8LB, UK*

[2]*Marine Biotechnology Program, Instituto de Estudos do Mar Almirante Paulo Moreira, 2 Kioto Street, Number 253, Praia dos Anjos, Arraial do Cabo, Rio de Janeiro, 28930000, Brazil*

[3]*School of Computer Science, North Haugh, University of St Andrews, Fife KY16 9SX*

**vj@st-andrews.ac.uk*

Intraspecific geographic variation in dolphin vocalizations is often assumed to be the cause for poor acoustic species classifier performance across different regions. However, such variation has only been quantified for a few species and it is often unknown whether some acoustic features are preserved across regions. The rough-toothed dolphin (Steno bredanensis) is a species in which the persistence of one feature (sudden frequency steps in the frequency modulation pattern of whistles) across oceans has been suggested. We therefore investigated parameter and whistle type variation in rough-toothed dolphins to assess whistle repertoire and acoustic feature stability across samples from the Tropical Pacific, the Southwest Atlantic and the Mediterranean Sea.

We used the Real-time Odontocete Call Classification Algorithm (ROCCA) module in PAMGuard to extract 24 acoustic parameters from the fundamental frequency modulation pattern of 107 randomly selected whistles from multiple encounters in each location. Kruskal-Wallis comparisons with Dunn tests were used to investigate whistle parameters across the three regions. In addition, we conducted an ARTwarp analysis of the whistle repertoires to compare whistle modulation patterns (types) across regions (https://doi.org/10.5281/zenodo.7713615).

Almost all parameters showed significant differences between at least two regions with the exception of the minimum frequency. Ten of 24 variables differed significantly between the Pacific and the Atlantic sites, while almost all variables (20 out of 24) differed significantly between the Mediterranean Sea and the other two sites. ARTwarp analysis found 7 shared whistle types across all three oceans, with an additional 16 types shared between two of the three oceans. Twenty-eight whistle types were unique to one location (11 in the Pacific, 13 in the Mediterranean and 4 in the Atlantic). However, whistles in all regions were mostly simple upsweeps or downsweeps that differed in specific parameters but showed very few inflection points. While we found whistles with abrupt frequency shifts in the Pacific sample, they were not apparent in our Mediterranean sample, suggesting that previous reports of these may have come from a different geographic area within the Mediterranean. Further work is needed to assess whether shared whistle types across regions are sufficiently different from those of other delphinids to use them as species identifiers for rough-toothed dolphins universally.

Leveraging citizen science in passive acoustic monitoring of cetaceans

G. Jankauskaite[1*], J. Menéndez-Blásquez[1], I. Abascal[1], M. Díaz[1], C. Cánovas[2], R. Sagarminaga[2], A. Sánchez Jones[2], J.A. Raga[1], D. March[1]

[1]*Cavanilles Institute of Biodiversity and Evolutionary Biology, University of Valencia, Paterna, Spain*

[2]*Alnitak, Madrid, Spain*

**greta.jankauskaite@uv.es*

Ecological monitoring of marine habitats is key to understanding and assessing the health of these ecosystems. Passive acoustic monitoring (PAM) is a powerful, non-invasive technique for detecting marine animal vocalizations and instances of anthropogenic disturbance. However, financial constraints limit PAM applications within marine environments due to typically high recorder costs. While citizen science initiatives hold the potential to contribute to the monitoring of underwater environments and expanding the pool of available acoustic data, the challenge lies in the absence of accessible, affordable, and user-friendly recorders. The development of low-cost hydrophones, such as the HydroMoth, offers new opportunities for remote monitoring of cetaceans and underwater noise. The elevated noise derived from the shipping industry, persisting over extended periods, poses considerable risks to cetaceans relying on sound for communication, navigation, and prey location. Beyond enabling remote marine mammal detections, these hydrophones offer the possibility to obtain in situ measurements of underwater noise levels. In this study, the HydroMoth recorder was deployed during fishing activities with FADs (Fish Agreggating Devices) and whale watching operations in the Western Mediterranean Sea. Cetacean detection capabilities of the hydrophone were assessed through testing with different recording configurations. The preliminary results showcase acoustic detections of one whale and four dolphin species. The findings underscore the effectiveness of the accessible hydrophones in capturing high-quality acoustic detections of a diverse range of marine mammal vocalizations. This capability provides an innovative approach for utilizing cost-effective recorders to support the monitoring of marine biodiversity and the associated anthropogenic threats. Engaging local fishermen, NGOs, and other organizations in citizen science projects using user-friendly recorders, is a valuable approach to involve citizens in marine conservation and expand passive acoustic monitoring of marine environments globally.

Evidence of Synchronized Calls of Likely Balaenoptera Musculus

D. Lechner[1*], S. Guan[2]

[1]*N/A*

[2]*The Catholic University of America*

**66lechner@cua.edu*

This presentation will review the analysis of approximately 300 calls made by a likely new subgroup of blue whales (Balaenoptera Musculus) in June 0f 2015 over a 3 day period, as detected by the Comprehensive Test Ban Treaty station's three hydrophones. The analysis will review the calls and their uniqueness from known subgroups and the apparent interaction shown by the data. It will also analyze the impacts of knowing the location and velocity of the callers, and the impacts on density estimation from this type of data. The location and velocity data is used to develop polar plots of the amplitude as a function of propagation angle, and support analysis of factors that could cause individual whales to take a turn making a (single) call during the ongoing sequence of calls. Variations in the amplitude are compared to the estimated velocity to test whether that might be an indication of fluke motion. Limitations of the analysis are also presented, with recommendations for corroboration of the data. Special appreciation is also noted to the former Mr. Thomas Norris, who provided insight and guidance of the initial analysis based on his field experience.

Free-ranging bearing joint to vocal analysis: application to whistles of short-beaked common dolphins

L.L. Lehnhoff[1,2*], H. Glotin[2], B. Mérigot[2]

[1]*MARBEC, Université de Montpellier, CNRS, IRD, Ifremer, Sète, France*

[2]*CIAN, Centre Intelligence Artificielle en Acoustique Naturelle, Toulon, France*

[3]*Université de Toulon, AMU Univ, CNRS, LIS, DYNI team, Marseille, France*

**loic.lehnhoff@gmail.com*

Passive acoustic monitoring methods have enabled the study of cetaceans sounds in both captive and free-ranging environments. Audio recordings often replace visual observations, so it is uncommon to have access at the same time to data related to cetaceans movements. To accurately study the whistles of free-ranging individuals, it is essential to first ensure that their recorded vocalisations are not varying based on their positioning. We investigate the variations of whistles emitted by free-ranging individuals by using estimations of their orientations deducted from their echolocation clicks.

We recorded wild short-beaked common dolphins (Delphinus delphis) off the coast of Brittany, France [1], using a small 4-hydrophone antenna (~1m wide) attached to a buoy and placed 3m below the sea surface. One-minute recordings with a sampling rate of 256kHz were made, 20 minutes of recordings containing common dolphin vocalisations were obtained. To determine bearings, we extracted dolphin echolocation clicks using a custom detector, then we determined the time delay of arrival (TDoA) of each click using a cross-correlation method. Finally, we estimated the direction of arrival (DoA) of each click knowing its TDoAs and the hydrophone's distances. In parallel, we also manually extracted the frequency contours of the whistles emitted during the experiments using a custom annotation tool.

We used the angular speed of dolphin echolocation click trains to determine their orientation towards the antenna. However, echolocation clicks are often inaudible, so dolphin orientation is sometimes unknown. We investigated the relation between the orientation of the dolphins and the features of the whistles recorded. Our results confirm studies [2] showing that the fundamental frequencies of dolphin whistles (10-15 kHz) are omnidirectional. Whistles emitted at frequencies >20 kHz should be studied with knowledge of the orientation and/or distance of the animal relative to the recording device. Detailed datasets with information on dolphin orientation and/or distance are needed to improve our understanding of dolphin communication in groups.

[1] Lehnhoff L, et al. Behavioural Responses of Common Dolphins Delphinus delphis to a Bio-Inspired Acoustic Device for Limiting Fishery By-Catch. Sustainability. 2022; 14, 13186. https://doi.org/10.3390/su142013186

[2] Lammers, Marc O. and Au, Whitlow W. L. “Directionality in the whistles of hawaiian spinner dolphins (Stenella longirostris): a signal feature to cue direction of movement?” Marine Mammal Science. 2003; 19, 249–264. https://doi.org/10.1111/j.1748-7692.2003.tb01107.x

Tools for the classification of small NBHF species in southern Chile

F. Malige[1*], J. Patris[1], M. Silvestri[2]

[1]*LIS laboratory, CNRS UMR 7020 LIS, DYNI team, France*

[2]*Austral university of Valdivia, Chile*

**franck.malige@lis-lab.fr*

Coastal odontocetes are resident species which are highly threatened by human activities such as fishing, harbour development or fish farms. In Southern Chile, four species of NBHF (Narrow Band High Frequency) dolphins and porpoise share this threat, in a coastal area that is of growing economical interest. Passive acoustic monitoring is an interesting tool to assess the presence of these species. However, their acoustic emission have been little studied and there is actually no means of acoustically classify the different species. We argue that actual parameters used to characterize the NBHF clicks are not precise enough and propose a new method of measuring the frequency and duration of a click, by fitting a mathematical function (a Gabor wavelet) to the click. In this work, we show some important biasses of classical NBHF clicks parameters, and discuss the interest of a Gabor wavelet model. Our method is applied on two dataset with two distinct species in southern Chile : the Peale dolphin (Lagenorhynchus australis) and the Chilean dolphin (Cephalorhynchus eutropia). The results show the striking inter-specific similarity of NBHF clicks, and point towards possible method to separate them.

Inter burst-pulse interval as a species indicator for Pacific delphinids

E.G. McCloskey[1*], V.M. Janik[1], S. Rankin[2], V.B. Deecke[3], J.N. Oswald[1]

[1]*Scottish Oceans Institute, University of St Andrews, Fife KY16 8LB, UK*

[2]*Southwest Fisheries Science Center, 8901 La Jolla Shores Drive, La Jolla, CA, USA*

[3]*Institute of Science and Environment, University of Cumbria, Ambleside, UK*

**egm25@st-andrews.ac.uk*

Autonomous recorders such as those in the Oceans Observatories Initiative (OOI) Coastal Endurance Array off the coast of Newport, Oregon, U.S.A. are commonly used to monitor marine mammals. The OOI recorders detect many sounds produced by marine mammals, including patterned sequences of burst pulses that do not also contain killer whale calls. Only Pacific white-sided dolphins (PWSD), Lagenorhynchus obliquidens, and northern right-whale dolphins (NRWD), Lissodelphis borealis, are known to produce such sequences in this area, however the OOI recordings lack visual observations to confirm the identity of vocalizing species. To investigate differences in burst pulse sequences produced by these species, we utilised dipping and towed hydrophone recordings with visual confirmation of species identity from both species to compare the median and inter-quartile range (IQR=25th-75th percentile) of five burst-pulse parameters: burst-pulse duration, inter burst-pulse interval (IBPI), number of clicks in a single burst-pulse unit, the number of burst pulse units in a sequence and sequence duration. We defined a burst-pulse sequence as two or more successive burst-pulse units produced within 758 milliseconds (ms) for PWSD and 537 ms for NRWD, determined using bout criterion interval (BCI) analysis. A burst-pulse unit was defined as containing at least 3 clicks with inter-click intervals of less than 9.9 ms for PWSD and 8.9 ms for NRWD, based on BCI. For PWSD, burst-pulse sequences were made up of 2-31 units and each unit was made up of 3-294 clicks. For NRWD, burst pulse sequences were made up of 5-22 units and each unit was made up of 3-537 clicks. We found that median IBPI for PWSD (42.8ms, IQR=94.6ms) was significantly shorter than median IBPI for NRWD (73.9 ms, IQR=65.5ms) ($W=21810$, $p=0.0004$, Mann-Whitney U test). The other four burst-pulse parameters did not differ significantly between species ($p>0.05$). Furthermore, there appears to be more variability in the structure of PWSD burst-pulse sequences compared to NRWD which were found to be more stereotyped and could be grouped into specific sequence types. We are currently carrying out a structural analysis of burst pulse sequences to quantify this in more detail. We suggest that IBPI and potentially burst pulse sequence structure can be used to aid species identification in the Pacific Northwest.

Combining visual and acoustic subgroup localization efforts to examine movement of false killer whales

J.L.K. McCullough[1*], P. Gruden[2], S. Fregosi[3], Y. Barkley[4], E. Oleson[1]

[1]*NOAA Fisheries: Pacific Islands Fisheries Science Center, Honolulu, USA*

[2]*University of Hawaii, Manoa, Honolulu, USA*

[3]*Ocean Associates, Inc. contracted to Pacific Islands Fisheries Science Center, Honolulu, USA*

[4]*Cooperative Institute for Marine & Atmospheric Research, Honolulu, USA*

**jennifer.mccullough@noaa.gov*

False killer whale (Pseudorca crassidens) groups are often organized into dispersed subgroups that can span dozens of miles and be visually cryptic near the surface. This subgroup behavior has been accounted for in analysis of visual line-transect surveys through treatment of the subgroup as the unit of detection. Integration of more numerous acoustic detections of subgroups may increase the accuracy and precision of the abundance estimate. In addition, it may provide a means for correcting for potential bias resulting from animals moving toward the trackline before they are detected by the visual team, a behavior that appears evident in the sighting data detection function. However, visual sightings and acoustic detections of the same subgroups rarely align in space and time with visual sightings captured ahead of the ship and acoustic localization finalized at the beam of the towed array (behind the ship). Seeking to overcome these issues, we have used towed hydrophone array acoustic recordings paired with visual survey data collected as part of the 2017 Hawaiian Islands Cetacean and Ecosystem Assessment Survey (HICEAS) (DCLDE 2022 dataset). Time-Difference-Of-Arrivals (TDOAs) for all false killer whale signals were automatically extracted and then aggregated into subgroups based on consistent TDOA trajectories using a multi-target Bayesian framework. Visual sighting time and bearing information was placed within the subgroup TDOA tracks to examine how well the modalities align in identifying subgroups. Perpendicular distance to the trackline was computed at the beam for acoustically-detected subgroups, allowing for comparison to the visual detection function. Further, distance at other points along the TDOA trajectory can be computed to see if segmentation can provide insight into responsive movement.

Dolphin whistle contour extraction in a noisy environment using the pyknogram representation.

R. Miralles[1,2*], C. Gallardo[1], E.J. Belda[2], G. Serrano[1], B. Feliu-Tena[2], G. Lara[2], J. Gosalbez[1], I. Bosch[1], A. Leon[5]

[1]*Institute of Telecommunications and Multimedia Applications (iTEAM), Universitat Politècnica de València (UPV), Camino de Vera S/N, 46022 Valencia, Spain*

[2]*Unidad Mixta de Investigación IEO-UPV, Tinglados Muelle Frutero, Grau de Gandia, 46370 Valencia, Spain*

[3]*Instituto de Inv. Para la Gestión Integrada de Zonas Costeras (IGIC), Universitat Politècnica de València (UPV), C/Paranimf 1, Grau de Gandía, 46730 Valencia, Spain*

[4]*Instituto Español de Oceanografía (IEO-CSIC), C.O. Murcia, C/el Varadero 1, Lo Pagan, 30740 Murcia, Spain*

[5]*Departamento de Comunicaciones, Universitat Politècnica de València (UPV), Camino de Vera S/N, 46022 Valencia, Spain*

**rmiralle@dcom.upv.es*

The maritime traffic of the Cartagena Port (Murcia, Spain) has exponentially increased in recent years. This is a significant concern for all authorities and organisms, and several efforts are being made to monitor the impact of the noise on different cetacean species. This study is being done as part of the LIFE PortSounds project" (LIFE20 ENV/ES/000387) that aims to implement mitigation measures to reduce the impact of marine traffic noise. The project studies the population abundance trends of three objective species considered as species of Community interest under the Habitats Directive (Council Directive 92/43/ECC): bottlenose dolphin (Tursiops truncatus), striped dolphin (Stenella coeruleoalba) and long-finned pilot whale (Globicephala melas). To study the noise impact on the abundance, and distribution of these species we must extract the whistle contours with the highest precision to classify the species by their sounds. The underwater ship noise present in the Cartagena harbor area can complicate the already difficult task of whistle contour extraction. Underwater noise usually breaks the extracted whistle contour into several fragments, which diminishes the identification of the objective species by their whistles.

In this work, we explore how to combine the whistle candidates obtained using the pyknogram representation with different whistle extraction algorithms. Pyknogram representation has been used many times for the extraction of formant in speech signals. It has proven to have better results than the spectrogram in the presence of noise or interfering signals. However, it produces fewer whistle candidates than the spectrogram, and thus, it does not work equally well with all whistle contour extraction algorithms. Different whistle extraction techniques: GM-PHD, and 2D curvilinear structures detector, will be used and tested using a manually annotated dataset from the LIFE PortSounds project. Various metrics will be obtained to compare the whistle contour fragmentation and accuracy when classifying the three objective species.

Deep Ocean Prey Mapping from bottom mounted bidirectional nodes

R. Morrissey[1*], A. Muniz[1], S. Watwood[1], K. Benoit-Bird[2]

[1]*Naval Undersea Warfare Center, 1176 Howell Street, Newport, RI, USA*

[2]*Monterey Bay Aquarium Research Institute, 7700 Sandholdt Road, Moss Landing, California 95039, USA*

**ronald.p.morrissey.civ@us.navy.mil*

In response to mid-frequency active sonar (MFAS), beaked whales appear to move from an area and return once the disturbance has diminished (McCarthy 2011). It is presumed these animals occupy areas of high prey concentration, but the relative distribution of prey at depths below 1,000m is little understood. This project evaluated the use of existing bottom-mounted transducers for mapping of relative prey availability over broad spatial and temporal scale by leveraging bidirectional nodes found at the Southern California Offshore Range (SCORE). Preliminary results and analysis are presented along with an overview of the system.

Synchronous aerial and acoustic surveys to estimate porpoise emission rate

J.P. Mura[1*], L.L. Oliveira[1], A. Andriolo[2]

[1]*Aqualie Institute - Brazil*

[2]*Zoology Department - Federal University of Juiz de Fora - Brazil*

**mura@aqualie.org*

Coastal porpoises face significant challenges, including high mortality rates, low population growth rates, and susceptibility to anthropogenic threats, potentially leading to the risk of extinction. Accurately estimating their distribution, abundance and population trend is paramount for effective species conservation. While traditional methods such as visual surveys have been used, passive acoustic monitoring (PAM) offers a promising alternative method for population studies. Using PAM for this purpose implies calculating vocalization emission rates to be considered as correction factors throughout the analysis. Here, we propose an approach to estimate porpoise emission rates using synchronous drones and PAM surveys. Understanding the habitat and behavior of the target species is crucial for adapting the tools for effective feasibility of the method. Several key pieces of information need to be integrated, including diving time, acoustic detection probability (estimated using the passive sonar equation), and visual range of detectability (estimated using Ground Sample Distance - GSD). Acoustic data processing involves automated click detection, click train identification, classification, and localization. The drone data was analyzed independently by manual visual group counting for estimation of perpendicular distances. An essential aspect consists in setting a threshold for acoustic detection probability, commonly established at >95%, and ensuring that the ratio between the acoustic detection range and boat speed remains below the species' diving time. Once all required data is collected, a step of the protocol is implemented to synchronize the detections by time. This step correlate drone sights with click trains counts to determine the acoustic rate accurately. This method was tested and applied for the first time for Pontoporia blainvillei, employing a small boat towing a hydrophone array. The results were used to correct acoustic density estimation of the species in one area in Brazil. This approach promises a more accurate assessment of population density and distribution, offering valuable insights for porpoise conservation efforts.

How can we improve acoustic classifier performance? A meta-analysis of acoustic species classifiers for odontocetes

J.N. Oswald[1*], T. Kleyn[1], H. Worthington[2], V.M. Janik[1]

[1]*Scottish Oceans Institute, School of Biology, University of St Andrews, UK*

[2]*Centre for Research into Ecological and Environmental Monitoring, University of St Andrews, UK*

**jno@st-andrews.ac.uk*

Acoustic species classifiers are important tools for monitoring and managing cetacean populations and this has led to a significant increase in their development over the last 20 years. Overall performance of these classifiers has been reported individually but what factors influence classifier accuracy is often unclear. We conducted a literature study on acoustic classifiers for odontocetes worldwide to investigate the influence of selected parameters such as the number of species targeted, geographic region, sample size of the training data set, classification model type, manual or automatic acoustic feature extraction method, classification of single vocalizations or groups of vocalizations, and the type of vocalization used for classification. We found 95 classifiers published with sufficient information to be included in our meta-analysis. We modelled accuracy of classifiers against a classical discriminant function classifier in the Atlantic using a binomial, generalized linear mixed model (GLMM). GLMMs fit with different combinations of variables and interactions between them were compared using analysis of variance to assess the impact of their inclusion in the model and to determine the best model.

Our model provided the following significant results. Classifiers that classified groups of vocalizations (e.g. all found in one encounter) showed higher accuracy than those classifying each individual vocalization ($p<0.001$). At small training data sample sizes, click classifiers had higher accuracy than whistle classifiers ($p<0.001$), while this effect disappeared at higher sample sizes. Using automated acoustic feature extraction improved accuracy over manual extraction ($p<0.05$), but it is possible that this is driven by clicks. Using independent test data for accuracy assessment decreased accuracy ($p<0.05$). This last result indicates that using all or parts of the training data in accuracy assessment leads to an overestimation of classifier accuracy. Number of species targeted for classification and the geographic area for the classifier did not influence classification accuracy. There was also no significant accuracy difference between the most common classification models used (discriminant function analysis, decision tree analysis, neural network, and gaussian mixture models), but higher accuracy was found with rarer methods including more recent machine learning algorithms. Our analysis demonstrates the importance of considering classifier design and highlights the role of new methods in classification models.

A new deep learning model evaluated on the Antarctic benchmark for baleen whale calls

C. Parcerisas[1*], I. Kaya[2], E. Debusschere[1], P. Devos[3], E. Schall[4]

[1]*Flanders Marine Institute (VLIZ), InnovOcean Campus, Jacobsenstraat 1, 8400 Oostende, Belgium*

[2]*Middle East Technical University, Üniversiteler Mahallesi, Dumlupınar Bulvarı No:1 06800 Çankaya Ankara, Turkey*

[3]*Ghent University, WAVES, Technologiepark-Zwijnaarde 126, 9052 Gent, Belgium*

[4]*Alfred-Wegener-Institute Helmhotz Center for Polar and Marine Research (AWI), Klußmannstr. 3d, 27570 Bremerhaven, Germany*

**clea.parcerisas@vliz.be*

Passive acoustic monitoring (PAM) is commonly used to continuously monitor marine habitats by providing valuable information on, for example, species distributions or ecosystem dynamics. This is often achieved by detecting and classifying calls from species of interest, but the unceasingly increasing amount of data makes it impossible for scientists to annotate it manually. Therefore, highly efficient automatic analysis techniques which limit human input to the minimum are necessary in order to exploit the full potential of the available data.

In the last decades several automatic algorithms have been developed to this effect, but their performances cannot be compared because of the differences in data used for training and testing. Here, we evaluate a new model in a benchmark which includes a public dataset, a well-defined task and evaluation procedure to develop and test automatic analysis techniques. This benchmark focuses on the special case of detecting baleen whale vocalizations in a real-world dataset from the Antarctic. We believe that such a benchmark is necessary to monitor the progress in the development of new detection algorithms in the field of marine bioacoustics.

The benchmark proposes to report the performance in a blocked cross-validation fashion with 11 site-year blocks for a multi-species detection scenario in a large marine passive acoustic dataset, to simulate how would the algorithm perform when applied to data obtained from a new deployment. Performance is measured with three simple metrics (i.e., true classification rate, noise misclassification rate, and call misclassification rate) and one combined fitness metric, which allocates more weight to the minimization of false positives created by noise. This decision is done to limit manual revision.

To the models reported in the benchmark we add a new state-of-the-art model which uses transfer learning from computer vision to detect objects in spectrograms. The proposed model is a further step to detect individual and simultaneous calls, and is added as a state-of-the-art algorithm to the benchmark. The presented benchmark is an important step to advance in the automatic processing of the continuously growing amount of PAM data which is collected throughout the world's oceans. To ultimately achieve usability of developed algorithms, the focus of future work should be laid on the reduction of the false positives created by noise.

A GAM-based classification of ranges of fin whale calls obtained from single seismic sensors

A. Pereira[1*], R. Hilmo[2], D. Harris[2], W. Wilcock[3], L. Matias[1]

[1]*Instituto Dom Luiz, University of Lisbon, Portugal*

[2]*School of Oceanography, University of Washington, US*

[3]*Centre for Research into Ecological and Environmental Modelling, University of St. Andrews, St. Andrews, Scotland, U.K.*

**afpereira@fc.ul.pt*

Ocean-bottom seismometers (OBS) have increasingly become common instruments for acquiring spatial information on baleen whales, including tracking movements, investigating habitat use, behavioural responses to sound sources, and animal density estimation. These instruments, depending on their configuration and propagation conditions, allow estimating ranges to calling animals, i.e. ranging, through various methodologies, including single-station techniques. Matias & Harris (2005) expanded on a seismic ranging method that utilized ground motion recorded by the three orthogonal channels of a single OBS to derive the range and azimuth of fin whale calls, the three-component method (3C). While the method provides range estimates robust enough for animal density estimates, it is limited to direct path signals within a validity range, which is linked to instrument depth and propagation conditions. Consequently, the method requires a classification process for ranges resulting from external sound sources. In the "Combining global OBS and CTBTO recordings to estimate abundance and density of fin and blue whales", or CORTADO project, fin whale call data from six deployment areas - Marianas trench, Hawaii, Oregon OOI, ENAM, Azores and Gulf of Cadiz - were used to apply the 3C and develop a classification process using Generalized Additive Models (GAM). The classification model, developed using 70% of the standardized data, incorporated a comprehensive set of parameters related to channel amplitude, signal quality, polarity, and estimated angles, represented with average values and measures of dispersion. It successfully captured 72.8% of the deviance of the data, with major contributions from the average and dispersion of correlation values and average values of vertical and horizontal channel amplitudes, and polarity parameters. Validation on the remaining 30% of the standardized data and a five-fold cross-validation on the training dataset showed the high robust discriminative capability of the model, resulting in an Area-Under-the-Curve (AUC) scores of 97.3% (CI% 96.0%-98.5%) and 96.6% (CI% 96.0-97.7%), respectively. Applying the model to an external dataset of fin whale calls from Alaska validated its strong predictive performance, achieving a 91% accuracy rate in classifying ranges. The requirement for an extensive array of explanatory variables, encompassing both average values and measures of variability, underscores the intricate nature of classifying 3C ranges.

Observations Regarding Pile Driving Noise Measurements on a Towed Hydrophone Array

V. Premus[1*], P. Abbot[1], J. Browning[1], T. Abbot[1], V. Kmelnitsky[1], A. Logan[1], J. Freise[1]

[1]*OASIS, Inc., a wholly owned subsidiary of ThayerMahan, Inc. Lowell, MA USA*

**premus@oasislex.com, vpremus@thayermahan.com*

An analysis of towed hydrophone array measurements of pile driving radiated noise collected on 1 August 2023 in the Martha's Vineyard offshore wind lease area is presented. Data was collected on an uncrewed surface vehicle (USV) instrumented with a 32-channel towed hydrophone array. In-band (i.e., 50-250 Hz) omni-directional hydrophone noise levels measured during the strike were observed to be as high as 135-140 dB re 1Pa at a range of 5 NM. Corrected to spectrum level, this is roughly 30 dB higher than historical Wenz heavy shipping. Array beamforming with a 32-channel towed array is shown to deliver array gain, or spatial noise rejection, of as much as 25-30 dB in the North Atlantic right whale (NARW) upcall frequency band in this highly anisotropic noise environment. To assess the relative impact of pile driving noise on NARW upcall classification using single hydrophone and beamformed acoustic measurements, an audio wavefile comprised of actual pre-recorded upcalls was coherently injected into the array element data to mimic the presence of right whales—NARW were not present during the at-sea recording event. The results demonstrate that the beamformer and a spectrogram-correlator based autodetector were able to detect and classify the injected upcall at an element-level signal to noise ratio (SNR) for which the single hydrophone could not. In this talk we will report on observations regarding noise level, sound exposure level, array gain, beam response, and the bearing-dependent impact of pile driving noise on array and single hydrophone detector-classifier performance.

Preliminary investigation of odontocete acoustics in French Polynesia

X. Raick[1*], E. Cabanne[1], E. Parmentier[1]

[1]*University of Liège, Belgium*

**xavier.raick@uliege.be*

Acoustic communication plays a crucial role for species such as odontocetes. These marine mammals emit distinct whistles, making their sounds valuable for assessing abundance and distribution. Despite this, no study in French Polynesia has investigated odontocete distribution using passive acoustic surveys. This study aims to evaluate odontocete species presence in French Polynesia by employing classifiers developed for two other geographical regions. The objective is to provide an initial assessment of these classifiers and to document any geographical or diurnal patterns. From March 2018 to April 2019, a total of 4094 whistles were recorded. Results revealed an abundance of Tursiops truncatus in the Tuamotu Archipelago, a prevalence of the genus Stenella in the Society Archipelago, and a dominance of whistling in the morning and evening. These findings align with existing literature, indicating differences between the two archipelagos, a north-to-south gradient decrease in the Tuamotu, and a daytime period of resting activity, consistent with observed acoustic patterns.

Glider-based passive acoustic monitoring of marine mammals

D. Risch[1*], E. Edwards[2], S. Beck[3]

[1]*Scottish Association for Marine Science (SAMS)*

[2]*XODUS*

[3]*Agri-Food and Bioscience Institute (AFBI)*

**denise.risch@sams.ac.uk*

Recent technological advances have resulted in the availability of self-contained, autonomous passive acoustic monitoring (PAM) devices, which can be attached to ocean gliders for multi-week deployments. It has been shown that gliders can be used effectively for marine mammal and underwater noise monitoring. Glider-based acoustics also enable 3D tracking of deep-diving species to derive information about their feeding behaviour and may be useful in estimating species abundance. The integration of such data with habitat specific parameters collected by ocean gliders will substantially improve our knowledge on important marine mammal habitat and potential threats such as underwater noise. Autonomous, mobile PAM has the advantage of collecting data across a wider spatial area, which could complement static data collection at a limited number of sites.

The INTERREG VA COMPASS project implemented a static passive acoustic monitoring (PAM) array to investigate the seasonal distribution of marine mammals and measure ambient sound at ten sites between Scotland and the island of Ireland. As part of this effort, two short-term (one month) Seaglider missions were equipped with a Soundtrap 300 HF and run in proximity to the existing static PAM array, to compare glider-based detection results with those obtained by stationary recorders.

The glider tracks were carried out in spring and autumn and covered the Islay Front and circled back north past Stanton Banks, a series of granite rises which outcrop from the seafloor south of the Outer Hebrides, and are known for their species diversity and relatively high abundance of cetaceans.

Despite a lot of glider related broadband noise in the recordings, many instances of porpoise and delphinid detections were found to be reliable detectable during both glider missions. In both the spring and autumn deployments, harbour porpoises were detected on every day of the deployment. Delphinids were also detected regularly during both deployments and most often when the glider was surveying close to Stanton Banks.

The additional payload added to the glider due to the attachment of the Soundtrap made the glider harder to pilot in high current areas and also reduced the overall battery life of the mission. This is a common drawback in studies aiming to combine PAM and oceanographic data collection. For long endurance missions the addition of PAM to traditional oceanographic sampling will have to be carefully considered and weighed against mission length and duration.

Communication of killer whales engaging in carousel feeding recorded with a large baseline hydrophone array.

J. Rychen[1*], A. Berg[1]

[1]*Department of Information Technology and Electrical Engineering, ETH Zurich, 8092 Zurich, Switzerland*

**jrychen@ethz.ch*

We used an array of four drifting hydrophone recorders that we deployed in a square of about 500 m side length around Norwegian killer whales that engaged in carousel feeding or subsequent socializing. These activities are well suited to study complex communication, because the animals need to cooperate as a group, vocalize very actively, and stay quite stationary and close to the surface. We used drones to simultaneously film the action from above with a view of about 10 m into the water. We designed the instruments and procedures in a way that allows to localize sounds and map them with the videos to assign them to individual animals. The special built hydrophone recorders are synchronized to the GPS clock with sub sample precision (100 ns) and the drone videos are synchronized to frame accuracy (20 ms). Each recorder has four hydrophones arranged in a tetrahedral structure suspended 10-20 m below the surface which allow for 3d sound field reconstruction. Our documented data management process yielded consolidated files with 16 hydrophone channels with calibrated sound pressure measurements and all relevant data like GPS tracks, tidal levels, temperatures and much more. Additionally we measured the vertical sound speed profile for every deployment to enhance the precision of localization algorithms. In the winter months Nov. 2023 till Jan. 2024 we deployed the array more than 100 times Northern Norway. We are in the process of publishing this dataset and are seeking collaborations to analyze the communication of killer whales to find out how they are coordinating the carousel of herrings.

Glider and Whales: using acoustic glider to monitor marine mammals

F. Samaran[1*], J. Béesau[1], M. Dupont[1], M. Torterotot[1]

[1]*ENSTA Bretagne, UMR CNRS 6285 Lab-STICC*

**flore.samaran@ensta-bretagne.fr*

Started in 2017, the project Gilder and Whales aims to demonstrate the opportunity provided by the acoustic glider technology to monitor the biodiversity and especially cetacean species in their environment. This mobile and autonomous device can move toward location of interest in remote area to record ocean soundscape including sounds produce by cetaceans. For this project, a Sea Explorer (ALSEAMAR) equipped with one hydrophone as a payload and different sensors for positioning and data transmission/reception has been deployed to collect acoustic data through 3 missions in 2 different oceans. Acoustic data were manually visualized for cetacean presence using the annotation platform APLOSE (ENSTA Bretagne) over two frequency bandwidths (0-500 Hz and 0-48kHz) to encompass the full spectra of 10-minute time windows.

In May 2017, the glider has made its first mission in the south of Madagascar southeast of Walters Shoal. During 10 days, he travelled for nearly 200 km and collected numerous recordings with a sampling frequency of 32 kHz. Marine mammal acoustic presence occurred in more than 40% of 10-minute time windows. Two blue whale acoustic population's songs, fin whales 20 Hz calls and sperm whales' clicks were detected. The next deployment was around St Paul & Amsterdam islands from February 28 to April 5 2019 with a sampling frequency of 48kHz. Marine mammal acoustic presence occurred in more than 70% of 10-minute time windows along the glider route. Here, three blue whale acoustic population's songs, spot or P-call (an unattributed baleen whale call), as well as some D-calls were detected and sperm whales' clicks were also highly present in the dataset. We observed significant presence of blackfish around Amsterdam Island. The last deployment was in the Bay of Biscay (Atlantic Ocean) in June 2023. The sampling frequency was 96 kHz. Marine mammal acoustic presence was detected in more than 90% of 10-minute time windows. We detected at least four cetacean species, mainly odontocetes. A next step would be to confront the acoustic detections regarding the depth of the glider to better understand the effect of the sound propagation on the received calls in the water column.

Using an acoustic glider allow to explore remote areas and complete acoustic monitoring collected by fixed stations. It also seems to be an accurate alternative to explore an area before the deployment of a long-term acoustic observatory.

CAB Guardian: Detections and Bearings from a 3 Month Deployment

S. Tabbutt[1*], P. King[1], H. Yurk[2], J. Wood[1]

[1]*SMRU Consulting, Friday Harbor, USA*

[2]*DFO/MPO, Vancouver, Canada*

**st@smruconsulting.com*

The Coastal Acoustic Buoy (CAB) is an autonomous PAM platform with real-time noise measurement and cetacean detection capabilities. Equipped with three HTI-96 Min hydrophones separated by 2 meters, the CAB Guardian calculates bearings to low- and mid-frequency cetaceans using the Pamguard Whistle and Moan detector. With cell communications, Pamguard data files are transmitted to a server at 20-minute intervals, and, in the event of at least one detection, a sample recording of the 20-minute period with the densest minute of detections is sent.

The system was configured for a 3-month deployment period (June – August 2022) on the outer coast of Vancouver Island, British Columbia, Canada in conjunction with Fisheries and Oceans Canada for technology testing and marine mammal monitoring.

Humpback and killer whales were detected throughout the deployment. During an agency response to the sinking of a fishing vessel ~130 km away, the CAB Guardian provided detections and bearings (to infer direction-of-motion) to endangered Southern Resident killer whales to the response team for situational awareness and response management. We will provide results from the 3-month deployment, including the effectiveness of direction-of-motion estimation for various species, and an analysis of the 539 distinct detection events that occurred over the deployment.

North Atlantic Right Whale (NARW) Line-Array Beamformer Performance using Energy Detection and Pitch-Tracking Classification Metrics

J.A. Theriault[1,2*], T.J. Leblanc[1], J. Xu[3], G. Macklin[3], D. Flogeras[2], L. Theriault[2]

[1]*Theriault Environmental Consulting*

[2]*Ocean Environmental Consulting*

[3]*Fisheries and Oceans Canada*

**jim@ocean-env.com*

A vertical line array (VLA) of hydrophones was deployed by Fisheries and Oceans Canada in Roseway Basin (Canadian Scotian Shelf) in 2020. The recordings of each of the eight hydrophones have been processed using the Baumgartner LFDCS (low frequency detection and classification system) to detect North Atlantic Right Whale (NARW) upcalls and verified by an operator. Two approaches are used to beamform the array. The first approach is time-domain conventional beamformer and the second is a Minimal Variance Distortionless Response (MVDR) beamformer. This study allocated 30 % of the verified LFDCS detections with low Mahalanobis Distance (MD<1) to train a similar pitch-tracking classifier. The remaining raw time-series data with all LFDCS detections (with MD < 5) are considered for further processing. Individual hydrophone and beamformed time series are processed with a band-limited energy detection and a previously 2018 DCLDE presented pitch-tracking classifier. The detector and classifier are used for comparison metrics. The metric comparison based on the hydrophone channels, the steered conventional beamformer, and MVDR beamformer are presented.

Right whale call density estimates from the workshop dataset via spatial capture-recapture

L. Thomas[1*]

[1]*Centre for Research into Ecological and Environmental Modelling, University of St. Andrews, St. Andrews, Scotland, UK*

**len.thomas@st-andrews.ac.uk*

The 2024 workshop dataset comprises two days of recordings from a 4 x 8 grid of sonobuoys with approximately 8km spacing, deployed in the Bay of Funday, Canada. Manually annotated data were provided for people like me who are not capable of running detectors and classifiers. By making simplistic assumptions about reasonable time differences of arrival between sensors I associated annotated right whale detections across sonobuoys that were likely the same call. I used this dataset of associated calls as the input to a spatial capture recapture (SCR) analysis that estimated spatial density of calls (accounting for false negatives) in the area within and around the array. Results illustrate how poorly things go when statisticians are allowed to do things by themselves, and emphasize the importance of inter-disciplinary teamwork in achieving optimal outcomes. For example, it would be much better to use call characteristics to help with the association task, and to be able to use in the SCR analysis the bearing information that could be extracted from analysis of the DIFAR buoy wavfiles.

Localization-Derived Acoustic Detection Function for Cuvier's Beaked Whales Offshore Guam

I. Tolkova[1*], D. Harris[2], T. Yack[3], D.P. Salisbury[1], H. Klinck[4]

[1]*K. Lisa Yang Center for Conservation Bioacoustics, Cornell Lab of Ornithology, Cornell University*

[2]*Centre for Research into Ecological and Environmental Modelling, University of St Andrews*

[3]*EcoSound Bioacoustics, LLC*

[4]*K. Lisa Yang Center for Conservation Bioacoustics, Cornell Lab of Ornithology, Cornell University & Marine Mammal Institute, Department of Fisheries, Wildlife, and Conservation Sciences, Oregon State University*

**itolkova@cornell.edu*

Beaked whales (Ziphiidae) are an elusive family of cetaceans comprising over 20 species globally. Due to their deep-diving behavior and deep-sea habitat, comparatively little is known about their ecology, abundance, and distribution. In particular, while beaked whales are known to be sensitive to anthropogenic noise, including mid-frequency active (MFA) sonar signals, the extent of these impacts on beaked whales is not well understood. Though difficult to observe visually, they regularly emit echolocation clicks for navigation, foraging, and communication, thereby motivating the use of passive acoustic monitoring (PAM). Our work addressed a data gap in studies of beaked whale abundance offshore Guam, a U.S. Island territory within the Mariana Islands of the Western North Pacific, where the U.S. Navy conducts training events using MFA sonar.

With this aim, we calculated an acoustic detection function for Cuvier's beaked whales through acoustic localization of echolocation clicks to enable density estimation using distance sampling in future work. An array of five Rockhopper acoustic recording devices spanning an area of 1 square km was deployed from June to November of 2022, recording continuously at a sampling rate of 197 kHz. Cuvier's beaked whale clicks were automatically detected throughout the six-month survey using PAMGuard software, followed by manual verification. After synchronization of the five units with a pinger system, localization was performed with a custom implementation of a time-difference-of-arrival (TDoA) likelihood surface algorithm. High-confidence localizations were then used to obtain the spatial distribution of Cuvier's beaked whales and derive a corresponding detection function for the Rockhopper hydrophone array. Additionally, we will discuss the advantages and drawbacks of this analysis pipeline and, more broadly, specific opportunities and challenges within large-scale automated synchronization and localization for marine bioacoustics.

CETIROISE : a cetacean passive acoustic observatory in a French Marine Natural Park

M. Torterotot[1*], J. Beesau[1], C. Gicquel[2], F. Samaran[1]

[1]*Lab-STICC UMR 6285, ENSTA Bretagne, Brest, France*

[2]*French biodiversity agency, Iroise marine park, Le Conquet, France*

**maelle.torterotot@ensta-bretagne.fr*

The CETIROISE project (funded under the French economic recovery plan) aimed at monitoring cetacean presence in the Iroise sea (Britanny, France) using passive acoustic monitoring. Targeted species were porpoise, delphinids (particularly common and bottlenose dolphins) and large baleen whales. From May 2022 until May 2023, 7 sites were equipped in the northern sector of the Iroise Marine Natural Park (PNMI). Two coastal moorings (20m depth) and five deepest moorings (100m depth) were each equipped with two recorders: the F-POD (Chelonia Limited), an echolocation-click logging device, and the Sylence LP (RTSys), a broadband recorder sampling at 128 kHz.

Delphinid whistles were detected with PAMGuard's whistle and moan detector while delphinid and porpoise clicks were directly assessed from the FPOD detections. Manual annotation was performed on the web-based annotation platform APLOSE. The annotations were used to assess delphinid clicks and whistles detection performances and also to monitor the low frequency baleen whale calls on the noisy low frequency band. Results highlight the continuous yearly presence of delphinids, especially at the farthest sites from the coast. But these results also point out the need to develop tools to be able to discriminate delphinids species from their sounds. Porpoise were recorded to a lesser extent, but they also appear to dwell in this region during most of the year. Among baleen whales, only a few minke whale calls were found in the data during the first analysed months.

These findings supplement the picture drawn by other observation technics and support the development of the PNMI conservation plan.

Classification performance in different signal types across odontocete's mixed species groups

Y. Viana[1,2,3*], T. Amorim[2,3], A. Andriolo[2,3]

[1]*Laboratório de Ecologia Comportamental e Bioacústica - LABEC, Departamento de Zoologia, Instituto de Ciências Biológicas, Universidade Federal de Juiz de Fora, Rua José Lourenço Kelmer, s/n, 36036-900, Juiz de Fora, Minas Gerais, Brazil*

[2] *Instituto Aqualie, Centro Comercial São Pedro - Av José Lourenço Kelmer, 1300 Salas 110, 122 e 114 - São Pedro, Juiz de Fora - MG, 36036-330, Brazil*

[3]

**yasminviana@aqualie.org*

Acoustic classification among cetaceans mixed-species groups (MSGs) remains a challenge given the difficulty of addressing which vocalization belongs to which species. Single species groups (SSGs) and MSGs of Stenella attenuata and Peponocephala electra were recorded with a 300m AUSET® hydrophone towed array (high-pass filter of 499Hz; operating frequency 100Hz-500kHz) and a SailDaq digitizer board (sampling frequency 500kHz/16bits), from 2017 to 2022, in the Western South Atlantic Ocean (04º to 22º S) totalizing nine distinct encounters. All data used in this study constituted recordings with simultaneous sightings. To verify whether the three groups of data (P. electra SSGs; P. electra + S. attenuata MSGs; S. attenuata SSGs) could be classified based on their whistles and clicks, a random forest analysis (RFA) was performed. The parameters low frequency, high frequency, delta frequency (i.e. the difference between high and low frequencies), center frequency, beginning and ending frequencies, and duration were extracted from whistles and peak frequency, 3 dB bandwidth, 10 dB bandwidth, and inter-click interval were extracted from clicks. For the models, 70% and 60% of the data were randomly selected as training sample, for whistles and clicks, respectively. Overall, the global accuracy was 60.5% for clicks and 56.6% for whistles. The MSGs clicks presented a low percentage of misclassification with the SSGs (12% with P. electra and 22% with S. attenuata SSGs) and a higher percentage of correct classification (66%) than whistles (39% of correct classification and 27% of misclassification with P. electra and 34% with S. attenuata SSGs). The 66% of correctly classified clicks produced during MSG present distinct characteristics that allow them to be separated by the SSGs clicks, making it feasible to hypothesize an acoustic modulation produced by one or both species in the MSG. The whistles similarity between mixed and single species groups may indicate that other factors may influence the used signal and modulation types, such as the nature of the mixed group, the behavior, and species composition. This study showed differences in the classification performance depending on the type of acoustic signal used to classify single and mixed species contexts. However, we do not know whether interaction enables acoustic changes to make some parts of mixed species signals unique. The results from this study shed light on future acoustic classification among sympatric species that share similar traits in their ecology.

Evaluating Temporal and Spatial Variability in North Atlantic Right Whale Upcall Detection and Classification Performance Over a 12-Year Period using PAM Guard Software

T.M. Yack[1*], A. Martinez[1], R. Schick[1,2], D. Nowacek[3], D. Gillespie[4], P. Tyack[4], C. Clark[5], S. Mayo[6]

[1]*Duke University, Marine Geospatial Ecology Lab, Nicholas School of the Environment, Duke University, Durham, NC USA 27708*

[2]*Southall Environmental Associates, 9099 Soquel Drive, Aptos, CA 95003*

[3]*Duke University, Nicholas School of the Environment, 135 Duke Marine Lab Rd., Beaufort, NC USA 28516*

[4]*Sea Mammal Research Unit, Scottish Oceans Institute, University of St Andrews East Sands, University of St Andrews, KY16 8LB, Scotland, Fife, United Kingdom*

[5]*K. Lisa Yang Center for Conservation Bioacoustics, Cornell University, 159 Sapsucker Woods Road, Ithaca, NY 14850, USA*

[6]*Center for Coastal Studies, 5 Holway Avenue, Provincetown, MA USA 02657*

**tina.yack@duke.edu*

North Atlantic Right Whales (NARWs) produce upcalls to communicate with conspecifics. Upcalls are produced by all age and sex classes, and upcall detection is typically used as a proxy for NARW presence. Automated methods are often used to detect and classify upcalls in long-term datasets. However, temporal and spatial variability in algorithm reliability due to environmental changes and presence of confounding sounds, is often overlooked. This variability in detection and classification performance can affect the interpretation of results and have significant impacts on subsequent density and abundance estimation. This study focuses on automated detection and classification of NARW upcalls using an edge detector and neural net classifier (Le Net) in PAMGuard software. The dataset consists of 301 days of recordings collected by Cornell on distributed arrays of 5-11 Marine Autonomous Recording Units (MARUs) in Cape Cod Bay spanning 12 winter-spring seasons (2007-2018). Upcall detections were previously validated on a subset of data using 17-days of recordings from two winter-spring seasons (2008-2009), for which all edge detections (101,468) were validated as true or false positives. The validation showed a precision of 0.72 and a recall of 0.89. Notably, most classification scores for validated upcalls exceeded 80%. However, there was a great deal of variation in precision across days and between seasons. Given the important ramifications of this variation on any resulting study, we expanded this effort to further investigate the factors that affected detection and classification performance in the entire 12-year dataset. To assess and validate detection and classification performance, we randomly selected a subset of data from each season to validate. Each selected 5 second period was manually assessed by a trained analyst to determine true or false positives and

negatives. The validation data were then used to calculate precision and recall of the edge detector and generate Receiver Operating Characteristic (ROC) curves. Classification scores were also evaluated against the true and false positive upcall validations. The results of the detector and classifier performance will be presented and discussed in relation to spatial and temporal dynamics, including ambient noise, environmental conditions, and the daily presence of humpback whales (from aerial surveys). We aim to explore factors that contributed most to detection and classification performance. This comprehensive analysis will enhance our understanding of the robustness of this detection and classification method under varying conditions. This work is expected to provide insights into how detection and classification performance influences the interpretation of results from long-term datasets.